Available online at www.sciencedirect.com

Resuscitation Plus

journal homepage: www.elsevier.com/locate/resuscitation-plus

AI and machine learning in resuscitation: Ongoing research, new concepts, and key challenges

Yohei Okada [a,b,*], *Mayli Mertens* [c,d], *Nan Liu* [a], *Sean Shao Wei Lam* [a], *Marcus Eng Hock Ong* [a,e]

Abstract

Aim: Artificial intelligence (AI) and machine learning (ML) are important areas of computer science that have recently attracted attention for their application to medicine. However, as techniques continue to advance and become more complex, it is increasingly challenging for clinicians to stay abreast of the latest research. This overview aims to translate research concepts and potential concerns to healthcare professionals interested in applying AI and ML to resuscitation research but who are not experts in the field.

Main text: We present various research including prediction models using structured and unstructured data, exploring treatment heterogeneity, reinforcement learning, language processing, and large-scale language models. These studies potentially offer valuable insights for optimizing treatment strategies and clinical workflows. However, implementing AI and ML in clinical settings presents its own set of challenges. The availability of high-quality and reliable data is crucial for developing accurate ML models. A rigorous validation process and the integration of ML into clinical practice is essential for practical implementation. We furthermore highlight the potential risks associated with self-fulfilling prophecies and feedback loops, emphasizing the importance of transparency, interpretability, and trustworthiness in AI and ML models. These issues need to be addressed in order to establish reliable and trustworthy AI and ML models.

Conclusion: In this article, we overview concepts and examples of AI and ML research in the resuscitation field. Moving forward, appropriate understanding of ML and collaboration with relevant experts will be essential for researchers and clinicians to overcome the challenges and harness the full potential of AI and ML in resuscitation.

Keywords: Prediction model, Natural language processing, Heterogeneity, Self-fulfilling prophecy, Feedback loop, Large language model, Emergency medicine

Introduction

Artificial intelligence (AI) and Machine learning (ML) are important areas of computer science that have recently attracted attention for their combined application to medicine. AI refers to technology in which computer systems have the ability to think and learn like humans and to automatically perform tasks that humans would normally perform such as cognition driven decision-making.[1] ML is used to develop algorithms and models that can learn from and make predictions or recommend decisions based on large datasets.[1] In resuscitation medicine, AI and ML hold the potential to revolutionize patient care by providing decision support and optimizing treatment strategies. However, as techniques continue to advance and become more complex, it is increasingly challenging for clinicians to stay abreast of the latest research involving AI and ML techniques in the resuscitation field.

This review aims to introduce recent AI and ML research to healthcare professionals interested in applying ML to resuscitation research but who are not experts in the field. We reviewed the relevant literatures searched as described in the Supplementary file to introduce prediction models, natural language processing (including large language models, LLM), consideration of treatment heterogeneity, and optimization of medical practice and resource management by reinforcement learning. We also discuss the limitations and challenges of implementing AI and ML tools in actual clinical settings.

* Corresponding author at: Health Services and Systems Research, Duke-NUS Medical School, National University of Singapore, Singapore.
E-mail addresses: yohei_ok@duke-nus.edu.sg (Y. Okada), mayli.mertens@uantwerpen.be (M. Mertens), liu.nan@duke-nus.edu.sg (N. Liu), gmslasws@nus.edu.sg (S.S.W. Lam), marcus.ong@duke-nus.edu.sg (M.E.H. Ong).
https://doi.org/10.1016/j.resplu.2023.100435

We aim to facilitate discussion on the potential for further research and enhance communication between clinicians, resuscitation researchers and AI and ML experts.

Prediction models

The most common use of ML is predictive modeling.[1] Prediction models (also known as supervised learning) are commonly used to predict a patient's diagnosis or outcomes based on clinical data. For example, ML models can be helpful to diagnose, estimate the severity in triage, and understand the risk of complications in decision-making for surgery, which can allow us to develop more appropriate treatment plans and potentially improve patient prognosis in a more objective manner.[2–4] This type of prediction model may also be applied to adjust for severity when considering the quality of care and assuming the counterfactual scenario (such as if a certain treatment was not performed with the resulting outcome) when discussing causal inference.[5,6]

Prediction models may incorporate a wide array of data including structured data such as demographic information, clinical variables, biomarkers, and blood test results, and also unstructured data such as images and bio-signals like electrocardiograms (EEG) and electroencephalograms (EEG), to predict outcomes (Fig. 1). We introduce some examples of research on prediction models based on the type of data.

Prediction models using structured data

Structured data is one of the most common sources of data for ML models in resuscitation research.[7,8] This type of data is typically presented in a tabular format with clear rows and columns, representing patients and their respective features or attributes. These may include demographic information, medical history, vital signs, laboratory test results, and more. For out-of-hospital cardiac arrest (OHCA) research, the Utstein format is established worldwide as a standardized data format. This enables us to easily develop ML models applied to the data.[9,10] One of the primary uses of ML with tabular data in resuscitation research are predictive models to estimate the likelihood of outcomes such as return of spontaneous circulation (ROSC), survival, or neurological recovery after cardiac arrest.[7,8,11] In another example, tabular data was also used to develop early warning systems (EWS) that predict the risk of cardiac arrest or other serious adverse events among patients admitted to hospital.[12–14] These systems use ML models to analyze various data such as heart rate, blood pressure, respiratory rate, oxygen saturation, temperature, and laboratory data, to identify patterns that may suggest a patient's condition is deteriorating. As a result, EWS can alert healthcare providers to intervene before a cardiac arrest occurs.[14] Further, these predictions are also valuable to estimate demand for bed capacity and to appropriately allocate medical resources.[15] Some of these ML models are implemented in electronic medical record systems or as applications on tablets or smartphones, which automatically input the data into the model and output the calculated results, improving user availability and accessibility.[14,15]

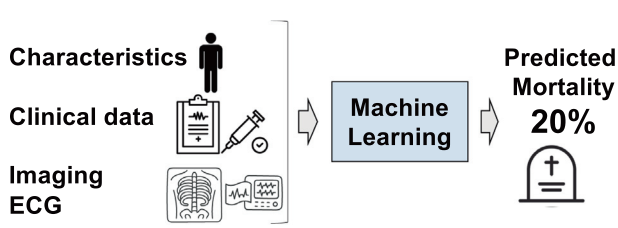

Fig. 1 – The concept of prediction models applied to predict mortality A prediction model is one type of ML developed to predict the outcome. Various patterns of clinical information can be utilized to develop prediction models.

Prediction models using unstructured data

Images and bio-signals (EEG, and ECG)

ML has been increasingly utilized in resuscitation research to enhance diagnostic and prognostic accuracy in unstructured data such as various imaging modalities, including CT scans, EEG, and ECG. For example, there are some researchers developing ML models to predict neurological outcomes using head CT images[16–18] and EEG,[19–21] potentially leading to more accurate and timely diagnoses. Similarly, ML models have been employed to analyze ECG data, enabling the prediction of critical events such as in-hospital cardiac arrest, ventricular arrhythmia, sudden cardiac death, and the success of defibrillation during resuscitation.[22–26] These applications of ML models using medical imaging and bio-signals are expected to contribute to facilitating early detection, improving predictive accuracy, and ultimately enhancing more appropriate resuscitation, emergency, or intensive care.

