The impact of artificial intelligence on surgical outcomes

Habeeb salman Alhamood, Ali Yousef Al Qadhi, Yahya Hussain AL haddaf, Abdullatif Fahad Alsaliemy, Hassan ahmed almahdi, Mustafa Mohammad Alqutef, ALI HUSSAIN ALRAJAA, Mohammed Khaled Alnajim, Abbad Ali AlNasser, Mohammed Abdul moheen alaqnam

Abstract

Artificial intelligence (AI) has the capacity to transform surgical results by offering immediate insights, enhancing diagnostic precision, and refining therapeutic approaches. **Aim of the Study:** This study attempts to evaluate the present status of artificial intelligence in the field of surgery, with a specific emphasis on its uses, constraints, and forthcoming consequences. **Methods:** A thorough literature search was performed in PubMed and Cochrane Library databases to find publications on the use of artificial intelligence in surgical operations. Articles were reviewed according to certain criteria, and information about AI applications, predictive modeling, risk assessment, and constraints was collected. **Results:** AI algorithms showed potential in forecasting postoperative complications, surgical results, and patient risk evaluation. Machine learning methods surpassed traditional statistical models in terms of accuracy and precision. Limitations such as data quality, model interpretability, and ethical issues were emphasized. **Conclusion:** Artificial intelligence has the potential to revolutionize surgical procedures by improving decision-making, customizing treatment strategies, and promoting patient safety. Additional research is required to overcome data constraints, enhance AI models, and provide ethical guidelines for the proper integration of AI in surgical procedures. It is essential for surgeons, data scientists, and ethicists to collaborate in order to fully use AI and transform surgical care. **Keywords:** Artificial intelligence, Machine learning, Surgical outcomes, Predictive modeling, Risk assessment.
**Introduction**

Artificial intelligence (AI) is the field that focuses on algorithms enabling computers to engage in reasoning and cognitive tasks including problem solving, image and word identification, as well as decision-making. Artificial intelligence (AI) is an applied science that employs concepts and computational techniques to enable machines to carry out activities often done by human intellect (McCarthy and Hayes, 1981; Korndorffer Jr et al., 2020). Significant progress in processing speeds, online storage, and the accessibility of documented large datasets has substantially eased the widespread use of AI in several sectors. AI is expected to greatly benefit medical procedures, patient satisfaction, and precise image analysis in medicine (Topol, 2019). AI has significant drawbacks such as prejudice, insufficient clarity in result parameters, and varying levels of precision throughout data and system kinds. AI has become very used in the field of surgery for analyzing surgical footage. Business and educational institutions are investigating methods to effectively gather, preserve, and analyze surgical recordings to provide decision assistance for enhancing surgical performance and patient results (Hashimoto et al., 2018; Samareh et al., 2019; Mascagni et al., 2020). During a recent Surgical Data Mining meeting, data scientists, engineers, and surgeons identified the main obstacle to widespread use of AI in surgery as the lack of availability of big surgical datasets with more than 1000 items (Maier-Hein et al., 2017). Lacking data accessibility hinders the development and validation of data analytical methods, data handling norms, and user input. Surgeons should possess a fundamental understanding of AI to comprehend its potential effects on medicine and to explore possible interactions with this form of technology. The review introduces AI by focusing on its impact in surgery. It discusses their limits and potential future impact on surgeons.

**Methodology:**

A search was conducted in the PubMed and Cochrane Library databases from their beginning until 2024. Articles were evaluated by examining their abstracts based on the specified criteria: (1) released in English, (2) released in a peer-reviewed publication, and (3) originally written or a review piece. Articles were chosen based on a thorough evaluation of abstracts and complete texts to meet the criteria of subject relevance, methodologic strength, and innovative or meritorious addition to existing literature. Articles referenced by the original search results were examined using the same standards.

**Artificial Intelligence Versus Conventional Risk Tools**

Knowing precisely a postoperative complication's risk is essential for perioperative decision-making, initial guidance, postoperative boundary assessment, and patient choice optimization prior to surgery (Merath et al., 2020). Several risk classifications and forecasting algorithms have been created to assess patients' postoperative complications and death risk, such as the Physiologic and Operative Severity Score for the Enumeration of Mortality and Morbidity (POSSUM), the American Society of Anesthesiologists (ASA) score, and American College of Surgeons Surgical Risk Calculator (ACS-SRC) (Bilimoria, et al., 2013; Hassan et al., 2023).

