

Introduction to Artificial Intelligence for General Surgeons: A Narrative Review

Blanche Lee¹, Nikhil Narsey¹

¹. General Surgery, Toowoomba Hospital, Toowoomba City, AUS

Corresponding author: Blanche Lee, blanche.lee@uqconnect.edu.au

Review began 02/21/2025

Review ended 02/28/2025

Published 03/01/2025

© Copyright 2025

Lee et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

DOI: 10.7759/cureus.79871

Abstract

Artificial intelligence (AI) has rapidly progressed in the last decade and will inevitably become incorporated into trauma and surgical systems. In such settings, surgeons often need to make high-stakes, time-sensitive, and complex decisions with limited or uncertain information. AI has great potential to augment the pre-operative, intra-operative, and post-operative phases of trauma care. Despite the expeditious advancement of AI, many surgeons lack a foundational understanding of AI terminology, its processes, and potential applications in clinical practice. This narrative review aims to educate general surgeons about the basics of AI, highlight its applications in thoraco-abdominal trauma, and discuss the implications of incorporating its use into the Australian health care system. This review found that studies of AI in trauma care have predominantly focused on machine learning and deep learning applied to diagnostics, risk prediction, and decision-making. Other subfields of AI include natural language processing and computer vision. While AI tools have many potential applications in trauma care, current clinical use is limited. Future prospective, locally validated research is required prior to incorporating AI into clinical practice.

Categories: General Surgery, Trauma, Healthcare Technology

Keywords: acute trauma care, artificial intelligence in surgery, computer vision, deep learning artificial intelligence, machine learning models, natural language processing (nlp)

Introduction And Background

Artificial intelligence (AI) is rapidly evolving in the 21st century and powering the Fourth Industrial Revolution [1]. This is made possible as global funding into the AI sector surges, exponential advances are made in computational power and data storage, and as the digitalisation of data in hospitals with electronic medical records (EMR) produce massive datasets [2,3]. AI can be broadly defined as technology that enables machines to perform human-like tasks such as image processing, problem-solving, decision-making, and understanding language. Four subfields of AI include machine learning (ML), computer vision (CV), natural language processing (NLP), and deep learning (DL), also known as deep neural networks (DNNs) [4]. AI marks the intersection between mathematics, statistics, computer and data sciences, linguistics, hardware and software engineering, cognitive science, ethics, and philosophy. AI is positioned to make an exigent entry into the healthcare sector and will inevitably be integrated into the future of medical and surgical services. Currently however, Australia significantly lags behind the rest of the world with respect to funding, research, and clinical applications of AI [5,6]. A *Lancet* scoping review identified 86 randomised controlled trials (RCTs) on AI performed worldwide which were published between 2018 and 2023: none were performed in Australia [7]. While AI technology has been extensively assessed in imaging sectors such as radiology, its adoption in surgery has been slow [8]. Many surgeons are aware of the concept of AI but lack specific knowledge and practical experience of its use. A global survey found general scepticism from surgeons on the use of AI to support clinical decision-making in emergency and trauma surgery [9]. Many surgeons conveyed a low degree of acceptance of AI, and a majority demonstrated a limited understanding of AI applications in surgery [9]. Surgeons preferred classical decision-making aids such as traditional training, clinical guidelines, and multidisciplinary committees [9]. A key concern may lie in the lack of technical understanding of AI, and lack of clarity on legal responsibility of its use. Another international survey of surgeons on their self-rated knowledge of AI terminology found that 24% were unable to define any AI terminology, and 50% were only able to partially define AI concepts such as narrow and general AI, ML, DL, supervised and unsupervised learning, CV, and NLP [10]. The low survey response rate of 2% from members of the World Society of Emergency Surgery increases the risk of selection bias where those interested in AI are more likely to respond [10]. As global AI innovation continues, it is imperative for surgeons to gain a basic understanding of AI. AI terminology can be unfamiliar to surgeons who have received limited or no education on this topic during medical training. The purpose of this narrative review is to educate surgeons about the basics of AI terminology, its clinical applications and implications in thoraco-abdominal trauma, and highlight the potential for those algorithms to transform surgical care in the future.

