Emergency Medicine Australasia

Emergency Medicine Australasia (2018) **30**, 870–874 doi: 10.1111/1742-6723.13145

PERSPECTIVE

Artificial intelligence and machine learning in emergency medicine

Jonathon STEWART, Peter SPRIVULIS and Girish DWIVEDI

Royal Perth Hospital, Perth, Western Australia, Australia

Abstract

Interest in artificial intelligence (AI) research has grown rapidly over the past few years, in part thanks to the numerous successes of modern machine learning techniques such as deep learning, the availability of large datasets and improvements in computing power. AI is proving to be increasingly applicable to healthcare and there is a growing list of tasks where algorithms have matched or surpassed physician performance. Despite the successes there remain significant concerns and challenges surrounding algorithm opacity, trust and patient data security. Notwithstanding these challenges, AI technologies will likely become increasingly integrated into emergency medicine in the coming years. This perspective presents an overview of current AI research relevant to emergency medicine.

Key words: *artificial intelligence, deep learning, emergency medicine, machine learning.*

Introduction

Artificial intelligence (AI) research is undergoing a resurgence, with some predicting that advances will transform society as fundamentally as electricity did.[1] Exponential advances in computing power, data storage and the increasing digitisation of information have contributed to the current growth in AI. AI technologies are likely to impact many aspects of emergency medicine in the near future. This perspective presents an overview of current AI research relevant to emergency medicine (Boxes 1–3).

Clinical image analysis

Advances in image recognition are directly applicable to many domains of medical image analysis. The availability of large labelled image datasets makes image analysis a fertile research domain. Deep learning (DL) techniques have been shown to detect pneumonia on chest X-ray accurately, achieve 3D segmentation of subdural haematomas on brain computed tomography (CT), and assess risk of cerebral aneurysm rupture and score CTs of patients with suspected acute ischaemic stroke as accurately as stroke specialists.[10–13] DL applied to brain magnetic resonance imaging has been used to distinguish patients with a first episode psychosis from controls, and predict lifetime alcohol consumption.[14,15] DL has also been applied to ultrasound (USG); demonstrating high accuracy in detecting abdominal free fluid on FAST (focused assessment with sonography for trauma) scans, classifying abdominal USG images and providing automated analysis of ejection fraction on echocardiogram.[16–18] DL has also enabled novel technologies, such as the use of a microwave based imaging helmet to accurately distinguish between ischaemic and haemorrhagic stroke in the prehospital environment.[19]

Clinical monitoring

Intelligent clinical monitoring may allow for early identification of deteriorating patients. The compensatory reserve index (CRI) uses various machine learning (ML) techniques to reveal subtle changes in finger arterial blood pressure (BP) waveform to predict impending cardiovascular instability.[20] CRI decreases in a linear fashion following voluntary haemorrhage, and the addition of CRI to the monitor screen has been associated with earlier identification of impending instability.[21,22] Another ML model trained to assess vital signs issues cardiovascular instability alerts for intensive care unit stepdown patients on average 17 min and 51 s before onset of a cardiovascular instability event.[23]

Sepsis outcomes improve significantly with earlier detection and treatment. ML applied to heart rate and BP dynamics can independently predict sepsis 4 h prior to clinical onset.[24] Similarly, InSight is an accurate ML based sepsis prediction algorithm that is robust to missing data and uses the change in vital signs over time to also accurately predict sepsis and severe sepsis hours prior to onset.[25] A DL algorithm applied to electrocardiogram (ECG) analysis is able to detect 15 different arrhythmias, sinus rhythm and noise, more accurately than individual cardiologists.[26]

Correspondence: Dr Jonathon Stewart, Royal Perth Hospital, 197 Wellington Street, Perth, WA 6000, Australia. Email: jonathon.e.stewart@gmail.com

Jonathon Stewart, MBBS, MMed (CritCare), Resident Medical Officer; Peter Sprivulis, MBBS, PhD, FACEM, FACHI, Emergency Physician, Chief Clinical Information Officer; Girish Dwivedi, MD, PhD (UK), MRCP (UK), FASE, Cardiologist, Wesfarmers Chair in Cardiology, Consultant Cardiologist, Adjunct Professor of Cardiology, Adjunct Research Professor.

