Journal of Surgery and Trauma 2025; 13(2):46-54.
jsurgery.bums.ac.ir

REVIEW ARTICLE

Artificial Intelligence and Machine Learning in General Surgery: A Narrative Review of Transformative Technologies

Abolfazl Torabi[1], Mohammad Ehsan Bayatpoor[2], Nasim Nouri[1], Maryam Abbasi[2], Nasser Malekpour Alamdari [3]

[1] Medical Student, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

[2] General Surgery Department, Modarres Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran

[3] Critical Care Quality Improvement Research Center, Shahid Modarres Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Received: September 25, 2024 ***Revised: March 16, 2025*** ***Accepted: March 20, 2025***

Abstract

As computational power and data science continue to advance at an unprecedented pace, their influence is reshaping various scientific fields, including medicine. While machine learning (ML) has already made substantial strides in diagnostic areas, such as radiology and pathology, its role in surgery is an emerging frontier. This narrative review examines the current literature on artificial intelligence (AI) and ML applications in general surgery, with a particular focus on their ability to support clinical decision-making, streamline surgical workflows, and improve patient outcomes. Key topics explored include predicting discharge dates, assessing preoperative risk for both elective and emergency surgeries, and the innovative use of AI in resident education and simulation training. By evaluating these developments, the practical challenges, ethical concerns, and future prospects of integrating AI into surgical practice were discussed. Ultimately, this review highlights the transformative potential of AI and ML in surgery, suggesting that these technologies will play a key role in enhancing care quality and the professional growth of surgeons.

Key words: Artificial Intelligence (AI), General Surgery, Machine Learning (ML)

Introduction

The rapid progression of computational power and data science has sparked a revolution across multiple scientific fields, with medicine standing at the forefront of this transformation. Among the most impactful innovations are machine learning (ML) and artificial intelligence (AI) systems designed to replicate human intelligence in analyzing data, recognizing patterns, and making predictions. These technologies are already making significant strides in diagnostic fields like radiology and pathology, where they have improved accuracy and efficiency. However, their role in surgery remains an emerging and evolving frontier with vast untapped potential.

AI and ML are becoming increasingly vital tools in enhancing clinical decision-making, patient care, and surgical education. For instance, these technologies help surgeons predict optimal discharge dates, assess preoperative risks for both elective and emergency surgeries, and advance resident training through simulation and decision-support systems. Despite their promising

©2025Journal of Surgery and Trauma
Tel: +985632381214
Fax: +985632440488
Po Bax 97175-379
Email: jsurgery@bums.ac.ir

Correspondence to:
Nasser Malekpour Alamdari, Critical Care Quality Improvement Research Center, Shahid Modarres Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran
Telephone Number: +989124730423
Email: nassermalekpour@sbmu.ac.ir

[DOI: 10.61186/jsurgtrauma.13.2.46]

capabilities, the integration of AI and ML into general surgery brings forth important questions regarding their current applications, future prospects, and the challenges and limitations they may present (1-8).

This narrative review focuses on the application of AI and ML within general surgery—a multifaceted field that spans subspecialties like gastrointestinal, endocrine, and trauma surgery. While the breadth of general surgery is vast, this review highlights key areas where these technologies have shown considerable potential, including clinical decision-making, workflow optimization, and education. By addressing existing gaps in the literature and examining the evolving role of these innovations, the present study aims to offer a clearer perspective on how AI and ML are shaping the present and future landscape of surgical practice.

The review will not only assess the current state of AI and ML in surgery but will also consider their future implications, ethical concerns, and challenges associated with their integration. By examining the dynamic intersection of surgery and data-driven technologies, we hope to shed light on their transformative potential and profound impact on the future of surgical care.

Methods

This narrative review adhered to a structured approach in examining the use of AI and ML in general surgery. We conducted a thorough literature search across electronic databases, including PubMed, Scopus, and Google Scholar, to identify pertinent articles. Boolean operators, such as "AND," "OR," and "NOT," were employed to refine the results. Specific search queries included combinations, such as "artificial intelligence AND general surgery," "machine learning OR artificial intelligence AND preoperative planning," and "surgical training NOT diagnostics." The final search was performed on December 15, 2024.