Exploring sub-phenotypes and treatment heterogeneity

ML is also used to explore sub-phenotypes, an emerging concept in precision medicine. (Fig. 2) Sub-phenotypes are distinct subgroups within a disease or condition characterized by different clinical features such as disease progression, outcomes, and underlying biological mechanisms.[27,28] Whereas phenotypes represent categories of patients with common features such as a specific syndrome, e.g., sepsis or acute respiratory distress syndrome,[27,28] sub-phenotype is particularly relevant when discussing subgroups with heterogeneity on treatment effect.[27] Heterogeneity on treatment effect refers to the variation in how different individuals or groups respond to the same treatment.[5] It means that not all patients respond to treatments in the same way due to various factors such as genetic differences, lifestyle factors, pre-existing health conditions, and more.[5] Understanding the concept of sub-phenotypes and the complexities of treatment effect heterogeneity are anticipated to advance the development of personalized medicine, moving beyond the conventional 'one-size-fits-all' treatment approach. For example, some research in the resuscitation context suggests the hypothesis that sub-phenotypes exhibit heterogeneity of effect of targeted temperature management, such as some subgroups may have the potential benefit of hypothermia (e.g., at 33), while others may not.[29,30] For exploring such treatment heterogeneity, ML such as "clustering" or "causal machine learning" are utilized in some research.[31,32]

Clustering

Clustering is a type of unsupervised machine learning that can be used to identify subgroups who share similar clinical characteristics and explore treatment heterogeneity or novel association between the subgroups and events, using data such as patients' characteris-

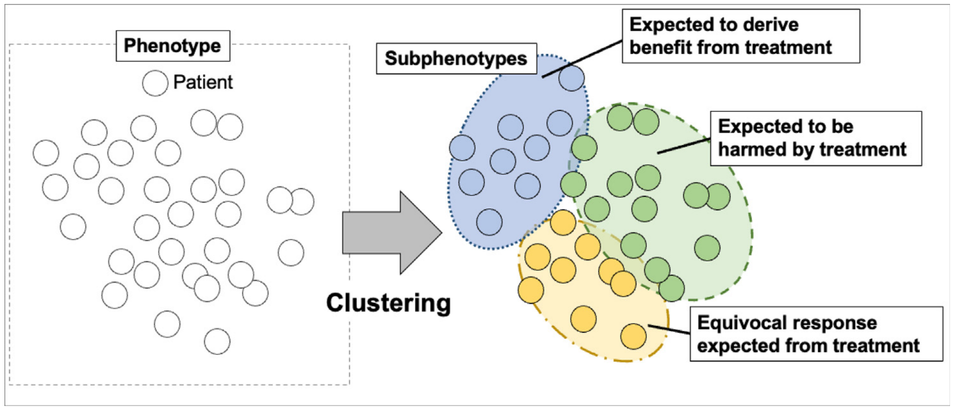

Fig. 2 – The concept of clustering and sub-phenotypes Phenotypes (e.g., sepsis, acute respiratory distress syndrome) are categorized by clustering to sub-phenotypes with different clinical features and the heterogeneous response to the treatment.

tics, biomarker values, and genomic data (Fig. 2).[31,33] One of the strengths of clustering is its ability to manage data complexity and discover hidden patterns, making the data easier to understand and visualize. Previously, this clustering analysis was used in research exploring novel sub-phenotypes among patients with various patterns in emergency medicine and critical care such as sepsis, ARDS, trauma, and cardiac arrest.[27,28,34–39]

For instance, various clinical patterns in coagulopathy among patients with severe head trauma are associated with different outcomes.[38] There are also subgroups among OHCA patients with different clinical outcomes when treated with ECPR.[39] Some research suggests the effect of early goal-directed treatment or the effect of drugs on coagulopathy are different among subgroups in sepsis.[36,37] This technique is also utilized to summarize the risk factors as a subgroup. One example is the subgroups with environmental features characterized by environmental parameters such as temperature, wind speed, and air pollution are suggested to be associated with the occurrence of acute myocardial infarction or acute ischemic stroke.[40,41]

Causal machine learning

Causal machine learning is an ML approach to investigate causal inference, which is particularly valuable in assessing heterogeneity in treatment effects (Fig. 3).[5,42,43] Causal forest, one approach within causal machine learning based on the random forest, works by splitting the data into different subgroups and assessing the treatment effect within each subgroup by handling the no-linear and/or hi-dimensional data.[5,42,43] For example in critical care fields, the causal forest was used on data from an RCT about the effect of using a bougie during intubation.[44] This RCT found that using a bougie did not increase the incidence of successful intubation on the first attempt in all critically ill adults; however, the causal forest analysis suggested some individuals who had the potential benefit of using a bougie.

The application of machine learning using genetic and molecular data (omics data) to treatment heterogeneity and precision medicine is also expected to result in a more personalized approach to health-care such as investigating the heterogeneity of the treatment response or adverse events of drugs among patients with certain genetic features.[45,46] Although this type of research is mainly focused on the oncology field because the drugs are commonly targeted to specific genetic features,[47,48] there will be an increasing number of studies on treatment heterogeneity and pharmacogenomics in the resuscitation field.

Reinforcement learning to optimize treatment

Reinforcement learning is a type of machine learning that autonomously chooses actions to maximize rewards obtained from the given environment. The system learns through trial and error to select actions that lead to the highest possible reward. Reinforcement learning has broad applications and is particularly useful for complex tasks, such as games, autonomous driving, robotics, and logistics.[49] For example, in 2015, AlphaGo, an AI developed using reinforcement learning, famously defeated a world champion Go player.[50]

In the field of medicine and healthcare, reinforcement learning has potential applications in optimizing treatment strategies. (Fig. 4) For example, one notable example of using reinforcement learning, in the context of intensive care, is the development of an "AI Clinician" for sepsis treatment in managing fluids and vasopressors.[51] This AI system analyzed two ICU databases and learned optimal treatment strategies by examining numerous treatment decisions to maximize the expected survival outcome. As a result, this AI model could select the optimal treatment strategy which showed the lowest mortality rates. Another model using reinforcement learning suggested personalized optimization of mechanical ventilation in patients staying at cardiovascular ICUs.[52] In other examples, some reinforcement learning programs were suggested to investigate the optimal dose of sedative agents in general anesthesia.[53] Although there are few published research using reinforcement learning in the resuscitation field, it has potential for future studies.