Accepted 21 June 2018

BOX 1. *What is…*

Artificial intelligence (AI)

- Can be defined as 'the theory and development of computer systems able to perform tasks normally requiring human intelligence'.[2]
- Currently, AI remains mostly narrow in its applications. Achieving generalisability is a challenging and unsolved problem, despite some recent progresses.[3]

Machine learning (ML)

- A subfield of AI that uses various methods to automatically detect patterns in data; then use these patterns to make predictions or decisions.[4]
- Models often commence randomly, then improve over time through a process of training.
- Uses learning machines rather than rule-based machines.
- Broadly, supervised ML is trained on labelled data, whereas the data used in unsupervised ML is unlabelled.

Deep learning (DL)

- A type of ML that learns to represents complex and abstract concepts in terms of multiple simpler concepts by passing inputs through a large number of layers of interconnected nonlinear processing units.[5]
- DL has been responsible for numerous recent breakthroughs in fields as diverse as image and speech recognition, speech synthesis, natural language processing and translation.[6–9]

BOX 2. *Possible uses of artificial intelligence in the ED*

- Patient's vitals monitored by algorithms, providing accurate early warning system for cardiovascular instability, sepsis or deterioration.
- Reduction in false alarms.
- Rapid screening, triaging and preliminary diagnosis of imaging and blood test results.
- Outcome predictions and risk stratification that outperform traditional metrics.
- Real time assisted ultrasound analysis.

Any system that attempts to give alerts must deal with the issue of false alarms that currently plague 'intelligent' monitors. A global research competition in 2015 focused on the use of ML to suppress false alarms in the intensive care unit. The winning model was able to suppress 80% of the false alarms while only suppressing 1% of the true alarms.[27] Future monitoring apparatus is likely to be less cumbersome than currently available systems. Novel technologies have been developed that utilise ML techniques to measure heart rate and respiratory rate using only video, without any patient contact.[28] ML has also been applied to cuff-less BP monitoring using a single arm band, and to accurately predict BP using ECG recording.[29,30]

Clinical outcome predictions

AI algorithms are becoming better at predicting the future, often outperforming current clinical scoring systems. Using raw data obtained from over 200 000 patients entire electronic medical records (over 46 billion data points), Rajkomar and Oren *et al.* were able to predict in-hospital mortality with an area under the receiver operating characteristic curve of 0.93–0.94, validated across two sites.[31] ML has been applied to vital sign and medical record data to develop a scoring system for predicting cardiac arrest within 72 h, and has also been used to retrospectively predict defibrillation success for out-of-hospital cardiac arrest.[32,33] ML based '*E*-triage' has demonstrated equivalent or improved identification of patient outcomes compared with the Emergency Severity Index.[34] ML models have been developed that outperform 'TIMI' and 'GRACE' scores for cardiovascular risk, even when trained on medical records with significant missing and noisy data.[35] Liu *et al.* developed ML model that incorporated heart rate variability, a 5 min 12 lead ECG, and vital signs for predicting adverse cardiovascular outcomes.[36] The ML model outperformed TIMI without the need for invasive investigations. ML has also been used to accurately predict 30 day mortality following 'ST elevation myocardial infarction'.[37] Molaei *et al.* developed ML model to identify trauma patients likely to have a positive CT head.[38] Their algorithm outperformed the commonly used Canadian head CT algorithm. ML models have also outperformed the Trauma and Injury Severity Score (TRISS), and are capable of predicting the need for life-saving intervention in trauma patients.[39,40]

Population and social media analysis

AI is likely to become useful in public health and disease surveillance. ML applied to twitter found that Tweets could be a useful supplementary influenza surveillance tool and correlate well with official statistics.[41,42] ML models have also been developed to classify suicide-related communication on twitter.[43] Natural language processing ML models have been applied to emergency medicine clinical documentation to detect influenza.[44] ML may also assist with the detection of novel