The search period covered studies published between 2013 and 2024, a timeframe selected to capture the most relevant advancements in AI and ML. This focus reflects the significant rise of these technologies in healthcare and surgery over the past decade. Earlier studies were included if they offered foundational insights or key methodologies that contributed to the application of AI/ML in surgery.

To ensure the rigor of the review, two independent reviewers evaluated the identified studies for eligibility and quality. Disagreements were resolved through discussion or by consulting a third reviewer. Studies were included if they met the following criteria: peer-reviewed articles, reviews, clinical studies, or expert opinions focusing on AI/ML in general surgery, with particular emphasis on preoperative planning, postoperative outcomes, surgical risk prediction, and resident training. Articles unrelated to general surgery or those focused solely on diagnostic fields without direct implications for surgical practice were excluded.

The quality of the included studies was assessed based on their methodological rigor, relevance to general surgery, and transparency in reporting AI/ML techniques. Emphasis was placed on the appropriateness of the models used, the robustness of the data sets, and the validity of the outcomes in real-world surgical environments. This systematic and evidence-based approach was designed to provide a comprehensive understanding of how AI and ML are reshaping the field of general surgery.

Results

Predictive modeling for surgical site infections (SSI)

In a retrospective cohort study, a predictive model achieved an Area under the ROC curve (AUC) of 0.84 for identifying the type of surgical site infection (SSI) and an AUC of 0.74 for predicting the postoperative week of SSI development. Unlike previous models, which primarily focused on identifying SSI development, this study also explored intraoperative characteristics as potential contributors (1-6). Distinguishing between superficial and deep/organ space infections is crucial due to their impact on patient outcomes, with risk factors varying between these types. For instance, irradiation therapy is associated with deep/organ space SSIs, while body mass index (BMI) is linked to incisional infections. The study underscores the importance of implementing systematic measures in SSI prevention, particularly in areas, such as antibiotic administration and intraoperative normothermia (9).

Predicting surgery cancellations using ml

Liu et al. employed ML techniques, primarily using electronic health record (EHR) data, to predict surgery cancellations with impressive accuracy (AUC up to 0.78) at two different campuses. Logistic regression models, especially a gradient-boosted variant, showed the highest predictive power. What is remarkable is that models trained on one campus performed nearly as well on the other despite differences in workload and patient demographics (10).

The ML models excelled in predicting "no-show" and nothing by mouth (NPO) violation cancellations compared to patient illness and patient/family refusal. They also identified important predictors for cancellations, including patient age, healthcare payer, and timing related to the intervention program. Notably, certain predictors differed between the two campuses, likely reflecting variations in patient populations.

The study's generalizability between the two campuses demonstrates the potential for ML methodology to be applied in hospitals with EHR systems (10).

Risk prediction for surgery cancellations

Regarding the assessment of the surgery risk, Luo et al. (2020) illustrated the potential to predict surgeries at high risk of cancellation. The ML models consistently achieve AUCs above 0.6 in the test set, with the best performance at 0.682 for Random Forest (RF) with oversampling. These models maintain stable performance, with a slight difference of under 0.04 between their upper and lower bounds. The ML models in this study exhibit high Negative Predictive Values (NPV), approximately 0.9, indicating that 90% of surgeries labeled as low risk are not canceled, which effectively filters out most negative cases, narrowing down the set of surgeries with potential issues (11).

Different sampling methods can be employed to adjust the model's performance, with oversampling and undersampling, which affect sensitivity and specificity compared to the original method. It is noteworthy that the practitioners' preferences play a role in determining the model's performance (11).

AI-based mortality and morbidity prediction in emergency surgery

Performance of the Predictive Optimal Trees in Emergency Surgery Risk (POTTER) tool, an AI-based calculator, in the context of emergency general surgery (EGS), is the topic of another study investigating the role of AI in predicting mortality and morbidity of surgeries (12). The mentioned article evaluated the effectiveness of the POTTER tool in predicting the 30-day outcomes of patients undergoing EGS. The tool accurately predicted both mortality and morbidity in this patient population. As a smartphone application, POTTER served as a valuable bedside tool for counseling patients and family members before EGS procedures. It could also be used as a risk adjustment tool for assessing the quality of EGS care (12).