Natural language processing

Natural language processing (NLP) is a subset of ML technology that enables computers to analyze the language that humans usually use in daily life. This technology is prevalent in our modern lives with applications using voice recognition such as voice assistant

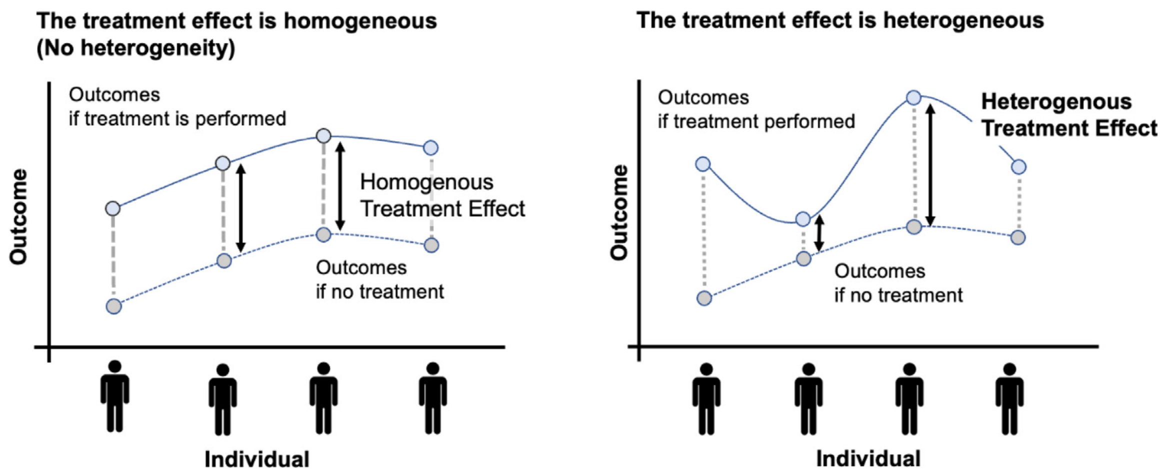

Fig. 3 – The concept of treatment heterogeneity (Left) Assuming that the difference between outcomes when treatment is performed and when it is not, is the same in each patient: treatment effect is homogenous between individual patients. (Right) Assuming that the difference between outcomes when treatment is performed and when it is not, is different in each patient: treatment effect is heterogenous between individual patients.

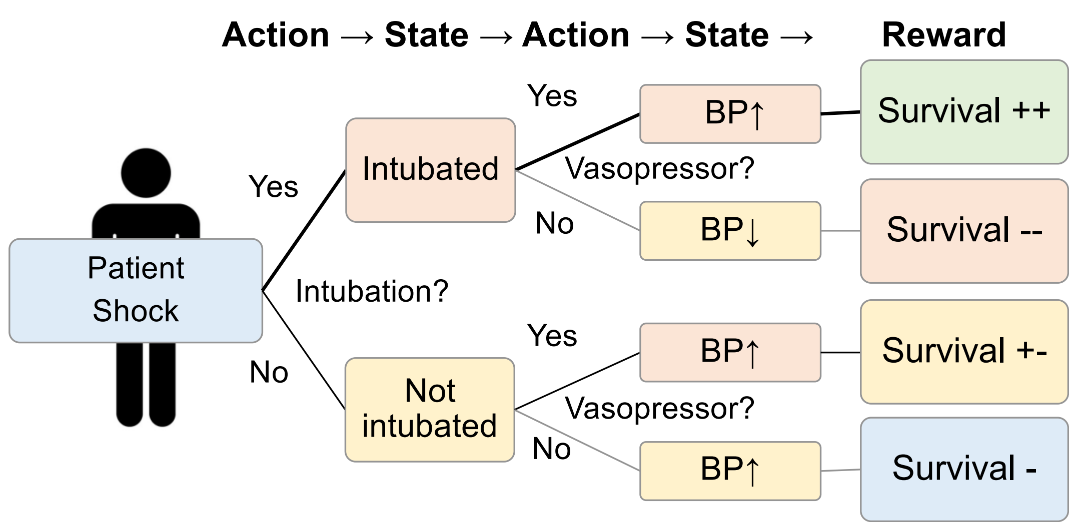

Fig. 4 – The concept of reinforcement learning in medical research. Patient status is changed to a different status by the action, and consequently, the reward is obtained based on the status. Reinforcement learning can find the best strategy to maximize the rewards based on many trials.

programs like Apple's Siri or Google Assistant and using text like chatbots or language translation tools.

In the field of research in resuscitation, NLP models are being utilized in innovative ways. One notable example of using voice data is ML programs to help recognize cardiac arrest and support initiating bystander-CPR during emergency calls to the dispatch center (Fig. 5).[54,55] These programs can analyze the caller's words during an emergency call and estimate the probability of the patient being in cardiac arrest. This kind of program has also been applied in research to detect other emergencies such as severe trauma after road trauma and stroke.[56,57] Additionally, NLP voice recognition technology offers practical benefits for paramedics in the field. Paramedics can use voice commands to create prehospital records thereby reducing the need for manual data entry and enabling them to focus more on patient care.[58] These programs have the potential to enable faster and more accurate deployment of emergency medical services, which can improve patient outcomes.

NLP technology can also be utilized to analyze clinical data from the free text in medical records such as medical history or physical findings.[59] Algorithms can be developed to predict emergency conditions such as in-hospital cardiac arrest or give decision support on the appropriate disposition of patients at the emergency department.[59–63] This technique can also be used to accurately predict neurological outcomes such as a modified Rankin scale by analyzing free text data in clinical notes.[64] Additionally, chatbot tools using NLP have also been developed in the resuscitation research fields. One example is a preliminary chatbot to guide users on how to perform bystander CPR.[62] In summary, NLP-applied research using voice or text is increasing and they can analyze communication or medical records to predict events and be a guide to action in resuscitation.

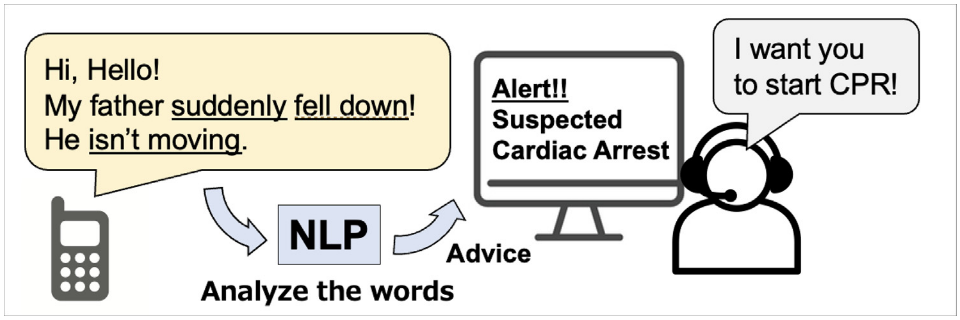

Fig. 5 – Example of Natural Language Processing for Activating Bystander CPR NLP: Natural Language Processing, CPR: Cardiopulmonary resuscitation In the emergency call dispatch center, the application utilizes natural language processing (NLP) to analyze the caller's words, aiding the dispatcher in identifying potential cases of cardiac arrest.[54]

Large language model (LLM) is one domain of research in NLP fields that can understand and generate natural language used by humans. Typically, by learning patterns from large amounts of textual data, these models can generate answers to new questions, or produce text to accomplish specific tasks such as translation or revising the text. Recently, the GPT-3 and GPT-4 developed by OpenAI have attracted a lot of attention for their wide adaptability and flexibility.[65] If you enter the prompt "What should we do if we encounter a patient who has suddenly collapsed?" into the application, the application can provide plausible answers as if they are provided by a healthcare professional. (However, it should be noted that these answers may be incorrect.) One representative example of using LLM is that the LLM can pass the medical licensing examination without any additional training data.[66,67] Further, some research indicated that LLM can provide quality and empathetic responses to patient questions.[68,69] Further, the LLM is also expected to summarize the clinical information from medical records like a professional or perform the systematic review instead of humans.[70,71] Although research in the resuscitation field has not yet been published, it is expected to develop in the future. In contrast, this LLM has also caused various controversial issues, such as the accuracy, validity, and responsibility of the generated sentences and ethical issues that may arise (more detail is discussed in the next paragraph).[65] Although several concerns, LLM has great potential to improve the burden on healthcare providers, especially in terms of decision-making, documentation, and summarizing medical information.