Conventional risk evaluation instruments like the POSSUM, ACS-SRC, ASA rating and have constraints that can reduce their forecasting accuracy, despite being important for evaluating the risk of challenges after surgery. Conventional risk evaluation techniques employ Cox exponential dangers regression assessment or logistic regression approaches, which may lead to overestimation or underestimation of risks due to their reliance on factors with statistically substantial impacts on results. These methods could overlook nuanced but significant elements that might impact results.

Conventional models frequently presume a linear relationship among factors and results and can neglect intricate relationships, especially at the extreme ends of parameter categories (Chen and Asch, 2017). Regression modeling is limited by
The impact of artificial intelligence on surgical outcomes

overestimation and plurality, which prevent the evaluation of several variables. Existing models for forecasting are generally restricted to a small number of parameters and may not include other significant factors that influence the result (Cruz and Wishart, 2006).

AI approaches often outperform classic statistical models such as logistic regression in terms of predicting accuracy (Hassan et al., 2022). Machine learning techniques enable the assessment of a greater quantity of medical information compared to conventional modeling methods, perhaps revealing subtle indicators or connections among factors that might enhance prediction precision (Hashimoto et al., 2018). Utilizing nonlinear approaches including different types of data like assessments, therapies, and laboratory measurements, machine learning has surpassed logistic regression in forecasting postoperative results (Soguero-Ruiz et al., 2015; Fei et al., 2017; Lee et al., 2018).

Recent study proved that machine learning outperformed multivariable logistic analysis in predicting problems and identifying variables in abdominal wall repair and mastectomy restoration (Hassan et al., 2022). Progressive learning allows machine learning to enhance its performance continually by incorporating fresh data, unlike traditional statistical methods. In contrast to conventional hazard calculators, which are fixed, machine learning systems are dynamic and evolve continually (Gepperth and Hammer, 2016).

Surgical Applications

Hepatobiliary And Colorectal Procedures

Merath et al. (2020) created a predictive model for anticipating problems in patients who have hepatic, pancreatic, and colorectal surgeries utilizing the National Surgical Quality Improvement Program (NSQIP) dataset. The algorithm was developed on a dataset of 15,657 individuals and produced an area under the curve (AUC) that varied from 0.76 for predicting operative area illnesses to 0.98 for predicting strokes. The study showed that the machine learning algorithms accomplished better than the ASA as well as ACS-SRC. Further studies have using deep learning algorithms to forecast difficulties for individuals with regionally advanced or recurring colorectal cancer undergoing pelvic exenteration (PelvEx Collaborative et al., 2020). An artificial neural network (ANN) model was created in a research with 1,147 patients, achieving area under the curve (AUC) values between .61 and .79. The scientists determined that their deep learning algorithm outperformed logistic regression models in forecasting a result based on a complicated set of individuals- and procedure-related characteristics.

Cardiothoracic Operation

Lapp et al. (2018) used machine learning algorithms to forecast serious postoperative problems in individuals who had undergone coronary artery bypass transplant along with valve operation. With information from 3,700 individuals, the random forest approach exhibited better results compared to other approaches, having an AUC of .72. Salati et al. (2021) later created a machine learning algorithm to forecast cardiopulmonary problems in patients after lung resection. The severe gradient enhancing approach, tested on 1,360 individuals, attained an AUC of .75 and a precision of seventy percent, surpassing previous algorithms. The researchers determined that machine learning algorithms offer specific forecasts by examining accessible patient data. These algorithms aid in operating decision-making by recommending tailored postoperative services, choosing preoperative treatments for patients with elevated risk, and assessing the quality of care.

Mastectomy

Hassan et al. (2022) have recently created machine learning algorithms to forecast skin flap degeneration among individuals after mastectomy for breast tumors. They assessed the prediction effectiveness of nine machine learning methods using information gathered from 694 individuals. The techniques, based on easily accessible perioperative information, correctly estimated the likelihood of mastectomy epidermal flap degeneration for specific patients through an AUC of 0.70 and 89 percent precision. They discovered risk variables
linked with this issue by PFI as well as ALE evaluation.