BOX 3. *Possible uses of artificial intelligence (non-ED scenarios)*

- Public health surveillance (e.g. influenza surveillance) through automated analysis of ED notes.
- Smart home monitoring allowing for earlier identification of falls or seizures.
- Personal monitoring allowing for early identification of asthma and chronic obstructive pulmonary disease exacerbations.
- Identification of individuals at high risk of suicide from social media analysis.

viral outbreaks, having been used to develop accurate and dynamic forecasting models for Zika virus using Google Trends.[45]

Home monitoring

Patients in the future may present to ED earlier at the prompting of algorithms, and with more information. ML has been used to analyse telemonitored respiratory sounds in order to predict acute exacerbations of chronic obstructive airway disease; detecting 75.8% of exacerbations early, at an average of 5 days in advance of the patient seeking medical attention.[46] Similarly, a ML model can predict a child's asthma control deterioration 1 week ahead of clinical symptoms.[47] Wrist-worn accelerometers can be used to sense seizures, and 'smart carpets' that use a floor based detection system and ML or smartphone audio systems can determine if an elderly fall had occurred.[48–50]

Issues and challenges

Many studies in this perspective are retrospective analyses of datasets. The use of AI technologies in clinical practice will require further validation in prospective studies and randomised control trials prior to widespread clinical adoption. Dealing with noisy data remains a challenge, although there have been good results recently using massive datasets.[31] Lack of universally recognised and utilised reporting/publication guidelines for ML and DL research in medicine adds difficulty to the evaluation of research quality; especially for clinicians without a strong mathematics/computer science background.

DL algorithms are often a 'black box' with the decision logic expressed in thousands (or millions) of numerical weights and biases. This opacity has ethical and legal implications and may foster distrust of AI systems by clinicians and patients. This has been recognised as a key issue for AI adoption and there is ongoing work to develop more human interpretable models, with some image recognition techniques now able to show 'hot spots' containing the pixels in the image that most influence the algorithm. Currently patients are likely to trust a doctor more than a machine, and the idea of their treatment being led by trained algorithmic models may not be accepted.

Humans are excellent at novelty detection; whereas state-of-the-art ML algorithms require significant amounts of labelled data to train and often perform poorly if shown something that deviates too significantly from their training dataset. Accuracy in these supervised models can also be inherently limited by the limits of human performance at labelling the data, and application of the technology is limited to domains where there are large amounts of labelled data available. Providing researchers with access to such large datasets can contravene privacy laws. Notably, the National Health Services (UK) was found to have failed to comply with data protection laws when it provided access to around 1.6 million patient records to Google DeepMind.[51]

Deep learning is one of the most powerful AI techniques currently available; however, as with any technique, there are inherent limitations. Fundamentally, DL maps an input to an output. Despite impressive results it is not indicative of high-level rationality and reasoning.[52] We must be careful to not personify the algorithms and extrapolate intelligence or imagine that these systems truly understand the task that they are performing.[52]

Perhaps the greatest challenge is integration of AI technologies into practice. Despite significant research successes, there remain very few instances of AI algorithms being successfully integrated into daily clinical practice in a complex and critical healthcare systems.

Future trends

AI research has seen significant benefit from the development of large and open data sets that provide high-quality training data, act as a benchmark for comparison between different models, and provide the opportunity for international competition. Ongoing efforts to create large freely accessible high-quality data sets, with international competitions providing financial incentives to solve important problems may benefit emergency medicine.

Current state-of-the-art models make use of architectures hand designed by experts and require the fine tuning of multiple 'hyper parameters'. There is interest in using ML itself to design and optimise ML architectures, and there have been recent advances in this area.[53] Just as computers were first only accessible to a very small group of highly skilled experts but have gone on to become easily usable consumer technology, advances in automating ML may allow easier access to powerful ML techniques to non-specialists.