Challenges for AI and ML in resuscitation research and implementation

Despite the extensive research conducted, actual implementation of AI and ML in the clinical setting remains limited, though some practices have implemented AI and ML-based algorithms in resuscitation and intensive care.[14,15,54,72,73] Widespread adoption may be slow due to several concerns and limitations.[74] Here we give an overview of the most important challenges and barriers that prevent proper implementation.

Data quality and availability

AI and ML algorithms heavily depend on the quality of data they are trained on. If the data is unreliable, missing, incomplete, or biased, the model's predictions or performance can be inaccurate or even harmful. A prediction model may simply be biased because of the original data it is trained on, reflecting the existing bias as is. For example, an AI model may reflect historical disparities in healthcare access and outcomes, and inadvertently perpetuate these biases by recommending differential treatment based on factors such as race, gender, age, or socioeconomic status.[75,76] It is therefore essential that the training data is diverse and representative of the patient population. However, in the actual scene of resuscitation, obtaining comprehensive and diverse datasets can be challenging. Clinical situations can change drastically in a short time, making it difficult to comprehensively collect data in a timely manner, such as in a resource-limited environment like the prehospital setting or a crowded emergency department.[77,78] Furthermore, in many settings of resuscitation, clinical data is still being recorded using paper and pen, and some backend data entry process is needed to integrate the data into electronic medical records for it to be utilized for ML application.[79] Yet, ensuring the availability of comprehensive and representative data is crucial to develop accurate and generalizable models.

Validation process to verify the reproducibility

Once ML models have been developed, they should be reproducable.[80] Previously, it has been reported that many prediction models have a high risk of bias, especially due to the lack of the validation process to confirm the reproducibility of the models using different datasets.[80–83] One of the problems to validate the ML and AI models using different datasets is the difficulty in obtaining different data from the original data with consistent format and definition of the variables. In the resuscitation fields, the Utstein format is broadly accepted as a universal data-collecting standard mainly in prehospital settings; however, some of the in-hospital data have not yet been standardized (e.g., some variables in the emergency department or intensive care unit have still not been strictly defined).[10]

Another concern is inappropriate reporting of the originally developed models.[83] Reproducibility can be difficult to ascertain as details of the models are not reported.[83] Furthermore, validation study risks selective reporting bias, meaning that validation studies reporting models with poor performance are less likely to get published.[81] Yet, ensuring robustness in AI and ML models, including their reliability and reproducibility, is essential to prevent or minimize unintended harm.

Generalizability and clinical integration

Verifying the Generalizability is also essential to validate the AI and ML models prior to clinical application. Again, ML models depend on data, and if the model too strongly fits certain features of the data ("overfitting"), the results may not be generalizable to the different population without those features. Resuscitation practices vary across different healthcare settings, geopolitical contexts, and patient populations.[84–87] AI models developed in one context may not generalize well when traveling to other settings. Ensuring the generalizability and applicability the models to diverse populations, different clinical protocols and resource-constrained environments is essential for their widespread application.[87]

Additionally, other practical barriers exist to implementing AI and ML in clinical settings. It includes not only regulatory approval but also integration into clinical workflows. Moreover, the adoption of ML models necessitates clear benefits in routine clinical practice, such as improving patients' outcomes and reducing workload or costs. However, few randomized controlled trials (RCTs) have shown the actual benefit of ML models in clinical settings.[42,80] If integration of ML models into general clinical workflows does not yield clear benefits for clinicians, patients, or other stakeholders, no one would use these models. The actual benefit of ML tools in clinical settings compared to existing clinical workflows need to be demonstrated in research before widespread adoption will follow.

Self-fulfilling prophecies and feedback loops

Another important issue to be focused on in the resuscitation field is the risk of hidden false positive bias by self-fulfilling prophecy and feedback loop when predicting the prognosis of cardiac arrest patients.[88,89] A self-fulfilling prophecy is a prediction that influences people's beliefs and behavior through which the prediction is then realized.[90] In resuscitation, if clinicians expect that a particular patient may not survive despite the best treatment, the expectation could influence their decision to forego further treatment, allowing the patient's death, thereby fulfilling the initial prediction (self-fulfilling prophecy). This becomes especially problematic if the initial prediction was incorrect (a false-positive), which could result in the patient not receiving the potentially beneficial care. While these issues have existed even before AI and ML are developed (because predictions of clinicians are sometimes inaccurate),[91] there is growing concern that AI and ML might amplify the bias due to self-fulfilling feedback loops (Fig. 6). If a model trained on biased data is applied to guide clinical decision-making, and the new data influenced by the model's results are then used as input data again to "improve" the model, there is a risk that the initial biases will be reinforced and amplified. To illustrate, if a prediction model is developed using data from a hospital where resuscitation efforts were consistently terminated early for OHCA patients aged over 70 years old during a specific period due to temporary limitation of resources (such as limitation of intensive care during the COVID-19 pandemic), the model may inevitably predict the lower probability of survival for similar patients than is accurate. This prediction merely reflects the flawed input data itself rather than the truth under ideal circumstances. Yet, if clinicians perceive this prediction as "accurate" and terminate resuscitation efforts based on such false positives, no one will notice the missed opportunities for successful resuscitation of OHCA patients over 70 years, since the outcome confirms the prediction.[88] If new models are then trained based on the confirmed biased data, it can further amplify the biased prediction and inappropriate withdrawal rates. In essence, past mistakes lead to new self-fulfilling prophecies, reinforcing predictions that generate inappropriate clinical judgments, creating a vicious cycle; an automated feedback loop of self-fulfilling poor outcomes for future cardiac arrest patients.[88] Furthermore, the lack of error signals due to confirmative outcomes combined with the lack of interpretability of ML models greatly hinders clinicians from recognizing such biased predictive feedback loops. Catching false positives retrospectively is near to impossible, since this would require counterfactual data. Clinical guidelines suggest the need for a multi-modal approach to predict the outcome of cardiac arrest patients to minimize the potential harm of false-positive of predictions.[92] When advanced AI models are developed, clinicians must remain aware of the risk for amplified bias through self-fulfilling prophecies and feedback loops.

Transparency, Interpretability, and trust

A key challenge when applying AI and ML to the actual resuscitation scene is the interpretability of and trust in ML models.[80–82] ML models are often described as a 'black box' due to the complexity of the models that generate the results. This lack of transparency can hinder clinicians' or patients' trust towards ML models. One example is, as mentioned above, an ML model was developed to detect potential cardiac arrest cases using the voice data of emergency calls at the dispatch center.[54] The retrospective observational study using the voice recordings indicated that the ML model outperformed human dispatchers.[77] However, the RCT comparing the dispatcher assisted by the ML model to those without such assistance, did not demonstrate any improvement in the performance to recognize the cardiac arrest cases.[54] One of the potential mechanisms of this result suggested by the research team was that the dispatcher could not understand the ML model's decision-making process and the dispatcher possibly did not trust the alert from the ML program.[93] Had the advice come from human experts instead of the ML model, the dispatchers might have asked the rationale why and how they concluded, considered accepting (or rejecting) their suggestion, and thereby improved their performance to recognize the cardiac arrest case. As such, achieving interpretability and trust in ML models may be essential to successfully implement AI and ML into real-world clinical practice.