The EGS represents a significant public health concern, making up a substantial portion of hospital admissions and surgical mortality (13). The nonlinearity of surgical risk in EGS patients requires advanced technology to detect and understand these complex patterns (12, 14).

Unlike some ML methods, POTTER is transparent and interpretable. It relies on the Optimal Classification Tree (OCT) technology, which can achieve high accuracy without sacrificing interpretability. This transparency allows physicians to understand how the risk estimates are calculated. The tool could assist in improving the benchmarking of EGS care quality and guiding discussions with patients and their families. Moreover, it may help in the postoperative triaging of high-risk patients (12).

Concerning the prediction of only a 30-day outcome, El Hechi et al. suggested the integration of OCT algorithms into a hospital's electronic medical record (EMR) as a way to address limitations and enhance the accuracy and applicability of risk predictions. Finally, it seems that POTTER is a valuable tool for predicting outcomes in EGS patients and has the potential to improve decision-making and quality benchmarking in this surgical specialty (12).

Another study conducted by Gao et al. represents the first attempt to develop an ML algorithm tailored to predict mortality following EGS. The ML algorithm exhibited exceptional performance in predicting EGS mortality when compared to existing risk-prediction models, such as the American Society of Anesthesiologists (ASA) classification, American College of Surgeons Surgical Risk Calculator (ACS-SRC), and Multi-Feature Learning (MFL) (15). The model achieved superior accuracy, as indicated by various metrics, including AUC, sensitivity, specificity, Positive Predictive Values (PPV), and NPV. To optimize the ML model and mitigate issues related to overfitting and parameter optimization, they employed an ensembled ML classifier, which combined two distinct ML algorithms, Gradient Boosting Machine (GBM) and Multi-layer Perceptron Artificial Neural Network (MLP ANN) (15). This approach maximized the benefits of each algorithm while minimizing their limitations. Parameter optimization was conducted using a grid search technique to determine the best combination of hyperparameters for accurate mortality prediction (15).

ML for automated surgical skill assessment

In the context of robot-assisted surgery, there is a growing interest in automating assessment and feedback using ML, as it allows for the transparent capture and analysis of high-quality surgical motion data (16). One key step is automated surgical activity recognition, which involves identifying surgical activities and their timing, even for new

[DOI: 10.61186/jsurgtrauma.13.2.46]

trainees without prior annotated data.

It is noteworthy that different research has focused on recognizing activities and gestures using different techniques like hidden Markov models (HMMs) and conditional-random-field (CRF) based methods (17). DiPietro et al. addressed both gestures and maneuvers, which are more complicated, such as suture throw and knot tying, and administered recurrent neural networks (RNNs) to evaluate their effectiveness.

Hyperparameter analysis in surgical skill assessment

An essential aspect of the study of DiPietro et al. (2019) is the hyperparameter analysis, which reveals the importance of factors, including the number of hidden units per layer and the learning rate. Additionally, it demonstrates that RNNs, especially long short-term memory (LSTM) and gated recurrent unit (GRU) models, outperform simple RNNs due to their ability to handle the vanishing gradient problem and maintain smooth predictions (17).

Skill-level classification in surgical trials

An investigation by Ismail Fawaz et al. (2018) introduces a skill-level classification task for surgical trials involving novice, intermediate, and expert levels. The study evaluates various models' performance in classifying these skill levels using micro and macro accuracy metrics. The results indicate that a particular model, a fully convoluted neural network (FCN), outperforms others with notably high accuracy, especially in needle passing and suturing tasks. However, its performance is relatively lower for the knot-tying task due to minor differences between expert and intermediate participants.

Class activation maps (CAM) for skill visualization

The article also highlights the use of CAM to visualize the most influential parts of surgical trials in skill classification. This visualization allows trainees to understand and improve their motor behaviors, potentially helping novices progress to expert levels. Additionally, the CAM technique's potential for providing feedback is demonstrated through heatmaps. These heatmaps reveal patterns that influence a subject's skill level classification. They can be used to guide novices in understanding which movements classify them as novices or experts. The article also applies the CAM technique to explain the prediction of the Objective Structured Assessment of Technical Skills (OSATS) scores. It shows how specific predictions rely on various regions of the input surgery, depending on the nature of the task (18).