A significant challenge is determining the best ways to implement AI systems into working clinical environments. It is likely that AI systems will be implemented initially as clinical decision support tools rather than replace clinicians. There are a number of areas of 'low hanging fruit', where current technology already equals or exceeds expert clinician performance that has the potential to make a large difference, including image analysis and clinical deterioration alerting.

Conclusions

Despite limitations, current AI techniques are very capable at solving well defined problems across a wide range of clinical domains. Such systems have the potential to augment many aspects of emergency patient care. The capabilities of AI technology will very likely improve over time, and the integration of such solutions into practice has potential to benefits patients, physicians, and the public through more efficient and accurate delivery of high-quality healthcare.

Competing interests

None declared.

References

1. Lynch S. Andrew Ng: why AI is the new electricity. Stanford News. 2017. [Cited 15 Mar 2018.] Available from URL: https://news.stanford.edu/thedish/2017/03/14/andrew-ng-why-ai-is-the-new-electricity/
2. Oxford. Artificial Intelligence. English Oxford Living Dictionaries. 2018. [Cited 16 May 2018.] Available from URL: https://en.oxforddictionaries.com/definition/artificial_intelligence
3. Silver D, Hubert T, Schrittwieser J *et al*. Mastering chess and Shogi by self-play with a general reinforcement learning algorithm. arXiv:171201815. 2017. [Cited 21 Mar 2017.] Available from URL: http://arxiv.org/abs/1712.01815
4. Murphy KP. *Machine Learning: A Probabilistic Perspective*, Vol. **25**. Cambridge, MA: MIT Press; 2012.
5. Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015; **521**: 436–44.
6. Russakovsky O, Deng J, Su H *et al*. ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* 2015; **115**: 211–52.
7. van den Oord A, Dieleman S, Zen H *et al*. WaveNet: a generative model for raw audio. arXiv: 160903499. 2016. [Cited 21 Mar 2017.] Available from URL: http://arxiv.org/abs/1609.03499
8. Shen J, Pang R, Weiss RJ *et al*. Natural TTS synthesis by conditioning WaveNet on mel spectrogram predictions. arXiv:171205884. 2017. [Cited 21 Mar 2017.] Available from URL: http://arxiv.org/abs/1712.05884
9. Wu Y, Schuster M, Chen Z *et al*. Google's neural machine translation system: bridging the gap between human and machine translation. arXiv:160908144. 2016. [Cited 21 Mar 2017.] Available from URL: http://arxiv.org/abs/1609.08144
10. Rajpurkar P, Irvin J, Zhu K *et al*. CheXNet: radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv: 171105225. 2017. [Cited 21 Mar 2017.] Available from URL: http://arxiv.org/abs/1711.05225
11. Farzaneh N, Soroushmehr SMR, Williamson CA *et al*. Automated subdural hematoma segmentation for traumatic brain injured (TBI) patients. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2017; 2017: 3069–72.
12. Liu J, Chen Y, Lan L *et al*. Prediction of rupture risk in anterior communicating artery aneurysms with a feed-forward artificial neural network. *Eur. Radiol.* 2018; 28: 3268–75.
13. Herweh C, Ringleb PA, Rauch G *et al*. Performance of e-ASPECTS software in comparison to that of stroke physicians on assessing CT scans of acute ischemic stroke patients. *Int. J. Stroke* 2016; **11**: 438–45.
14. Squarcina L, Perlini C, Peruzzo D *et al*. The use of dynamic susceptibility contrast (DSC) MRI to automatically classify patients with first episode psychosis. *Schizophr. Res.* 2015; **165**: 38–44.
15. Guggenmos M, Scheel M, Sekutowicz M *et al*. Decoding diagnosis and lifetime consumption in alcohol dependence from grey-matter pattern information. *Acta Psychiatr. Scand.* 2018; **137**: 252–62.
16. Sjogren AR, Leo MM, Feldman J, Gwin JT. Image segmentation and machine learning for detection of abdominal free fluid in focused assessment with sonography for trauma examinations: a pilot study. *J. Ultrasound Med.* 2016; **35**: 2501–9.
17. Cheng PM, Malhi HS. Transfer learning with convolutional neural networks for classification of abdominal ultrasound images. *J. Digit. Imaging* 2017; **30**: 234–43.
18. Knackstedt C, Bekkers SCAM, Schummers G *et al*. Fully automated versus standard tracking of left ventricular ejection fraction and longitudinal strain. *J. Am. Coll. Cardiol.* 2015; **66**: 1456–66.
19. Persson M, Fhager A, Trefna HD *et al*. Microwave-based stroke diagnosis making global prehospital thrombolytic treatment possible. *I. E.E.E. Trans. Biomed. Eng.* 2014; **61**: 2806–17.
20. Convertino VA, Grudic G, Mulligan J, Moulton S. Estimation of individual-specific progression to impending cardiovascular instability using arterial waveforms. *J. Appl. Physiol.* 2013; **115**: 1196–202.
21. Convertino VA, Howard JT, Hinojosa-Laborde C *et al*. Individual-specific, beat-to-beat trending of significant human blood loss: the compensatory reserve. *Shock* 2015; **44**: 27–32.
22. Muniz GW, Wampler DA, Manifold CA *et al*. Promoting early diagnosis of hemodynamic instability during simulated hemorrhage with the use of a real-time decision-assist algorithm. *J. Trauma Acute Care Surg.* 2013; **75**: 184–9.
23. Guillame-Bert M, Dubrawski A, Wang D, Hravnak M, Clermont G, Pinsky MR. Learning temporal rules to forecast instability in continuously monitored patients. *J. Am. Med. Inform. Assoc.* 2017; **24**: 47–53.
24. Shashikumar SP, Stanley MD, Sadiq I *et al*. Early sepsis detection in critical care patients using multiscale blood pressure and heart rate