In a recent research, Hassan et al. (2022) used machine learning algorithms to forecast the exploratory lymph node condition in elderly individuals diagnosed with invasive breast cancer. An analysis was conducted on data from 1,706 consecutive patients to create a support vector machine model that can predict node positive before surgery based on patients' demographics, cancer phase, genetic description, and scan information. The algorithm accurately determined the sentinel lymph node condition in those individuals with an overall precision of eighty-four percent (95% CI: 80.88%) and had strong predictive ability with an AUC of .70 (95% CI: .62-.77). The examination of the model identified characteristics linked to lymph node positive, including illness phase, young age, and familial background of breast tumors, border position, and progesterone and estrogen receptor positivity. The simulation may assist patients in comprehending their likelihood of developing node-positive illness, perhaps impacting decisions about surgical procedures and further medication.

**Plastic and Reconstructive Operation**

Machine learning algorithms have been created to forecast difficulties during implant-based breast augmentation (Hassan et al., 2022). Researchers used perioperative information collected on 481 individuals to develop machine learning algorithms for predicting periprosthetic illness and the requirement for device transplantation. The machine learning algorithms showed strong efficiency for forecasting illness (AUC, .73; precision, eighty-three percent and the requirement for equipment removal (AUC, .78; accuracy, 84%). ML models surpassed standard multivariable logistic regression techniques in recognizing risk variables including gadget implantation plane, kind of a cellular dermal matrix employed, and supplemental treatment.

When analyzing the information, multivariable logistic regression revealed two indicators of disease, whereas machine learning (ML) found nine. Such algorithms assist surgeons in making well-informed judgments while offering a reliable and precise measure to utilize when advising patients on potential restorative options and outcomes. The simulations may help a patient throughout the written permission procedure by forecasting the potential dangers and advantages of a certain treatment and recognizing adjustable factors that can be dealt with before restoration to enhance the patient's eligibility for the surgery.

**Neurosurgery**

Van Niftrik et al. (2019) created a severe gradient enhancement algorithm to forecast difficulties in patients having brain tumor operations. The approach, based on information gathered from 668 patients, achieved a precision of seventy percent and an AUC of .74, outperforming a traditional statistical approach in prognostic capability. Farrokhi et al. (2020) used machine learning techniques to forecast difficulties during cerebral stimulation operation. The trained models showed excellent discriminatory efficacy in forecasting several outcomes: any incident (AUC, .86), problem within a year (AUC, .91), requirement for a second surgical treatment (AUC, .88), and illness (AUC, .97). The researchers determined that machine learning may enhance risk evaluation, preoperative informed approval, and therapeutic scheduling for patients undergoing neurosurgery.

**Orthopedic Operation**

Kim and colleagues (2018) used the NSQIP database to create a machine learning model for forecasting problems after posterior lumbar spine surgery. An artificial neural network (ANN) model was created utilizing data from 22,629 individuals to forecast occurrences of cardiac problems, complications of wounds, venous thromboembolism (VT), and mortality. The artificial neural network (ANN) algorithm attained an AUC of 0.71, surpassing the benchmark ASA results.

Devana and colleagues (2021) employed information gathered from 156,750 patients to create a set of machine learning models aimed at forecasting difficulties following complete knee arthroplasty. The algorithms showed strong prejudiced ability with an AUC of .68. The authors suggested that the team could be beneficial for
The impact of artificial intelligence on surgical outcomes

Preoperative therapy, collaborative decision-making, informed approval, risk modification reimbursement applications, and handling postoperative standards for patients as well as surgeons.

General Operation

Hassan et al. (2022) is currently developing machine learning algorithms to forecast the results of abdominal wall restoration. A set of 9 trained machine learning models was constructed utilizing information from 725 individuals for predicting hernia repetition, surgical area incidences, as well as hospitalizations during 30 days. The group employed numerous machine learning algorithms to achieve improved prediction accuracy by the majority principle. Machine learning algorithms shown strong predictive capacity in predicting problems throughout an extended follow-up duration of around 3 years. This included recurrent hernias with a precision of eighty-five percent and an AUC of 0.71, along with short-term consequences like surgical area incidences with a reliability of 72 percent and an AUC of 0.75, and a thirty-day readmission with reliability of eighty-four percent and an AUC of 0.73.

Model assessment helped uncover elements linked to negative consequences that were not evident in conventional statistical methods like logistic regression. The parameters considered were surgical procedures, previous abdominal operations, and the extent of wound infection. From comparable records, multivariable logistic analysis found five variables of surgical area incidences, while machine learning discovered 12 variables. ML algorithms can enhance surgical preparation, preoperative improvement, and collaborative decision-making.