Feature extraction for real-time surgical skill analysis

In a study, the authors extensively compared various feature extraction techniques for analyzing surgical skill levels in near real-time. Notably, four deep learning models, including Convolutional Neural Network (CNN), CNN-LSTM, LSTM, and Principal Component Analysis (PCA), outperformed other techniques, with CNN exhibiting the highest accuracy. Ensemble techniques, while potentially effective, might not be suitable for near real-time assessment due to computational demands (19).

The optimal time window for skill assessment varies among different skills, suggesting the importance of considering skill duration. The deep learning models were designed for efficiency in near real-time applications. Overall, the study offers valuable insights for medical engineers and researchers seeking feature extraction techniques for automated surgical skill assessment. Future work aims to explore additional techniques and datasets for a comprehensive evaluation (19).

Learning curve analysis and surgical skill progression

A study by Gao et al. demonstrates the potential of ML in assessing a trainee's surgical skill progression during repetitive trials. It shows that from these early trials, the number of trials required to achieve proficiency and the final performance level (average Fundamental of Laparoscopic Surgery [FLS] score after the 40th trial) can be accurately predicted. A single factor, referred to as "learning ability (LI)," is introduced to capture common variations within these parameters and the initial performance level (20).

The study challenges the conventional use of log-linear models for learning curve analysis. While log-linear models rely on predefined curve forms and group data, the proposed Kernel Partial Least Squares (KPLS) regression model offers a data-driven approach without assumptions about curve shapes, making it more suitable for capturing the complex process of surgical skill learning (20).

Furthermore, the study highlights the potential applications of ML in surgical training programs. The early prediction of learning curve variables in training can enable personalized training, more focused feedback, and adaptive learning strategies. By clustering trainees based on learning curve characteristics, the study identifies distinct groups with unique performance profiles, indicating that individuals with different initial skill levels require different practice to reach proficiency (20).

[DOI: 10.61186/jsurgtrauma.13.2.46]

[DOI: 10.61186/jsurgtrauma.13.2.46]

Deep learning for Real-time bacteremia detection

Using deep learning, Park et al. (2020) introduced a model for real-time bacteremia detection and prediction in surgical inpatients. Their model achieved an AUC of 0.978 for bacteremia detection every 8 h, outperforming existing criteria. In the prediction of bacteremia 24 h in advance, the model achieved an AUC of 0.929, albeit slightly lower compared to that of the detection model (21).

The occlusion analysis revealed that vital signs played a crucial role in bacteremia detection. The model using longer time steps (10 days vs. 1 day) provided more accurate results, indicating that patterns of time-variant variables, such as vital signs, contribute significantly to the characterization (21).

This study is valuable for continuous bacteremia monitoring, assisting clinicians in assessing the risk of uncontrolled infection. Unlike other sepsis prediction studies relying on clinical assessments, this model uses direct blood culture results, simplifying interpretation. The resulting predicted probabilities offer clear guidance for identifying infection sources, initiating antibiotics, and monitoring responses (21).

Surgical skill development using kinematic data

Lefor et al. (2021) presented a learning curve analysis of the Johns Hopkins–Intuitive Gesture and Skill Assessment Working Set (JIGSAWS) dataset, examining global rating scores (GRS) and kinematic parameters, such as time, path length (PL), and movements. Their analysis compared performance between the first and fifth trials for the exercises, offering insights into surgical skill development (22).

In a work conducted by Lefor et al., GRS score analysis did not reveal a learning curve effect for any exercise, unlike a prior study involving a more complex task (23). However, the analysis indicated a learning curve effect for time in the suturing exercise and PL in the needle-passing exercise. The researchers explained that this might be due to the smaller size compared to other studies and the nature of the exercises.

In this regard, kinematic data from the da Vinci surgery system is used for skill assessment in clinical surgery; however, its association with GRS scores is not always strong. Some studies have found weak correlations between kinematic parameters and skill levels (24, 25).

AI for mitigating distractions in the operating room (OR)

A novel study addressed the issue of distraction in surgical environments, which can negatively impact the efficiency of surgical procedures and potentially lead to medical errors and patient safety issues. Despite the acknowledged importance of mitigating distractions in the OR, the current methods for assessing distraction are subjective and rely on human observation (26).