Introduction to AI

The term 'artificial intelligence' was coined in 1955 by Stanford Professor McCarthy as "the science and engineering of making intelligent machines" [2]. Most AI applications in use fall under the category of narrow AI. This is defined as an intelligent system designed to perform a single task such as image

How to cite this article

Lee B, Narsey N (March 01, 2025) Introduction to Artificial Intelligence for General Surgeons: A Narrative Review. *Cureus* 17(3): e79871. DOI 10.7759/cureus.79871

recognition or translation of language. In contrast, artificial general intelligence refers to an intelligent system which aims to perform actions that mimic human general intelligence, self-learns and reasons, and has human-like cognitive abilities [2]. Figure 1 provides an overview of narrow AI taxonomy.

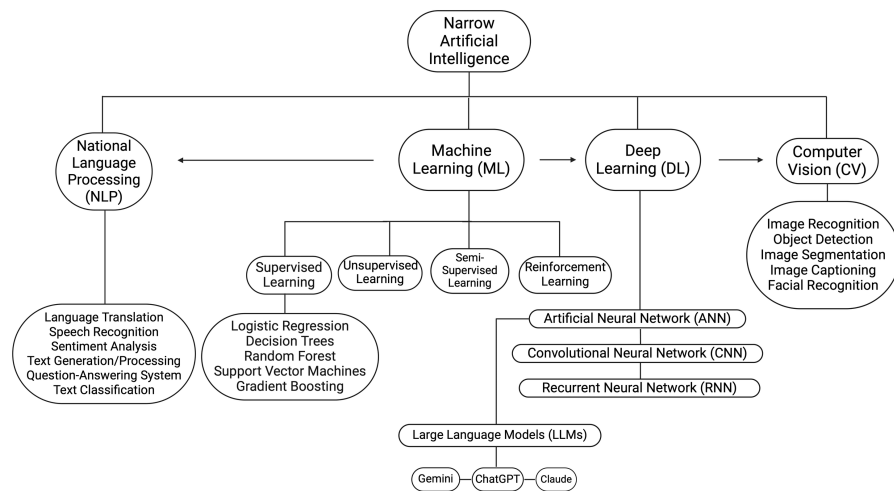


FIGURE 1: Narrow AI Taxonomy

Machine learning

ML is a subfield of AI and uses algorithms to imitate the way humans learn. ML algorithms can make a prediction or classification by making an estimate about patterns in data. The model will evaluate the prediction and improve its accuracy through an iterative optimisation process [11]. ML can be further classified into supervised (labelled data), unsupervised (unlabelled data), and semi-supervised (labelled and un-labelled data). Reinforcement learning is another branch of ML where decision-making is performed by an autonomous agent such as a robot or self-driving car through an interaction between an agent, environment and goal [12]. Within trauma medicine, supervised ML algorithms are the most studied in published literature [13]. Supervised ML is more dependent on human intervention to learn because the input data is labelled, and the model delivers a predefined output. Supervised ML algorithms include logistic regression, decision trees, random forest, support vector machines (SVM), and gradient boosting. The receiver-operating characteristic (ROC) curve is drawn by calculating the true-positive rate (TPR) and false-positive rate (FPR) at selected intervals, and then the TPR is graphed over the FPR [14]. A perfect model at a certain threshold would have a TPR of one and a FPR of zero. Generally, the model with a greater area under the ROC curve (AUROC) is the better one, and this can help compare the performance of two different models if the dataset is balanced.

Deep learning

DL is a subset of ML and is advantageous in that it can be scalable to large data sets. DL can use labelled, unlabelled, or unstructured data in raw forms (e.g. text or images) to develop its own representation for pattern recognition [3,11]. The ability to use unstructured data is particularly useful in a healthcare setting where data is often heterogenous in nature [3]. DL uses neural networks, similar to the neuronal connections in the human brain, to make complex decisions. Other DL networks include recurrent neural networks (RNNs) and convolutional neural networks (CNNs). DL differs from traditional 'non-deep' ML in that it utilises three or more computational layers, normally hundreds or thousands, to train the models. Neural networks are comprised of node layers including an input, one or more hidden layers, and an output layer. These node layers are connected to each other via 'interneurons' and each hold a threshold value and weight [15]. The input and output node layers are termed 'visible' layers. Computational progression through the layers can occur via forward or backpropagation to enable the algorithm to predict and correct for errors. With time, the model become more accurate as it learns. While DL offers many advantages, it requires significant computational power to train algorithms and is costly to scale [15].