dynamics. *J. Electrocardiol.* 2017; **50**: 739–43.
25. Mao Q, Jay M, Hoffman JL *et al.* Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department, general ward and ICU. *BMJ Open* 2018; **8**: e017833.
26. Rajpurkar P, Hannun AY, Haghpanahi M, Bourn C, Ng AY. Cardiologist-level arrhythmia detection with convolutional neural networks. arXiv:170701836. 2017. [Cited 21 Mar 2017.] Available from URL: http://arxiv.org/abs/1707.01836
27. Plesinger F, Klimes P, Halamek J, Jurak P. False alarms in intensive care unit monitors: detection of life-threatening arrhythmias using elementary algebra, descriptive statistics and fuzzy logic. *Comput. Cardiol.* 2015; 42: 281–4.
28. Monkaresi H, Calvo RA, Yan H. A machine learning approach to improve contactless heart rate monitoring using a webcam. *IEEE J. Biomed. Heal. Informatics* 2014; **18**: 1153–60.
29. Zhang Q, Zhou D, Zeng X. Highly wearable cuff-less blood pressure and heart rate monitoring with single-arm electrocardiogram and photoplethysmogram signals. *Biomed. Eng. Online* 2017; **16**: 1–20.
30. Simjanoska M, Gjoreski M, Gams M, Madevska Bogdanova A. Non-invasive blood pressure estimation from ECG using machine learning techniques. *Sensors (Basel)* 2018; **18**: 1160.
31. Rajkomar A, Oren E, Chen K *et al.* Scalable and accurate deep learning with electronic health records. *npj Digital Medicine* 2018; 1: 18.
32. Liu T, Lin Z, Ong ME *et al.* Manifold ranking based scoring system with its application to cardiac arrest prediction: a retrospective study in emergency department patients. *Comput. Biol. Med.* 2015; **67**: 74–82.
33. Shandilya S, Kurz MC, Ward KR, Najarian K. Integration of attributes from non-linear characterization of cardiovascular time-series for prediction of defibrillation outcomes. *PLoS One* 2016; **11**: e0141313.
34. 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**: 565–74.
35. VanHouten JP, Starmer JM, Lorenzi NM, Maron DJ, Lasko TA. Machine learning for risk prediction of acute coronary syndrome. *AMIA Annu. Symp. Proc.* 2014; **2014**: 1940–9.
36. Liu N, Lee MAB, Ho AFW *et al.* Risk stratification for prediction of adverse coronary events in emergency department chest pain patients with a machine learning score compared with the TIMI score. *Int. J. Cardiol.* 2014; **177**: 1095–7.
37. Shouval R, Hadanny A, Shlomo N *et al.* Machine learning for prediction of 30-day mortality after ST elevation myocardial infraction: an acute coronary syndrome Israeli survey data mining study. *Int. J. Cardiol.* 2017; **246**: 7–13.
38. Molaei S, Korley FK, Soroushmehr SMR *et al.* A machine learning based approach for identifying traumatic brain injury patients for whom a head CT scan can be avoided. *38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2016: 2258–61.