In this pilot study, the researchers developed a deep CNN algorithm that leveraged electroencephalography (EEG) data from participants performing robotic-assisted surgical tasks. The results of the study indicate that EEG data can be effectively used in a deep learning model to objectively evaluate the distraction level of surgeons, potentially contributing to improved patient safety in the OR (26).

AI in surgical simulation training

A pioneering clinical trial in the literature assessed the effectiveness of AI-driven tutoring systems compared to traditional expert-led instruction in surgical simulation training. The research presents several noteworthy insights. First, it demonstrates that AI tutoring, particularly through an AI-powered virtual reality simulation platform, proves to be a valuable tool in surgical simulation training. The AI system offers feedback, establishes performance objectives, and contributes to the enhancement of participants' performance during practice sessions and realistic scenarios. This enhancement is quantitatively measured using expertise scores (27).

The study introduces AI-based feedback for surgical training, proving to be more effective and efficient than remote expert guidance. It reduces supervision time by approximately 53 h over 13 weeks while improving performance. Surprisingly, it does not evoke negative emotions, and participants prefer a combination of AI and expert instruction. This finding highlights the potential of hybrid approaches for optimal learning (27).

ML for predicting postoperative respiratory failure (PRF)

A study on EGS patients revealed a 10% PRF rate linked to higher mortality and costs. To address this, an ML model was introduced for PRF prediction using preoperative data, outperforming traditional logistic regression. Patients with PRF often have comorbidities like congestive heart failure and renal dysfunction, guiding perioperative strategies. Existing PRF risk models faced adoption challenges. The ML models provide more straightforward and accurate predictions, are suitable for EMRs, enhance patient care, and standardize protocols (28).

ML for predicting anastomotic leakage (AL)

In a retrospective study conducted at Xinhua Hospital with a cohort involving 297 patients from the Department of General Surgery, researchers found that postoperative AL occurred in approximately 6.1% of the patients. This rate was consistent with findings from previous investigations (29).

Multiple factors contribute to postoperative AL, including malnutrition, local inflammation, and patient characteristics. Researchers have explored systemic immune nutritional markers as predictors. According to findings, C-reactive protein (CRP) on postoperative day 4 emerged as a common predictor, along with variables like neutrophil-to-lymphocyte ratio (NLR) on day 4 and minimum serum albumin level. The ML models, particularly support vector machine (SVM), integrating both inflammatory and nutritional variables, demonstrated promising results with an AUC of 0.89. This approach enhances AL prediction and improves model performance while avoiding overfitting (30).

A comparison between ML and non-ML models highlighted the superiority of ML in predicting postoperative AL. The ML models considered multiple factors, acknowledging the multifaceted nature of AL. Feature analysis revealed the significance of variables like CRP and albumin levels, as well as variational indexes, including changes in albumin levels, white blood cell counts, and CRP. This research emphasized the value of analyzing clinical variable trajectories for prognostic studies (30).

An interesting finding was that CRP measured on postoperative day 7 exhibited the best AUC, surpassing other variables. However, due to a substantial number of AL cases presenting symptoms within the first week after surgery, CRP on postoperative day 7 may be more indicative of AL onset rather than a predictive risk factor (30).

Decision trees for mortality prediction in egs patients

High mortality rates in EGS patients are considered a global concern. A study used decision trees to predict mortality in EGS patients undergoing laparotomy. Key findings include the varying significance of comorbidities, with age and specific comorbidities strongly linked to mortality. Physiological factors, such as base excess (BE) and serum urea, emerged as essential predictors for high-risk patients. Shock Index was valuable for patients with enteric breaches. This risk assessment approach allows for the application of damage control principles, potentially improving outcomes in the EGS population. These findings establish the foundation for improved risk assessment and tailored interventions to enhance the prognosis of EGS patients (31).

AI and ML-based risk prediction models in surgery

In El Moheb et al.’s study (2023), the AI model POTTER demonstrated superior performance over surgeons in predicting postoperative outcomes for EGS cases. While performance varied for different outcomes, POTTER and surgeons performed similarly for septic shock but fell short when used in combination. The study stressed the importance of assessing outcomes individually, with variations in surgeon experience levels potentially favoring the use of POTTER in clinical practice. It highlighted the limitations of human cognitive capacity in dealing with complex data compared to AI. The POTTER's capability in analyzing intricate nonlinear interactions of risk factors surpassed human judgment, which can be influenced by biases and tends to overestimate risks, especially in mortality and adverse outcomes. The study concluded that AI algorithms like POTTER should complement clinical judgment, enhance risk assessments, and improve patient counseling.