Computer vision

CV is a subfield of DL which uses computers and systems to understand digital images and videos, and then acts or make a recommendation based upon these inputs. CV enable computers to see and observe through image detection, object classification, and segmentation. CV utilises a type of DL called CNNs to perform these tasks. While it has greater scalability to large data sets, it is also computationally demanding, requires graphical processing units, and necessitates a vast amount of data to train itself. CV has particularly

promising results in complex, image-heavy diagnostics fields relating to radiology, dermatology, ophthalmology, and pathology [3]. CV has also been used to study the intra-operative phase of care to augment surgical decision-making, support safer surgery, and expand access to surgical data [16]. However, there are currently no widely used CV tools in surgery.

Natural language processing

NLP enables computers and devices to recognise, understand, and generate text and speech by combining ML and DL with computational linguistics [17]. Unstructured text and voice data can be used to train large language models to perform tasks like language translation and sentiment analysis. NLP is commonly used to power search engines, chatbots, and voice-operated digital assistants on smartphones or global positioning system (GPS). The potential benefit of utilising NLP in a healthcare system with EMRs is a rapidly expanding area of research [3]. Within a single hospital admission, ~150,000 pieces of data are generated [3]. This data ranges from pathology results, imaging reports, ward round notes, and patient outcome data. Analysis of such data could provide significant diagnostic, prognostic, and decision-making support for clinicians.

Review

AI in thoracic and abdominal trauma

AI has a wide range of potential applications in trauma systems in the pre-operative, intra-operative, and post-operative phases. However, this review found that AI tools remain largely conceptual and are almost non-existent in clinical settings in Australia [6]. AI-derived ML and DL tools can enhance diagnostics, predictive analytics, risk stratification, patient selection for operative intervention, and guide shared decision-making between the patient and the clinician, triage allocation of resources, and aid in discharge planning. Studies of CV tools remain preliminary but have the potential to enhance surgical education and training through augmented reality and real-time intra-operative assistance [16]. NLP tools could streamline documentation and communication, data collection and auditing, and inform patient consent and follow-up [18].

Pre-operative

Prediction of Haemorrhage Control

Zhang et al. evaluated various ML models in predicting the need for urgent haemorrhage control, which was defined as the activation of a massive transfusion protocol, endovascular embolisation, or damage control surgery [19]. Five ML models were compared, and data was collected from the United States National Trauma Data Bank. Seventeen input variables including vital signs, Glasgow Coma Scale (GCS), injury patterns to the head, neck, torso, abdomen, pelvis, and limbs were collected within four hours of the trauma incident. The XGBoost model performed the best across training and internal and external data validation sets. It achieved an AUROC of 0.88 with an accuracy of over 81% [19].

Thoracic Trauma Imaging

Rib fractures are the most common sequelae of traumatic chest injury and result in significant morbidity and mortality. The rate of missed rib fractures by radiologists is up to 20.9% [20]. ML has been used to reduce clinical time for the detection of fractures in chest radiography and CT with promising results [20-22]. Wang et al. used DNNs to segment ribs and detect rib fractures with an AUROC of 0.94, sensitivity of 86.2%, and specificity of 98.8%, which was on par with attending radiologists [23]. AI misdiagnoses were significantly reduced when a radiologist reviewed the same images [21]. This highlights the utility of AI as a second-reader in trauma imaging and the central role a radiologist plays in verifying the accuracy of AI. Similar studies have been performed to study the use of ML in the detection of pulmonary contusions, pneumothorax, and hemothorax [24].

Abdominal Trauma Imaging

DL demonstrates significant potential in the automated detection of visceral organ injury and lesions through image segmentation and volumetric calculations [25,26]. Cheng et al. used 1302 venous phase CT scans from a single trauma centre in Taiwan to train a model to detect intra-abdominal injury and tested it on 194 CT scans [26]. The spleen injury model had an AUROC of 0.95 and accuracy of 0.94, the liver model had an AUROC of 0.87 and accuracy of 0.82, and the kidney model had an AUROC of 0.87 and accuracy of 0.95. While the individual organ models demonstrated reasonably good accuracy, the whole image model lacked the capacity to identify specific injured organs reliably. Among 72 patients, 19 cases (26.4%) of solid organ injuries were missed [26]. Shen et al. trained a DL model to identify multiple organ injuries simultaneously with a database of 4000 CT images from 14 countries [25]. This model was trained to detect extravasation of contrast media to identify bleeding, as well as simultaneous detection of injuries to the liver, spleen, kidney, or bowel. The accuracy of the model was best for renal injuries with an AUROC of 0.88 and accuracy of 0.93. Liver injury detection had an AUROC of 0.82 and accuracy of 0.87, while splenic injury had an AUROC of 0.85 and an accuracy of 0.77. In contrast, bowel injuries had a low positive predictive value (PPV) and poor sensitivity despite having an AUROC of 0.83 and accuracy of 0.98. This suggests the model can be used to reliably rule out intestinal injury, but is poor at flagging positive cases [25]. Similarly, contrast blush could be reliably ruled out, but the model had low PPV and poor sensitivity.