Risk Estimators

Machine learning can provide current risk evaluations, allowing healthcare professionals to evaluate patients' risk and determine whether their condition is suitable for operation at that moment. Two verified machine learning-based risk assessments have been created to forecast significant surgical consequences. MySurgeryRisk was created and verified employing information gathered from 51,457 individuals who underwent significant hospitalized operation to forecast eight significant postoperative problems (acute kidney damage, septicemia, venous thromboembolism, and enrollment to critical care following 48 h, ventilatory assistance following forty-eight h, injury, and neurological as well as cardiac problems) along with mortality throughout two years complying with operations (Bihorac et al., 2019).

The model produced an Area Under the Curve (AUC) between 0.77 and 0.94. Brennan et al. (2019) later evaluated the usefulness and reliability of the MySurgeryRisk calculator in comparison to clinicians' medical judgment. MySurgeryRisk was shown to surpass clinicians' initial risk estimations for nearly every postoperative problems. The doctors' risk evaluations showed considerable improvement after their interaction with the ML system.

Bertsimas et al. (2018) created the POTTER risk estimator for Emergency Surgery using Predicted opTimal Trees. Thirty nine The framework was created employing a decision tree machine learning method with information from 382,960 individuals in the NSQIP data. POTTER demonstrated superior predictive accuracy for illness (AUC, .84) as well as death (AUC, .92) compared to the ASA calculator, Urgent Surgery rating, and NSQIP risk estimator.

Constraints Of AI Models

The efficacy of AI prediction relies on the precision and thoroughness of the incoming data. Obstacles in medical information collecting may influence the trends that AI identifies and the recommendations it generates (Moore et al., 2023). The majority of AI algorithms were created employing registry-based information like NSQIP, which has constraints such as short follow-up time, restricting the predicted precision to short-term results. Moreover, AI may uncover complex trends and hidden risk elements that are not apparent in conventional risk estimation. However, neglecting to follow guidelines in building models might lead to unfavorable results and decreased model effectiveness. Most research employing AI in surgeries have mostly been empirical.

Although AI has limits, the utilization and examination of ML techniques are crucial
for comprehending the different elements that forecast real-world surgical results. Substantial hurdles and hazards persist in the use of AI in surgery, including dependability, openness, responsibility, security, and confidentiality of information, effectiveness, systemic inequity, workforce replacement, and ethical considerations. A comprehensive authority structure is necessary for the creation, installation, and use of AI in surgeries.

**Conclusion**

This review’s results provide data that supports the incorporation of artificial intelligence in surgical procedures. AI systems have shown exceptional ability in forecasting surgical results, recognizing risk factors, and enhancing therapeutic approaches. Surgeons may enhance their decision-making, minimize problems, and enhance patient care by using AI. Although there have been encouraging improvements, certain difficulties must be overcome to assure the general use of AI in surgery. Ensuring data quality and uniformity is crucial for creating precise and dependable AI models.

It is crucial for healthcare institutions, data scientists, and regulatory agencies to collaborate in order to create strong data gathering and sharing systems. Moreover, the interpretability and explainability of AI models are essential for establishing confidence with surgeons and patients. Creating AI algorithms with transparent explanations for their predictions will make it easier to incorporate them into therapeutic settings. Ethical issues such as patient privacy, data security, and possible biases need to be thoroughly handled to enable the appropriate use of AI in surgery.

Future research in the field of AI in surgery should prioritize the following areas:

- Enhance data quality and standardization by implementing measures to improve data collecting, assure data integrity, and develop standardized data formats to increase the reliability and generalizability of AI models.
- Development and Validation of the Model: Perform thorough research to create and confirm AI models with extensive, high-caliber datasets. Examine several AI methods like deep learning and natural language processing to enhance predicted accuracy and tackle intricate surgical situations.
- Explore techniques to enhance the interpretability and comprehensibility of AI models for surgeons. Create visualization tools and explanations for AI forecasts to aid decision-making and establish confidence with healthcare practitioners.
- Research the ethical implications of artificial intelligence in surgery, focusing on patient permission, data privacy, and possible biases. Create standards and frameworks that guarantee the ethical and acceptable utilization of AI in surgical procedures.
- Investigate methods for incorporating artificial intelligence into clinical procedures and surgical decision-making protocols. Examine the effects of artificial intelligence on surgical results, cost efficiency, and patient contentment by conducting real-world implementation research. Exploring these study areas will enable us to maximize the capabilities of AI in surgery, providing surgeons with advanced tools to provide accurate, customized, and secure patient care.
The impact of artificial intelligence on surgical outcomes

References