While AI excels in data analysis, it requires active human engagement to interpret and contextualize its predictions. The study emphasized that the underutilization of AI tools in surgery is not due to a lack of acceptance but rather a lack of access and trust in their capabilities. Efforts should focus on increasing understanding and integration of AI-based risk prediction tools into clinical practice to enhance healthcare efficiency and delivery. Overall, the study highlights the potential benefits of AI in improving clinical risk assessment and decision-making, emphasizing the need for further research and AI tool integration in healthcare (32).

Another investigation aimed to improve the prediction of overall survival after gastroesophageal cancer resections using ML, particularly the random survival forest (RSF) model, which outperformed the traditional Cox proportional hazards (CPH) model with a c-index of 0.736. While the study benefited from detailed patient data, it recognized the need for practicality and selected the 20 most crucial features to create extended and compact RSF models that surpassed the CPH model. The findings could pave the way for an online application similar to the surgical risk calculator by the American College of Surgeons. Clinicians could input 20 variables anonymously, allowing for personalized risk assessments and survival curves based on risk groups, potentially guiding treatment decisions and surveillance in clinical trials.

Systematic follow-up, high data quality, and

[DOI: 10.61186/jsurgtrauma.13.2.46]

advanced imputation methods set the mentioned study apart from many medical publications. However, further prospective analyses and external validation are required to confirm the clinical relevance of the RSF model and other ML models. Minor differences between survival forest models may have limited clinical significance, but the study demonstrated the algorithms' ability to automatically select relevant predictors for long-term prognosis (33).

Another study introduced the Adelaide Score, an ML-based algorithm for predicting hospital discharge within 12-24 hours for general surgery patients using clinical observations and laboratory data. The RF model demonstrated superior performance and calibration, making the Adelaide Score a reliable clinical tool for healthcare systems. This research addresses the need to assess efficiency and outcomes in general surgery, offering a data-driven approach to postoperative discharge prediction that could potentially transform healthcare management. It signifies a move towards the incorporation of AI algorithms into healthcare systems to enhance efficiency and risk management (34).

AI in surgical suture performance analysis

Mansour et al. employed CNN models to analyze and evaluate suture images, focusing on assessing suture performance. The study assessed the models using various standard metrics, such as accuracy, specificity, precision, recall, and F1 score. The results were impressive, with the models demonstrating notably high accuracy, particularly the Xception model, which achieved a 95% accuracy, 96% precision, 95% recall, and 95% F1 score, highlighting its effectiveness in accurately identifying and categorizing suture images.

Conclusions

This research stands out from previous methods in the field, as it introduced a novel approach using CNN models, achieving an impressive 96% accuracy in suture image evaluation. This approach is user-friendly and accessible through an application interface, offering significant potential benefits for medical professionals. It is noteworthy that it can advance surgical education by reducing errors resulting from insufficient practice and providing efficient digitized tools.

Furthermore, this study uniquely contributes to the existing literature by introducing a user-friendly graphical interface. This interface enhances the accessibility and usability of the approach, setting it apart from previous methods. These findings shed light on the potential of deep learning techniques to enhance the accuracy and efficiency of suture training. By leveraging CNN models, the study demonstrated the ability to achieve high accuracy in assessing suture images, suggesting that deep learning approaches could be valuable for improving future training programs. The simplicity and accessibility of the approach through an application interface make it a practical tool for medical professionals seeking to enhance their skills (35).

Conflict of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

We would like to express our gratitude to our colleagues and mentors for their invaluable assistance and guidance throughout this project.