ICU Admission and Length of Stay

ML has been used to predict ICU admission and length of stay following thoraco-abdominal trauma using clinical parameters and CT imaging findings [27]. Clinical parameters included age, sex, vital signs, haemoglobin, haematocrit, Revised Trauma Score, GCS, and lactate. Imaging findings included the presence or absence of solid organ injury as per the American Association for the Surgery of Trauma Grading, intra-peritoneal fluid, thoracic aortic injuries, rib fractures, pelvic fractures, and haemothorax among other injuries. The AUROC of the SVM models for ICU admission and predicting length of stay was 0.87 and 0.80, respectively [27]. The predictions of outcomes were superior when combining clinical and CT imaging parameters.

Intra-operative

Surgical Data Sciences

Surgery demands a high physical and cognitive load, and any error can translate to significant harm for patients. While the effectiveness of surgery has increased in the last few decades, there is still significant potential to improve surgical safety. Surgical Data Science has emerged as a novel discipline where digital data can be analysed and audited to improve surgical care [28]. Surgical big data incorporates studying and recording the operating room (OR) environment, intra-operative videos, surgical device logs, robot kinematics, patient's physiological parameters, and measures of surgeon's heart rate variability as a proxy of cognitive workload [28]. This vast amount of data can be analysed offline or in real-time with DL and ML algorithms to inform and enhance surgical care.

Surgical Education, Training, and Feedback

The use of a surgical 'black box' and 'surgical control towers' has been proposed by Mascagni et al. as a method to systematically capture, analyse, and provide feedback to assist the operator [28]. These OR black boxes allow for post-operative analysis of near-misses or errors, and draw from parallels with their use in aviation and Formula One. The surgical black box is a multi-port device which captures synchronised audio-visual, environmental, and physiological data from the OR. Similarly, OR control towers, analogous to air traffic control, could live stream data in and out of the OR to permit real-time analysis and provide rapid feedback to the training or operating surgeon [28]. CV technology allows algorithms to interpret and analyse intra-operative visual data to support surgeon decision-making, enable safer surgery, and could democratise surgical care through international collaboration and education [16]. CV tools are still being refined in their ability to define intra-operative anatomy, tissue and dissection planes, detect a haemorrhagic event, identify specific surgical instruments, and to predict the next procedural step [8]. While there are many potential use cases, there are currently no widely used CV tools in clinical practice.

Prediction of Intra-Operative Hypotension and Transfusion

The Hypotension Prediction Index (HPI) is an ML-derived early warning system, which has been extensively studied. Two RCTs performed in elective non-cardiac surgery patients found that the use of the HPI resulted in less intra-operative hypotension compared to standard care [29,30]. Notably, all studies were performed in patients undergoing elective surgery and their results cannot be generalised to trauma patients. Zapf et al. used an ML model to predict the need for intra-operative red blood cell (RBC) transfusion based on pre-operative variables [31]. The LightGBM model outperformed six other ML algorithms with an accuracy of 76.1% and sensitivity of 91.2% [31]. Predictive factors for RBC transfusion included the type of procedure, haemoglobin level, and was surgeon- and anaesthesiologist-dependent [31]. The use of individual surgeon and anaesthesiologist identifiers limits the generalisability of these results to other institutions. It also highlights the importance of training ML models with institution-specific or local data sets.