39. Sefrioui I, Amadini R, Mauro J, Fallahi A, Gabbrielli M. Survival prediction of trauma patients: a study on US National Trauma Data Bank. *Eur. J. Trauma Emerg. Surg.* 2017; **43**: 805–22.
40. Liu NT, Holcomb JB, Wade CE, Darrah MI, Salinas J. Utility of vital signs, heart rate variability and complexity, and machine learning for identifying the need for lifesaving interventions in trauma patients. *Shock* 2014; **42**: 108–14.
41. Aslam AA, Tsou MH, Spitzberg BH *et al.* The reliability of tweets as a supplementary method of seasonal influenza surveillance. *J. Med. Internet Res.* 2014; **16**: e250.
42. Allen C, Tsou MH, Aslam A, Nagel A, Gawron JM. Applying GIS and machine learning methods to twitter data for multiscale surveillance of influenza. *PLoS One* 2016; **11**: e0157734.
43. Burnap P, Colombo G, Amery R, Hodorog A, Scourfield J. Multi-class machine classification of suicide-related communication on Twitter. *Online Soc. Networks Media* 2017; **2**: 32–44.
44. Ye Y, Wagner MM, Cooper GF *et al.* A study of the transferability of influenza case detection systems between two large healthcare systems. *PLoS One* 2017; **12**: e0174970.
45. Teng Y, Bi D, Xie G *et al.* Dynamic forecasting of zika epidemics using google trends. *PLoS One* 2017; **12**: e0165085.
46. Fernandez-Granero MA, Sanchez-Morillo D, Leon-Jimenez A. Computerised analysis of telemonitored respiratory sounds for predicting acute exacerbations of COPD. *Sensors (Basel)* 2015; **15**: 26978–96.
47. Luo G, Stone BL, Fassl B *et al.* Predicting asthma control deterioration in children. *BMC Med. Inform. Decis. Mak.* 2015; **15**: 1–8.
48. Kusmakar S, Karmakar CK, Yan B, O'Brien TJ, Muthuganapathy R, Palaniswami M. Detection of generalized tonic-clonic seizures using short length accelerometry signal. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS.* 2017: 4566–9.
49. Muheidat F, Tyrer HW. Can we make a carpet smart enough to detect falls? *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS.* 2016: 5356–9.
50. Cheffena M. Fall detection using smartphone audio features. *IEEE J. Biomed. Heal. Informatics* 2016; **20**: 1073–80.
51. ICO. Royal Free - Google DeepMind trial failed to comply with data protection law. Information Commissioner's Office. [Cited 21 Mar 2017.] Available from URL: https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2017/07/royal-free-google-deepmind-trial-failed-to-comply-with-data-protection-law/
52. Chollet F. *Deep Learning with Python*. Shelter Island, NY: Manning; 2017.
53. Zoph B, Le QV. Neural architecture search with reinforcement learning. arXiv:161101578. 2016. [Cited 21 Mar 2017.] Available from URL: http://arxiv.org/abs/1611.01578