References

1. Hopkins BS, Mazmudar A, Driscoll C, Svet M, Goergen J, Kelsten M, et al. Using artificial intelligence (AI) to predict postoperative surgical site infection: A retrospective cohort of 4046 posterior spinal fusions. Clin Neurol Neurosurg. 2020;192:105718.
2. Chen W, Lu Z, You L, Zhou L, Xu J, Chen K. Artificial intelligence–based multimodal risk assessment model for surgical site infection (AMRAMS): development and validation study. JMIR medical informatics. 2020 Jun 15;8(6):e18186.
3. Tunthanathip T, Sae-Heng S, Oearsakul T, Sakarunchai I, Kaewborisutsakul A, Taweesomboonyat C. Machine learning applications for the prediction of surgical site infection in neurological operations. Neurosurg Focus. 2019;47(2):E7.
4. Fletcher RR, Olubeko O, Sonthalia H, Kateera F, Nkurunziza T, Ashby JL, et al. Application of Machine Learning to Prediction of Surgical Site Infection. Annu Int Conf IEEE Eng Med Biol Soc. 2019:2234-7.
5. Kuo PJ, Wu SC, Chien PC, Chang SS, Rau CS, Tai HL, et al. Artificial neural network approach to predict surgical site infection after free-flap reconstruction in patients receiving surgery for head and neck cancer. Oncotarget. 2018;9(17):13768-82.
6. Bartz-Kurycki MA, Green C, Anderson KT, Alder AC, Bucher BT, Cina RA, et al. Enhanced neonatal surgical site infection prediction model utilizing statistically and clinically significant variables in combination with a machine learning algorithm. The American Journal of Surgery. 2018;216(4):764-77.
7. Farrokhi M, Taheri F, Moeini A, Farrokhi M, Sayed Alireza MZ, Farahmandsadr M, et al. Artificial Intelligence for Remote Patient Monitoring: Advancements, Applications, and Challenges. Kindle. 2024;4(1):1-261.
8. Amiri H, Ghaneiyan M, Farjami P, Mehran H, Mohammadzadeh S, Ghaderi S, et al. AI-Enhanced Virtual Clinics and Telemedicine for Cancer