Post-operative

Prediction of Survival and Complications

The power of AI algorithms lies in their ability to intelligently adjust the weight of input variables, capture non-linear relationships, and identify patterns in data sets. In contrast, traditional risk assessments such as the American Society of Anesthesiologists (ASA) score and Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity (POSSUM) often assume an additive and linear correlation [32]. The use of subjective scores like the ASA can result in misclassification, and conventional regression analysis used in POSSUM scores can over- or under-estimate risk, especially at extreme ranges of variables [33]. Novel smartphone-based AI risk prediction tools include Predictive Optimal Trees in Emergency Surgery Risk (POTTER), Trauma Outcome Predictor (TOP), and MySurgeryRisk [34-36]. POTTER uses a decision tree ML algorithm to predict post-operative morbidity and mortality after emergency surgery based on the American College of Surgeons (ACS) National Surgical Quality Improvement Program (NSQIP) data. POTTER was able to predict mortality and morbidity with great accuracy and surpassed surgeon gestalt [37,38]. POTTER's c-statistic was 0.91 and outperformed ASA, Emergency Surgery Score, and the ACS-NSQIP calculator for predicting mortality [33]. TOP utilises optimal classification trees to predict mortality with an AUROC of 0.94 and 0.88 for penetrating and blunt injuries respectively, but it was less accurate at predicting complications such as acute kidney injury (AKI), pulmonary embolism, acute respiratory distress syndrome, deep venous thrombosis, and sepsis with an AUROC ranging between 0.69 to 0.84 [34]. MySurgeryRisk is based on a random forest model, which predicts 30- and 90-day mortality, as well as post-operative complications. It uses 135 input parameters and has an AUROC of 0.84 for predicting

mortality [36]. MySurgeryRisk was better at predicting complications such as AKI, cardiovascular, delirium, prolonged mechanical ventilation, sepsis, wound, and venous thromboembolism compared to TOP with an AUROC ranging between 0.81 and 0.91 [36]. The platform display provides an output in real time and includes the risk of each post-operative complication, the top three risk factors for each complication, and patterns of complications for individual surgeons [36].

Discharge Prediction

Understanding patient disposition for discharge can aid bed management and pre-emptive discharge planning. Kovoor et al. demonstrated that a ML-derived random forest model called the Adelaide Score can successfully predict discharge at 12 and 24 hours in general surgical patients with an AUROC of 0.85 and 0.84, respectively [39]. Input parameters included vital signs, pain scores, recent bowel movements, Charlson Comorbidity Index scores, and blood results, which were used to generate a score between 0 and 100, which represents a patient's readiness for discharge following general surgery. The limitations of the Adelaide score include its lack of integration of socio-cultural factors in determining patient safety and readiness for discharge, and the requirement of an EMR system to make the score calculation practical for daily use [39].

Discussion

Although there is great potential for AI in trauma systems, a multitude of challenges and risks exist. It will be imperative to have appropriate regulation, awareness of bias, and adequate education of the patient and the surgeon. National governance from external regulators, alongside the Royal Australian College of Surgeons, will be required to evaluate these novel AI tools to protect patients from profit-driven corporations or commercial entities [40]. AI technologies that are classified as Software as a Medical Device will need approval and regulation from the Therapeutic Goods Administration [41]. Additionally, revision of current malpractice guidelines will be required to define how responsibility is determined if both AI and human factors contribute to adverse events or outcomes [42].

Surgical decision-making often has time-sensitive and significant consequences, and the use of AI can introduce automation bias and safety concerns. Algorithmic outputs lack the nuances of human decision-making and require critical appraisal when applied to individual patient care. Clinicians should be cautious to prevent an over-reliance on AI-based tools. Other potential biases inherent in training data may contribute to or perpetuate existing inequalities in age, race, social-economic status, or sex, and must be scrutinised [43]. The performance of individual AI applications is heavily dependent on the characteristics, quality, and completeness of available datasets. The generalisability of outcome data may be limited by whether local or global datasets are used. The prevalence and incidence of blunt versus penetrating trauma, such as gunshot wounds, can vary significantly around the world. For example, AI models which have derived predictive, diagnostic, or decision-aid tools from databases from level I trauma centres in the United States may not be relevant to a regional trauma centre in Australia.

Finally, it is important to consider public opinion and attitude on the incorporation of AI into healthcare [40]. For many, AI is referred colloquially as a 'black box' with its inner workings obscure and difficult to explain. The clinician has a duty to counsel the patient about the factors considered by the AI decision-support tool and its limitations. Even amongst clinicians, it may be challenging to educate the patients on the conclusions or recommendations from AI tools outside of detailing the inputs and outputs of algorithms [44]. Clear policy is required to ensure patient privacy and data security. Moreover, patients need to be informed that their data could be retained or used for continued training of AI models, and opt-out options should be offered where possible [40].