Regulatory and legal challenges

While proper data collection and management is an essential prerequisite for developing and applying ML models to clinical settings, such data collection and management must of course respect privacy and comply with the law.[94] Furthermore, liability and responsibility frameworks need to be developed and implemented for AI-driven and ML-based resuscitation interventions, in order to ensure accountability and patient safety. As seen in this article, AI and ML can raise several ethical concerns when it is applied to the actual medical system and care, although the ethical concerns far exceed the ones we mention here. Generally speaking, the Ethics Guideline for Trustworthy AI suggested seven key requirements including human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, environmental and societal well-being, and, accountability.[94] While we have selected several significant issues particular to resuscitation, these ethical principles should be addressed across all AI applications in medicine, regardless of the specialty. Indeed, many non-profit institutions, regulatory, and gov-

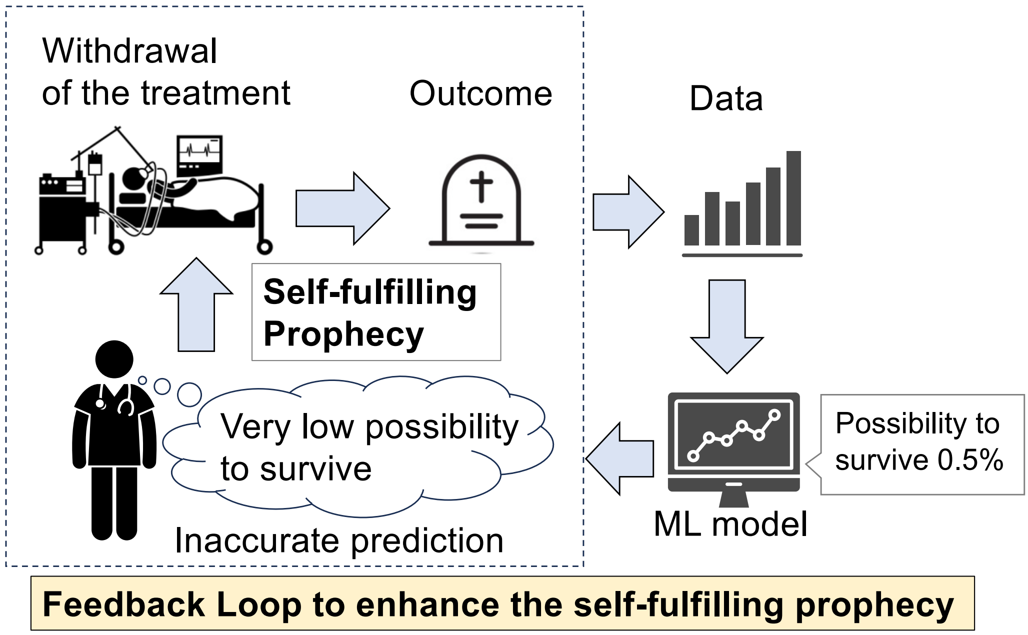

Fig. 6 – Concept of self-fulfilling prophecy and its feedback loop. A particular patient who could be saved is assessed as "Very low possibility to survive" by the inaccurate and biased prediction, which can potentially lead to the decision-making of treatment withdrawal. As a result, the initial prediction "Very low possibility to survive" is realized. If this data is utilized to develop the ML models, it can amplify and reproduce the false prediction, which lead to the potential harm that the patients lose the opportunity to be treated by the enhanced inaccurate prediction.

ernmental bodies across the world are currently collaborating to ensure (inter)national laws that better protect citizens from the rapidly increasing impacts of AI and ML-driven models.

Conclusion

In this article, we introduce and illustrate important concepts within AI and ML research in the resuscitation field. The application of AI and ML in resuscitation research holds significant potential to revolutionize the field by improving prediction, supporting decision-making, and developing personalized treatment strategies. However, various limitations and ethical concerns must be addressed to ensure the responsible and effective implementation of these technologies in actual clinical practice. As more high-quality data becomes available, it is expected that AI-driven models and ML-based algorithms will play an increasingly important role in resuscitation research and practice. Moving forward, it will be essential for researchers, computer scientists, clinicians, ethicists, policymakers, and other stakeholders to work together to overcome the challenges and harness the full potential of AI and ML in resuscitation, ultimately leading to better patient outcomes and more efficient healthcare systems.

Ethical approval

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

Not applicable.

Funding

This study was supported by a scientific research grant from the JSPS KAKENHI of Japan (JP22K21143) and the Zoll foundation. YO has received an overseas scholarship from the Japan Society for the Promotion of Science, the FUKUDA Foundation for medical technology, and the International medical research foundation. MM is funded by the European Union, through the HORIZON-MSCA-2022-PF-01-01 Marie Curie Postdoctoral Fellowship, Project 101107292 'PredicGenX'.

CRediT authorship contribution statement

Yohei Okada: Conceptualization, Writing – original draft. **Mayli Mertens:** Conceptualization, Writing – original draft. **Nan Liu:** Writing – review & editing. **Sean Shao Wei Lam:** Writing – review & editing. **Marcus Eng Hock Ong:** Writing – review & editing.

Declaration of Competing Interest

YO has received a research grant from the ZOLL Foundation and overseas scholarships from the Japan Society for Promotion of Science, the FUKUDA Foundation for medical technology, and the International medical research foundation. These organizations have no role in conducting this research. MEHO reports grants from the Laerdal Foundation, Laerdal Medical, and Ramsey Social Justice Foundation for funding of the Pan-Asian Resuscitation Outcomes Study an advisory relationship with Global Healthcare SG, a commercial entity that manufactures cooling devices; and funding from Laerdal Medical on an observation program to their Community CPR Training Centre Research Program in Norway. MEHO is a Scientific Advisor to TIIM Healthcare SG and Global Healthcare SG.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.resplu.2023.100435.