Treatment. Kindle. 2024;4(1):1-207.
9. Rafaqat W, Fatima HS, Kumar A, Khan S, Khurram M. Machine learning model for assessment of risk factors and postoperative day for superficial vs deep/organ-space surgical site infections. Surgical innovation. 2023;30(4):455-62.
10. Liu L, Ni Y, Zhang N, Nick Pratap J. Mining patient-specific and contextual data with machine learning technologies to predict cancellation of children's surgery. International journal of medical informatics. 2019;129:234-41.
11. Luo L, Zhang F, Yao Y, Gong R, Fu M, Xiao J. Machine learning for identification of surgeries with high risks of cancellation. Health informatics journal. 2020;26(1):141-55.
12. El Hechi MW, Maurer LR, Levine J, Zhuo D, El Moheb M, Velmahos GC, 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. Journal of the American College of Surgeons. 2021;232(6):912-9.
13. Gale SC, Shafi S, Dombrovskiy VY, Arumugam D, Crystal JS. The public health burden of emergency general surgery in the United States: A 10-year analysis of the Nationwide Inpatient Sample —2001 to 2010. J Trauma Acute Care Surg. 2014;77(2):202-8.
14. Bertsimas D, Dunn J, Velmahos GC, Kaafarani HMA. 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. Annals of surgery. 2018;268(4):574-83.
15. Gao J, Merchant AM. A machine learning approach in predicting mortality following emergency general surgery. The American Surgeon. 2021;87(9):1379-85.
16. Khoshravesh S, Arab M, Nouri N, Nouri B, Pirastehfar Z, Ghavami B. A Comparative Cohort Study of Laparoscopy Versus Laparotomy for the Treatment of Iranian Endometrial Cancer Patients. International Journal of Cancer Management. 2024;17(1):e141894.
17. DiPietro R, Ahmidi N, Malpani A, Waldram M, Lee GI, Lee MR, et al. Segmenting and classifying activities in robot-assisted surgery with recurrent neural networks. Int J Comput Assist Radiol Surg. 2019;14(11):2005-20.
18. Ismail Fawaz H, Forestier G, Weber J, Idoumghar L, Muller PA. Accurate and interpretable evaluation of surgical skills from kinematic data using fully convolutional neural networks. International journal of computer assisted radiology and surgery. 2019;14(9):1611-7.
19. Anh NX, Nataraja RM, Chauhan S. Towards near real-time assessment of surgical skills: A comparison of feature extraction techniques. Comput Methods Programs Biomed. 2020;187:105234.
20. Gao Y, Kruger U, Intes X, Schwaitzberg S, De S. A machine learning approach to predict surgical learning curves. Surgery. 2020;167(2):321-7.
21. Park HJ, Jung DY, Ji W, Choi CM. Detection of bacteremia in surgical in-patients using recurrent neural network based on time series records: development and validation study. Journal of Medical Internet Research. 2020;22(8):e19512.
22. Lefor AK, Harada K, Dosis A, Mitsuishi M. Motion analysis of the JHU–ISI gesture and skill assessment working set II: learning curve analysis. International Journal of Computer Assisted Radiology and Surgery. 2021;16(4):589-95.
23. Hernandez JD, Bann SD, Munz Y, Moorthy K, Datta V, Martin S, et al. Qualitative and quantitative analysis of the learning curve of a simulated surgical task on the da Vinci system. Surgical Endoscopy And Other Interventional Techniques. 2004;18(3):372-8.
24. Hung AJ, Oh PJ, Chen J, Ghodoussipour S, Lane C, Jarc A, et al. Experts vs super-experts: differences in automated performance metrics and clinical outcomes for robot-assisted radical prostatectomy. BJU international. 2019;123(5):861-8.
25. Hung AJ, Chen J, Gill IS. Automated performance metrics and machine learning algorithms to measure surgeon performance and anticipate clinical outcomes in robotic surgery. JAMA surgery. 2018;153(8):770-1.
26. Shafiei SB, Iqbal U, Hussein AA, Guru KA. Utilizing deep neural networks and electroencephalogram for objective evaluation of surgeon’s distraction during robot-assisted surgery. Brain research. 2021;1769:147607.
27. Fazlollahi AM, Bakhaidar M, Alsayegh A, Yilmaz R, Winkler-Schwartz A, Mirchi N, Langleben I, Ledwos N, Sabbagh AJ, Bajunaid K, Harley JM. Effect of artificial intelligence tutoring vs expert instruction on learning simulated surgical skills among medical students: a randomized clinical trial. JAMA network open. 2022;5(2):e2149008-.
28. Hadaya J, Verma A, Sanaiha Y, Ramezani R, Qadir N, Benharash P. Machine learning-based modeling of acute respiratory failure following emergency general surgery operations. PLoS One. 2022;17(4):e0267733.
29. Roh CK, Choi S, Seo WJ, Cho M, Kim HI, Lee SK, et al. Incidence and treatment outcomes of leakage after gastrectomy for gastric cancer: Experience of 14,075 patients from a large volume centre. Eur J Surg Oncol. 2021;47(9):2304-12.
30. Liu X, Lei S, Wei Q, Wang Y, Liang H, Chen L. Machine Learning-based Correlation Study between Perioperative Immunonutritional Index and Postoperative Anastomotic Leakage in Patients with Gastric Cancer. Int J Med Sci. 2022;19(7):1173-83.
31. Smith MTD, Bruce JL, Clarke DL. Using Machine Learning to Establish Predictors of Mortality in Patients Undergoing Laparotomy for Emergency General Surgical Conditions. World J Surg. 2022;46(2):339-46.
32. El Moheb M, Gebran A, Maurer LR, Naar L, El Hechi M, Breen K, et al. Artificial intelligence versus surgeon gestalt in predicting risk of emergency general surgery. J Trauma Acute Care Surg. 2023;95(4):565-72.
33. Jung JO, Crnovrsanin N, Wirsik NM, Nienhüser H, Peters L, Popp F, et al. Machine learning for optimized individual survival prediction in resectable upper gastrointestinal cancer. J Cancer Res Clin Oncol. 2023;149(5):1691-702.

[DOI: 10.61186/jsurgtrauma.13.2.46]

34. Kovoor JG, Bacchi S, Gupta AK, Stretton B, Malycha J, Reddi BA, et al. The Adelaide Score: An artificial intelligence measure of readiness for discharge after general surgery. ANZ J Surg. 2023;93(9):2119-24.

35. Mahmud S, Mansour M, Cumak EN, Kutlu M. Deep learning based suture training system. Surgery open science. 2023;15:1-11.

[DOI: 10.61186/jsurgtrauma.13.2.46]