Conclusions

Despite the growing presence of AI in health care, many surgeons have a limited understanding about basic AI terminology and how these technologies work. This review highlights the multifaceted potential of AI to improve patient outcomes and streamline the pre-operative, intra-operative, and post-operative phases of thoraco-abdominal trauma care. Many trials have focused on assessing ML and DL in predictive analytics, diagnostics, and decision-making with promising results. There is evidence that these algorithms can complement, augment, and in some cases, surpass traditional risk prediction tools. However, current clinical applications of AI are limited. Surgeons should stay informed on the evolving role of AI so that it can be safely and effectively integrated into clinical care. Future prospective, locally validated research, rooted in Australian data sets, is required.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Blanche Lee, Nikhil Narsey

Acquisition, analysis, or interpretation of data: Blanche Lee

Drafting of the manuscript: Blanche Lee

Critical review of the manuscript for important intellectual content: Blanche Lee, Nikhil Narsey

Supervision: Nikhil Narsey

Disclosures

Conflicts of interest: In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

1. What are Industry 4.0, the Fourth Industrial Revolution, and 4IR? . (2022). <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-are-industry-4-0-the-fourth-industrial-revolution...>
2. "The AI Index 2024 Annual Report," AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. (2024). https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_2024_AI-Index-Report.pdf.
3. Esteva A, Robicquet A, Ramsundar B, et al.: A guide to deep learning in healthcare . *Nat Med.* 2019, 25:24-9. [10.1038/s41591-018-0316-z](https://doi.org/10.1038/s41591-018-0316-z)
4. Hashimoto DA, Rosman G, Rus D, Meireles OR: Artificial intelligence in surgery: promises and perils. *Ann Surg.* 2018, 268:70-6. [10.1097/SLA.0000000000002695](https://doi.org/10.1097/SLA.0000000000002695)
5. A National Policy Roadmap for Artificial Intelligence in Healthcare. Australian Alliance for Artificial Intelligence in Healthcare, NSW, Australia. (2021). https://aihealthalliance.org/wp-content/uploads/2021/12/AAAIH_Roadmap_1Dec2021_FINAL.pdf.
6. van der Vegt A, Campbell V, Zuccon G: Why clinical artificial intelligence is (almost) non-existent in Australian hospitals and how to fix it. *Med J Aust.* 2024, 220:172-5. [10.5694/mja2.52195](https://doi.org/10.5694/mja2.52195)
7. Han R, Acosta JN, Shakeri Z, Ioannidis JP, Topol EJ, Rajpurkar P: Randomised controlled trials evaluating artificial intelligence in clinical practice: a scoping review. *Lancet Digit Health.* 2024, 6:e367-73. [10.1016/S2589-7500\(24\)00047-5](https://doi.org/10.1016/S2589-7500(24)00047-5)
8. Varghese C, Harrison EM, O'Grady G, Topol EJ: Artificial intelligence in surgery. *Nat Med.* 2024, 30:1257-68. [10.1038/s41591-024-02970-3](https://doi.org/10.1038/s41591-024-02970-3)