Author details

[a]Duke-NUS Medical School, National University of Singapore, Singapore [b]Preventive Services, Graduate School of Medicine, Kyoto University, Kyoto, Japan [c]Antwerp Center for Responsible AI, Antwerp University, Belgium [d]Centre for Ethics, Department of Philosophy, Antwerp University, Belgium [e]Department of Emergency Medicine, Singapore General Hospital

REFERENCES

1. Kühl N, Schemmer M, Goutier M, Satzger G. Artificial intelligence and machine learning. Electronic Markets 2022;32:2235–44. https://doi.org/10.1007/s12525-022-00598-0.
2. Goto T, Camargo Jr 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:e186937. https://doi.org/10.1001/jamanetworkopen.2018.6937.
3. Okada Y, Matsuyama T, Morita S, et al. Machine learning-based prediction models for accidental hypothermia patients. J Intensive Care 2021;9:6.
4. Bihorac A, Ozrazgat-Baslanti T, Ebadi A, et al. MySurgeryRisk: development and validation of a machine-learning risk algorithm for major complications and death after surgery. Ann Surg 2019;269:652.
5. Gong X, Hu M, Basu M, Zhao L. Heterogeneous treatment effect analysis based on machine-learning methodology. CPT Pharmacometrics Syst Pharmacol 2021;10:1433–43. https://doi.org/10.1002/psp4.12715.
6. Riascos A, Romero M, Serna N. Risk adjustment revisited using machine learning techniques. Documento Cede 2017.
7. Nishioka N, Kobayashi D, Kiguchi T, et al. Development and validation of early prediction for neurological outcome at 90 days after return of spontaneous circulation in out-of-hospital cardiac arrest. Resuscitation 2021;168:142–50. https://doi.org/10.1016/j.resuscitation.2021.09.027.
8. Liu N, Liu M, Chen X, et al. Development and validation of an interpretable prehospital return of spontaneous circulation (P-ROSC) score for patients with out-of-hospital cardiac arrest using machine learning: A retrospective study. EClinicalMedicine 2022;48:101422. https://doi.org/10.1016/j.eclinm.2022.101422.
9. Jacobs I, Nadkarni V, Bahr J, et al. Cardiac arrest and cardiopulmonary resuscitation outcome reports: update and simplification of the Utstein templates for resuscitation registries: a statement for healthcare professionals from a task force of the International Liaison Committee on Resuscitation (American Heart Association, European Resuscitation Council, Australian Resuscitation Council, New Zealand Resuscitation Council, Heart and Stroke Foundation of Canada, InterAmerican Heart Foundation, Resuscitation Councils of Southern Africa). Circulation 2004.
10. Perkins GD, Jacobs IG, Nadkarni VM, et al. Cardiac Arrest and Cardiopulmonary Resuscitation Outcome Reports: Update of the Utstein Resuscitation Registry Templates for Out-of-Hospital Cardiac Arrest. Circulation 2015;132:1286–300. https://doi.org/10.1161/CIR.0000000000000144.
11. Nanayakkara S, Fogarty S, Tremeer M, et al. Characterising risk of in-hospital mortality following cardiac arrest using machine learning: A retrospective international registry study. PLoS Med 2018;15:e1002709. https://doi.org/10.1371/journal.pmed.1002709.
12. Pimentel MAF, Redfern OC, Malycha J, et al. Detecting Deteriorating Patients in the Hospital: Development and Validation of a Novel Scoring System. Am J Respir Crit Care Med 2021;204:44–52. https://doi.org/10.1164/rccm.202007-2700OC.
13. 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;269:1059–63.
14. Winslow CJ, Edelson DP, Churpek MM, et al. The Impact of a Machine Learning Early Warning Score on Hospital Mortality: A Multicenter Clinical Intervention Trial. Crit Care Med 2022;50:1339–47. https://doi.org/10.1097/ccm.0000000000005492.
15. Goldstein BA, Cerullo M, Krishnamoorthy V, et al. Development and Performance of a Clinical Decision Support Tool to Inform Resource Utilization for Elective Operations. JAMA Netw Open 2020;3:e2023547. https://doi.org/10.1001/jamanetworkopen.2020.23547.
16. Kawai Y, Kogeichi Y, Yamamoto K, Miyazaki K, Asai H, Fukushima H. Explainable artificial intelligence-based prediction of poor neurological outcome from head computed tomography in the immediate post-resuscitation phase. Sci Rep 2023;13:5759. https://doi.org/10.1038/s41598-023-32899-5.
17. Mansour A, Fuhrman JD, Ammar FE, et al. Machine Learning for Early Detection of Hypoxic-Ischemic Brain Injury After Cardiac Arrest. Neurocrit Care 2022;36:974–82.
18. Elmer J, Liu C, Pease M, et al. Deep learning of early brain imaging to predict post-arrest electroencephalography. Resuscitation 2022;172:17–23. https://doi.org/10.1016/j.resuscitation.2022.01.004.
19. Zheng WL, Amorim E, Jing J, et al. Predicting neurological outcome in comatose patients after cardiac arrest with multiscale deep neural networks. Resuscitation 2021;169:86–94.
20. Zheng WL, Amorim E, Jing J, et al. Predicting Neurological Outcome From Electroencephalogram Dynamics in Comatose Patients After Cardiac Arrest With Deep Learning. IEEE Trans Biomed Eng 2022;69:1813–25.
21. Jonas S, Rossetti AO, Oddo M, Jenni S, Favaro P, Zubler F. EEG-based outcome prediction after cardiac arrest with convolutional neural networks: Performance and visualization of discriminative features. Hum Brain Mapp 2019;40:4606–17. https://doi.org/10.1002/hbm.24724.
22. Hajeb-M S, Cascella A, Valentine M, Chon KH. Deep Neural Network Approach for Continuous ECG-Based Automated External Defibrillator Shock Advisory System During Cardiopulmonary Resuscitation. J Am Heart Assoc 2021;10:e019065.
23. Kwon J-M, Kim K-H, Jeon K-H, Lee SY, Park J, Oh B-H. Artificial intelligence algorithm for predicting cardiac arrest using electrocardiography. Scand J Trauma, Resus Emergency Med 2020;28:98. https://doi.org/10.1186/s13049-020-00791-0.