9. Cobianchi L, Piccolo D, Dal Mas F, et al.: Surgeons' perspectives on artificial intelligence to support clinical decision-making in trauma and emergency contexts: results from an international survey. *World J Emerg Surg.* 2023, 18:1. [10.1186/s13017-022-00467-3](https://doi.org/10.1186/s13017-022-00467-3)
10. De Simone B, Abu-Zidan FM, Gumbs AA, et al.: Knowledge, attitude, and practice of artificial intelligence in emergency and trauma surgery, the ARIES project: an international web-based survey. *World J Emerg Surg.* 2022, 17:10. [10.1186/s13017-022-00413-3](https://doi.org/10.1186/s13017-022-00413-3)
11. IBM. AI vs. machine learning vs. deep learning vs. neural networks: What's the difference? . (2023). <https://www.ibm.com/think/topics/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks>.
12. What is reinforcement learning? IBM. (2024). <https://www.ibm.com/topics/reinforcement-learning>.
13. Hunter OF, Perry F, Salehi M, Bandurski H, Hubbard A, Ball CG, Morad Hameed S: Science fiction or clinical reality: a review of the applications of artificial intelligence along the continuum of trauma care. *World J Emerg Surg.* 2023, 18:16. [10.1186/s13017-022-00469-1](https://doi.org/10.1186/s13017-022-00469-1)
14. Classification: ROC and AUC. <https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc..>
15. IBM. What is deep learning?. <https://www.ibm.com/topics/deep-learning>.
16. Mascagni P, Alapatt D, Sestini L, et al.: Computer vision in surgery: from potential to clinical value . *NPJ Digit Med.* 2022, 5:163. [10.1038/s41746-022-00707-5](https://doi.org/10.1038/s41746-022-00707-5)
17. What is NLP (natural language processing)? IBM. (2024). <https://www.ibm.com/think/topics/natural-language-processing>.
18. Le KD, Tay SB, Choy KT, Verjans J, Sasanelli N, Kong JC: Applications of natural language processing tools in the surgical journey. *Front Surg.* 2024, 11:1403540. [10.3389/fsurg.2024.1403540](https://doi.org/10.3389/fsurg.2024.1403540)
19. Zhang J, Jin Z, Tang B, Huang X, Wang Z, Chen Q, He J: Enhancing trauma care: a machine learning approach with XGBoost for predicting urgent hemorrhage interventions using NTDB data. *Bioengineering (Basel).* 2024, 11:768. [10.3390/bioengineering11080768](https://doi.org/10.3390/bioengineering11080768)
20. Yao L, Guan X, Song X, et al.: Rib fracture detection system based on deep learning . *Sci Rep.* 2021, 11:23513. [10.1038/s41598-021-03002-7](https://doi.org/10.1038/s41598-021-03002-7)
21. Zhou Q, Qin P, Luo J, Hu Q, Sun W, Chen B, Wang G: Evaluating AI rib fracture detections using follow-up CT scans. *Am J Emerg Med.* 2023, 72:34-8. [10.1016/j.ajem.2023.07.018](https://doi.org/10.1016/j.ajem.2023.07.018)
22. Lee K, Lee S, Kwak JS, Park H, Oh H, Koh JC: Development and validation of an artificial intelligence model for detecting rib fractures on chest radiographs. *J Clin Med.* 2024, 13:3850. [10.3390/jcm13133850](https://doi.org/10.3390/jcm13133850)
23. Wang S, Wu D, Ye L, Chen Z, Zhan Y, Li Y: Assessment of automatic rib fracture detection on chest CT using a deep learning algorithm. *Eur Radiol.* 2023, 33:1824-34. [10.1007/s00330-022-09156-w](https://doi.org/10.1007/s00330-022-09156-w)
24. Zhao T, Meng X, Wang Z, et al.: Diagnostic evaluation of blunt chest trauma by imaging-based application

- of artificial intelligence. *Am J Emerg Med.* 2024, 85:35-45. [10.1016/j.ajem.2024.08.019](https://doi.org/10.1016/j.ajem.2024.08.019)
25. Shen X, Zhou Y, Shi X, et al.: The application of deep learning in abdominal trauma diagnosis by CT imaging. *World J Emerg Surg.* 2024, 19:17. [10.1186/s13017-024-00546-7](https://doi.org/10.1186/s13017-024-00546-7)
 26. Cheng CT, Lin HH, Hsu CP, et al.: Deep learning for automated detection and localization of traumatic abdominal solid organ injuries on CT scans. *J Imaging Inform Med.* 2024, 37:1113-25. [10.1007/s10278-024-01038-5](https://doi.org/10.1007/s10278-024-01038-5)
 27. Staziaki PV, Wu D, Rayan JC, et al.: Machine learning combining CT findings and clinical parameters improves prediction of length of stay and ICU admission in torso trauma. *Eur Radiol.* 2021, 31:5434-41. [10.1007/s00330-020-07534-w](https://doi.org/10.1007/s00330-020-07534-w)
 28. Mascagni P, Padoy N: OR black box and surgical control tower: Recording and streaming data and analytics to improve surgical care. *J Visc Surg.* 2021, 158:S18-25. [10.1016/j.jvisurg.2021.01.004](https://doi.org/10.1016/j.jvisurg.2021.01.004)
 29. Tsoumpa M, Kyttari A, Matiatou S, et al.: The use of the hypotension prediction index Integrated in an algorithm of goal directed hemodynamic treatment during moderate and high-risk surgery. *J Clin Med.* 2021, 10:5884. [10.3390/jcm10245884](https://doi.org/10.3390/jcm10245884)
 30. Wijnberge M, Geerts BF, Hol L, et al.: Effect of a machine learning-derived early warning system for intraoperative hypotension vs standard care on depth and duration of intraoperative hypotension during elective noncardiac surgery: the HYPE randomized clinical trial. *JAMA.* 2020, 323:1052-60. [10.1001/jama.2020.0592](https://doi.org/10.1001/jama.2020.0592)
 31. Zapf MA, Fabbri DV, Andrews J, Li G, Freundlich RE, Al-Droubi S, Wanderer JP: Development of a machine learning model to predict intraoperative transfusion and guide type and screen ordering. *J Clin Anesth.* 2023, 91:111272. [10.1016/j.jclinane.2023.111272](https://doi.org/10.1016/j.jclinane.2023.111272)
 32. Hassan AM, Rajesh A, Asaad M, Nelson JA, Coert JH, Mehrara BJ, Butler CE: Artificial intelligence and machine learning in prediction of surgical complications: current state, applications, and implications. *Am Surg.* 2023, 89:25-30. [10.1177/00031348221101488](https://doi.org/10.1177/00031348221101488)
 33. Bertsimas D, Dunn J, Velmahos GC, Kaafarani HM: Surgical risk is not linear: derivation and validation of a novel, user-friendly, and machine-learning-based predictive OpTimal trees in emergency surgery risk (POTTER) calculator. *Ann Surg.* 2018, 268:574-83. [10.1097/SLA.0000000000002956](https://doi.org/10.1097/SLA.0000000000002956)
 34. Maurer LR, Bertsimas D, Bouardi HT, et al.: Trauma outcome predictor: an artificial intelligence interactive smartphone tool to predict outcomes in trauma patients. *J Trauma Acute Care Surg.* 2021, 91:93-9. [10.1097/TA.00000000000003158](https://doi.org/10.1097/TA.00000000000003158)
 35. Maurer LR, Chetlur P, Zhuo D, et al.: Validation of the AI-based predictive OpTimal trees in emergency surgery risk (POTTER) calculator in patients 65 years and older. *Ann Surg.* 2023, 277:e8-e15. [10.1097/SLA.00000000000004714](https://doi.org/10.1097/SLA.00000000000004714)
 36. Ren Y, Loftus TJ, Datta S, et al.: Performance of a machine learning algorithm using electronic health record data to predict postoperative complications and report on a mobile platform. *JAMA Netw Open.* 2022, 5:e2211973. [10.1001/jamanetworkopen.2022.11973](https://doi.org/10.1001/jamanetworkopen.2022.11973)
 37. El Hechi MW, Maurer LR, Levine J, et al.: Validation of the artificial intelligence-based predictive optimal trees in emergency surgery risk (POTTER) calculator in emergency general surgery and emergency laparotomy patients. *J Am Coll Surg.* 2021, 232:912-919.e1. [10.1016/j.jamcollsurg.2021.02.009](https://doi.org/10.1016/j.jamcollsurg.2021.02.009)
 38. El Moheb M, Gebran A, Maurer LR, et al.: Artificial intelligence versus surgeon gestalt in predicting risk of emergency general surgery. *J Trauma Acute Care Surg.* 2023, 95:565-72. [10.1097/TA.0000000000004030](https://doi.org/10.1097/TA.0000000000004030)
 39. Kovoor JG, Bacchi S, Gupta AK, et al.: The Adelaide Score: an artificial intelligence measure of readiness for discharge after general surgery. *ANZ J Surg.* 2023, 93:2119-24. [10.1111/ans.18546](https://doi.org/10.1111/ans.18546)
 40. Kovoor JG, Bacchi S, Sharma P, et al.: Artificial intelligence for surgical services in Australia and New Zealand: opportunities, challenges and recommendations. *Med J Aust.* 2024, 220:234-7. [10.5694/mja2.52225](https://doi.org/10.5694/mja2.52225)
 41. Therapeutic Goods Administration. (2022). <https://www.tga.gov.au/resources/guidance/understanding-regulation-software-based-medical-devices>.
 42. Holm S, Stanton C, Bartlett B: A new argument for no-fault compensation in health care: the introduction of artificial intelligence systems. *Health Care Anal.* 2021, 29:171-88. [10.1007/s10728-021-00430-4](https://doi.org/10.1007/s10728-021-00430-4)
 43. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH: Ensuring fairness in machine learning to advance health equity. *Ann Intern Med.* 2018, 169:866-72. [10.7326/M18-1990](https://doi.org/10.7326/M18-1990)
 44. London AJ: Artificial intelligence and black-box medical decisions: accuracy versus explainability. *Hastings Cent Rep.* 2019, 49:15-21. [10.1002/hast.973](https://doi.org/10.1002/hast.973)