24. Kolk MZH, Deb B, Ruipérez-Campillo S, et al. Machine learning of electrophysiological signals for the prediction of ventricular arrhythmias: systematic review and examination of heterogeneity between studies. eBioMedicine 2023:89. https://doi.org/10.1016/j.ebiom.2023.104462.
25. Sem M, Mastrangelo E, Lightfoot D, Aves T, Lin S, Mohindra R. The ability of machine learning algorithms to predict defibrillation success during cardiac arrest: a systematic review. Resuscitation 2023;185. https://doi.org/10.1016/j.resuscitation.2023.109755.
26. Kenet AL, Pemmaraju R, Ghate S, et al. A pilot study to predict cardiac arrest in the pediatric intensive care unit. Resuscitation 2023;185:109740. https://doi.org/10.1016/j.resuscitation.2023.109740.
27. Reddy K, Sinha P, O'Kane CM, Gordon AC, Calfee CS, McAuley DF. Subphenotypes in critical care: translation into clinical practice. Lancet Respir Med 2020;8:631–43. https://doi.org/10.1016/s2213-2600(20)30124-7.
28. Wildi K, Livingstone S, Palmieri C, LiBassi G, Suen J, Fraser J. The discovery of biological subphenotypes in ARDS: a novel approach to targeted medicine? J Intensive Care 2021;9:14. https://doi.org/10.1186/s40560-021-00528-w.
29. Callaway CW, Coppler PJ, Faro J, et al. Association of Initial Illness Severity and Outcomes After Cardiac Arrest With Targeted Temperature Management at 36 °C or 33 °C. JAMA Netw Open 2020;3:e208215–e. https://doi.org/10.1001/jamanetworkopen.2020.8215.
30. Nishikimi M, Ogura T, Nishida K, et al. Outcome Related to Level of Targeted Temperature Management in Postcardiac Arrest Syndrome of Low, Moderate, and High Severities: A Nationwide Multicenter Prospective Registry. Crit Care Med 2021;8:e741–50.
31. Loftus TJ, Shickel B, Balch JA, et al. Phenotype clustering in health care: A narrative review for clinicians. Front Artif Intell 2022;5:842306. https://doi.org/10.3389/frai.2022.842306.
32. Jawadekar N, Kezios K, Odden MC, et al. Practical Guide to Honest Causal Forests for Identifying Heterogeneous Treatment Effects. *Am J Epidemiol* 2023. Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health For permissions, please e-mail: journals.permissions@oup.com.; 2023.
33. Sinha P, Calfee CS, Delucchi KL. Practitioner's Guide to Latent Class Analysis: Methodological Considerations and Common Pitfalls. Crit Care Med 2021;49:e63–79. https://doi.org/10.1097/CCM.0000000000004710.
34. Okada Y, Komukai S, Kitamura T, et al. Clustering out-of-hospital cardiac arrest patients with non-shockable rhythm by machine learning latent class analysis. *Acute Med Surg* 2022;9:e760.
35. Wilson JG, Subphenotypes CCSA. Understanding a Heterogeneous Syndrome. Crit Care 2020;24:102. https://doi.org/10.1186/s13054-020-2778-x.
36. Seymour CW, Kennedy JN, Wang S, et al. Derivation, validation, and potential treatment implications of novel clinical phenotypes for sepsis. JAMA 2019;321:2003–17.
37. Kudo D, Goto T, Uchimido R, et al. Coagulation phenotypes in sepsis and effects of recombinant human thrombomodulin: an analysis of three multicentre observational studies. Crit Care 2021;25:114.
38. Fujiwara G, Okada Y, Shiomi N, Sakakibara T, Yamaki T, Hashimoto N. Derivation of Coagulation Phenotypes and the Association with Prognosis in Traumatic Brain Injury: A Cluster Analysis of Nationwide Multicenter Study. Neurocrit Care 2023.
39. Okada Y, Komukai S, Kitamura T, et al. Clinical Phenotyping of Out-of-Hospital Cardiac Arrest Patients With Shockable Rhythm - Machine Learning-Based Unsupervised Cluster Analysis. Circ J 2022;86:668–76. https://doi.org/10.1253/circj.CJ-21-0675.
40. Koo GPY, Zheng H, Pek PP, et al. Clustering of Environmental Parameters and the Risk of Acute Myocardial Infarction. Int J Environ Res Public Health 2022;19. https://doi.org/10.3390/ijerph19148476.
41. Koo GPY, Zheng H, Aik JCL, et al. Clustering of Environmental Parameters and the Risk of Acute Ischaemic Stroke. Int J Environ Res Public Health 2023:20. https://doi.org/10.3390/ijerph20064979.
42. Athey S, Wager S. Estimating treatment effects with causal forests: an application. Observational Studies 2019;5:37–51.
43. Wager S, Athey S. Estimation and inference of heterogeneous treatment effects using random forests. J Am Stat Assoc 2018;113:1228–42.
44. Seitz KP, Spicer AB, Casey JD, et al. Individualized Treatment Effects of Bougie vs Stylet for Tracheal Intubation in Critical Illness. Am J Respir Crit Care Med 2023. https://doi.org/10.1164/rccm.202209-1799OC.
45. Syrowatka A, Song W, Amato MG, et al. Key use cases for artificial intelligence to reduce the frequency of adverse drug events: a scoping review. Lancet Digital Health 2022;4:e137–48. https://doi.org/10.1016/S2589-7500(21)00229-6.
46. Kline A, Wang H, Li Y, et al. Multimodal machine learning in precision health: a scoping review. npj Digital Med 2022;5:171. https://doi.org/10.1038/s41746-022-00712-8.
47. Azuaje F. Artificial intelligence for precision oncology: beyond patient stratification. *npj Prec* Oncol 2019;3:6. https://doi.org/10.1038/s41698-019-0078-1.
48. Lauschke VM, Ingelman-Sundberg M. Emerging strategies to bridge the gap between pharmacogenomic research and its clinical implementation. *npj* Genom Med 2020;5:9. https://doi.org/10.1038/s41525-020-0119-2.
49. Liu S, See KC, Ngiam KY, Celi LA, Sun X, Feng M. Reinforcement learning for clinical decision support in critical care: comprehensive review. J Med Internet Res 2020;22:e18477.
50. Silver D, Huang A, Maddison CJ, et al. Mastering the game of Go with deep neural networks and tree search. Nature 2016;529:484–9. https://doi.org/10.1038/nature16961.
51. Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. Nat Med 2018;24:1716–20. https://doi.org/10.1038/s41591-018-0213-5.
52. Peine A, Hallawa A, Bickenbach J, et al. Development and validation of a reinforcement learning algorithm to dynamically optimize mechanical ventilation in critical care. NPJ Digit Med 2021;4:32. https://doi.org/10.1038/s41746-021-00388-6.
53. Yun WJ, Shin M, Jung S, Ko J, Lee HC, Kim J. Deep reinforcement learning-based propofol infusion control for anesthesia: A feasibility study with a 3000-subject dataset. Comput Biol Med 2023;156:106739. https://doi.org/10.1016/j.compbiomed.2023.106739.
54. Blomberg SN, Christensen HC, Lippert F, et al. Effect of Machine Learning on Dispatcher Recognition of Out-of-Hospital Cardiac Arrest During Calls to Emergency Medical Services: A Randomized Clinical Trial. JAMA Netw Open 2021;4:e2032320–e. https://doi.org/10.1001/jamanetworkopen.2020.32320.
55. Byrsell F, Claesson A, Ringh M, et al. Machine learning can support dispatchers to better and faster recognize out-of-hospital cardiac arrest during emergency calls: A retrospective study. Resuscitation 2021;162:218–26. https://doi.org/10.1016/j.resuscitation.2021.02.041.
56. Chin KC, Cheng YC, Sun JT, et al. Machine Learning-Based Text Analysis to Predict Severely Injured Patients in Emergency Medical Dispatch: Model Development and Validation. J Med Internet Res 2022;24. https://doi.org/10.2196/30210.
57. Scholz ML, Collatz-Christensen H, Blomberg SNF, Boebel S, Verhoeven J, Krafft T. Artificial intelligence in Emergency Medical Services dispatching: assessing the potential impact of an automatic speech recognition software on stroke detection taking the Capital Region of Denmark as case in point. Scand J Trauma Resusc Emerg Med 2022;30:36. https://doi.org/10.1186/s13049-022-01020-6.
58. Fukaguchi K, Goto T, Yamamoto T, Yamagami H. Experimental Implementation of NSER Mobile App for Efficient Real-Time Sharing of Prehospital Patient Information With Emergency Departments: Interrupted Time-Series Analysis. JMIR Formative Research 2022;6:e37301.

59. Goto T, Hara K, Hashimoto K, et al. Validation of chief complaints, medical history, medications, and physician diagnoses structured with an integrated emergency department information system in Japan: the Next Stage ER system. *Acute Med Surg* Jan-Dec 2020;7:e554.
60. 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 Inf 2019,;129:184–8. https://doi.org/10.1016/j.ijmedinf.2019.06.008.
61. Brown JR, Ricket IM, Reeves RM, et al. Information Extraction From Electronic Health Records to Predict Readmission Following Acute Myocardial Infarction: Does Natural Language Processing Using Clinical Notes Improve Prediction of Readmission? Journal of the American Heart Association 2022;;11:e024198.
62. 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,;47:265–278.e7. https://doi.org/10.1016/j.jen.2020.11.001.
63. Okada Y, Okada A, Ito H, Sonoo T, Goto T. External validation of the POP score for predicting obstetric and gynecological diseases in the emergency department. *The.* Am J Emerg Med 2022,;51:348–53. https://doi.org/10.1016/j.ajem.2021.11.022.
64. Fernandes MB, Valizadeh N, Alabsi HS, et al. Classification of neurologic outcomes from medical notes using natural language processing. Expert Syst Appl 2023:214. https://doi.org/10.1016/j.eswa.2022.119171.
65. Haupt CE, Marks M. AI-Generated Medical Advice—GPT and Beyond. JAMA 2023;329:1349–50. https://doi.org/10.1001/jama.2023.5321.
66. Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. PLoS digital health 2023;2:e0000198.
67. Kasai J, Kasai Y, Sakaguchi K, Yamada Y, Radev D. Evaluating gpt-4 and chatgpt on japanese medical licensing examinations. *arXiv* 2023. *preprint arXiv:230318027*.
68. 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. https://doi.org/10.1001/jamainternmed.2023.1838.
69. Sarraju A, Bruemmer D, Van Iterson E, Cho L, Rodriguez F, Laffin L. Appropriateness of Cardiovascular Disease Prevention Recommendations Obtained From a Popular Online Chat-Based Artificial Intelligence Model. JAMA 2023;329:842–4. https://doi.org/10.1001/jama.2023.1044.
70. Patel SB, Lam K. ChatGPT: the future of discharge summaries? Lancet Digital Health 2023;5:e107–8. https://doi.org/10.1016/S2589-7500(23)00021-3.
71. Qureshi R, Shaughnessy D, Gill KAR, Robinson KA, Li T, Agai E. Are ChatGPT and large language models “the answer” to bringing us closer to systematic review automation? Syst Rev 2023;12:72. https://doi.org/10.1186/s13643-023-02243-z.
72. Pham SD, Keijzer HM, Ruijter BJ, et al. Outcome Prediction of Postanoxic Coma: A Comparison of Automated Electroencephalography Analysis Methods. Neurocrit Care 2022;37:248–58.
73. Aellen FM, Alnes SL, Loosli F, et al. Auditory stimulation and deep learning predict awakening from coma after cardiac arrest. Brain 2023.
74. Chan SL, Lee JW, Ong MEH, et al. Implementation of Prediction Models in the Emergency Department from an Implementation Science Perspective-Determinants, Outcomes, and Real-World Impact: A Scoping Review. *Ann Emerg Med* 2023;82:22–36. https://doi.org/10.1016/j.annemergmed.2023.02.001. Epub 2023 Mar 14.
75. Di Nucci E, Lee J-Y, Wagner IA. The Rowman & Littlefield Handbook of Bioethics. Rowman & Littlefield; 2022.
76. Suresh H, Guttag J. A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle. *Equity and Access in Algorithms, Mechanisms, and Optimization.* Association for Computing Machinery 2021. Article 17.
77. Frisch A, Reynolds JC, Condle J, Gruen D, Callaway CW. Documentation discrepancies of time-dependent critical events in out of hospital cardiac arrest. Resuscitation 2014;85:1111–4. https://doi.org/10.1016/j.resuscitation.2014.05.002.
78. Sundermann ML, Salcido DD, Koller AC, Menegazzi JJ. Inaccuracy of patient care reports for identification of critical resuscitation events during out-of-hospital cardiac arrest. Am J Emerg Med 2015;33:95–9.
79. Hani M, Christine N, Gerard O, et al. Emergency care surveillance and emergency care registries in low-income and middle-income countries: conceptual challenges and future directions for research. BMJ Glob Health 2019;4:e001442.
80. Volovici V, Syn NL, Ercole A, Zhao JJ, Liu N. Steps to avoid overuse and misuse of machine learning in clinical research. Nat Med 2022,;28:1996–9. https://doi.org/10.1038/s41591-022-01961-6.
81. Andaur Navarro CL, Damen JAA, Takada T, et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review. BMJ 2021;375:n2281. https://doi.org/10.1136/bmj.n2281.
82. Ramspek CL, Jager KJ, Dekker FW, Zoccali C, van Diepen M. External validation of prognostic models: what, why, how, when and where? Clin Kidney J 2021;14:49–58. https://doi.org/10.1093/ckj/sfaa188.
83. Yang C, Kors JA, Ioannou S, et al. Trends in the conduct and reporting of clinical prediction model development and validation: a systematic review. J Am Med Inform Assoc 2022;29:983–9. https://doi.org/10.1093/jamia/ocac002.
84. Ong MEH, Do Shin S, De Souza NNA, et al. Outcomes for out-of-hospital cardiac arrests across 7 countries in Asia: The Pan Asian Resuscitation Outcomes Study (PAROS). Resuscitation 2015;96:100–8.
85. Nichol G, Thomas E, Callaway CW, et al. Regional variation in out-of-hospital cardiac arrest incidence and outcome. JAMA 2008;300:1423–31. https://doi.org/10.1001/jama.300.12.1423.
86. Tagami T, Tanaka H, Shin SD, et al. Impact of population aging on the presentation of out-of-hospital cardiac arrest in the Pan Asian Resuscitation Outcomes Study. *Acute Med Surg* 2020;7:e430. https://doi.org/10.1002/ams2.430.
87. Van Calster B, Steyerberg EW, Wynants L, van Smeden M. There is no such thing as a validated prediction model. BMC Med 2023;21:70. https://doi.org/10.1186/s12916-023-02779-w.
88. Mertens M, King OC, van Putten M, Boenink M. Can we learn from hidden mistakes? Self-fulfilling prophecy and responsible neuroprognostic innovation. J Med Ethics 2022;48:922–8. https://doi.org/10.1136/medethics-2020-106636.
89. De-Arteaga M, Elmer J. Self-fulfilling prophecies and machine learning in resuscitation science. Resuscitation 2023;183:109622. https://doi.org/10.1016/j.resuscitation.2022.10.014.
90. King OC, Mertens M. Self-fulfilling Prophecy in Practical and Automated Prediction. Ethical Theory Moral Pract 2023;26:127–52. https://doi.org/10.1007/s10677-022-10359-9.
91. Detsky ME, Harhay MO, Bayard DF, et al. Discriminative Accuracy of Physician and Nurse Predictions for Survival and Functional Outcomes 6 Months After an ICU Admission. JAMA 2017;317:2187–95. https://doi.org/10.1001/jama.2017.4078.
92. Nolan JP, Sandroni C, Böttiger BW, et al. European Resuscitation Council and European Society of Intensive Care Medicine guidelines 2021: post-resuscitation care. Intensive Care Med 2021;47:369–421. https://doi.org/10.1007/s00134-021-06368-4.
93. Zicari RV, Brusseau J, Blomberg SN, et al. On assessing trustworthy AI in healthcare. Machine learning as a supportive tool to recognize cardiac arrest in emergency calls. Front Hum Dynam 2021;3:673104.
94. Floridi L. Establishing the rules for building trustworthy AI. Nat Mach Intell 2019;1:261–2. https://doi.org/10.1038/s42256-019-0055-y.