Intraoperative Decision Support and Complication Risk Stratification: Machine Learning Models for Surgical Outcome Prediction and Enhancement

Dr. Evelyn Figueroa, Professor of Industrial Engineering, University of Chile

1. Introduction to Machine Learning in Surgery

Technology spillovers from other fields have transformed surgery. In minimal access procedures, current robotic instruments now offer performance in dexterity, stability, and visualization far exceeding the surgeon's capabilities. However, artificial intelligence is a complementary and not a direct substitute technology that has transformative potential in surgery. At its core, machine learning is a subset of artificial intelligence, and both terms have taken on a number of definitions through their evolutions. For the purposes of this review, machine learning refers to a set of mathematical techniques that have shown promise for the discovery of new and subtle relationships among a suite of data inputs. Consequently, these techniques excel as a supplement to human cognitive processes by enhancing both sensory perception and decision streams.

The practical surgical need for intelligently informed and integrated data-driven operating rooms continues to grow. Intraoperative decision-making, greatly influenced by imperfect information, often lacks guiding structure. Integrating machine learning into the surgical field can therefore hold the promise of enhancing decision-making processes under conditions of uncertainty. As we progress to digital transformation, manual decision-making processes across all three major surgical domains need to be urgently replaced with data-driven automated processes. Organizations are moving quickly to devise strategies that embrace and exploit AI, and are actively exploring and researching AI in numerous surgical subspecialties. Many complex decisions need to be resolved to enable the integration and acceptance of AI into surgical practice. Many surgical teams worldwide are conducting research with AI. The majority of this research is based on machine learning approaches, covering over half the articles published in the last decade.

1.1. Overview of Machine Learning and Artificial Intelligence

Machine learning and artificial intelligence (AI) are terms often used interchangeably, but they do have specific meanings. AI encompasses a wide range of computing methods that are "smart." That is, they exhibit at least some of the characteristics we associate with intelligence in human behavior, such as understanding language, learning, problem solving, reasoning, and the like. Within the AI realm, machine learning represents one way that machines can function and interact more intelligently with people. As the name implies, machine learning is based on the idea that systems can learn from data, identify patterns, and make decisions with minimal human intervention. The more data that machines can access, the better they can learn to perform complex tasks and improve cognitive functionality.

The machine learning model is trained using the data and imports essential features from a high-dimensional data set to classify observations. Like humans, the more experience a machine or AI model has, the more accurately it can perform a task. Machine learning models are capable of recognizing patterns and trends in both structured and unstructured data to improve decision-making without specific programming. The patient exam example illustrates the potential power and generalization capabilities of such an approach. In the past, a form of machine learning was applied to clinical decision support modeling and demonstrated a statistically significant increase in patient outcomes, such as predicting the probability of surgical site infections in real-time cases with national data. In the future, as digital ecosystems and data landscapes evolve and continue to drive technology advancements, machine learning-based models will become more powerful and accurate. The first half of this section will discuss supervised and unsupervised learning systems and their application to preoperative planning in a clinical setting.

1.2. Applications of Machine Learning in Surgery

Machine learning methods have played an increasingly important role in many aspects of surgical practice. From preoperative diagnosis, personalized treatment planning, and prediction of postoperative risks, datasets and data collection for machine learning-based algorithms have enriched the information we can learn from routine clinical practice to aim for better decision-making. The most exciting recent development has been combining results from large-scale neuroimaging datasets with algorithmically derived metrics, which have been demonstrated to be useful in predicting major complications and rehabilitation. Intraoperatively, machine learning can also offer assistance with multiple aspects of surgical decision-making. These include preoperative personalization based on the risk of bleeding and excessive surgical trauma, based on the form of heat mapping of predicted activity and damage based on neuroanatomy, to establishing the viability of resection or excision in the context of baseline investigations and patient anatomy. Several increasingly cited series report on the utility of machine learning techniques to personalize the extent and choice of interventions during image-guided neurosurgery, using a range of datasets, imaging methods, and outcomes derived from the Perioperative Enhancement Approach for Patients Undergoing Neurosurgery series. In addition to supportive applications, collaborative advances are also possible when developing machine learning-based navigation tools that send and receive workflow-relevant data from the overall surgical process. Not only as reported but also possible is to learn surgical models, i.e., learning to map preoperative assessments to a map of the expected intraoperative surgical anatomy, to anticipate scale resection probabilities, surgical complication likelihoods, and dose-adapted radiotherapy targeting more likely recurrence areas. The opportunities reside with utilizing employers who can also learn this network to consider relationships between brain regions with surgical exploration subtler than that convention and advanced presented here. Some researchers have already used this strategy to train neural networks toward boosting cognitive functionality while also sustaining plasticity. The iterative nature of these analyses is, therefore, fruitful. By formal data collection protocols and AI operationalization for robotic neuronal procedures, we expect these methods to be established in the neurosurgical field.

2. Preoperative Planning

The preoperative phase, as the first part of the surgical care process, presents several potential roles for machine learning algorithms to be applied for effective surgical workflow improvements. Traditionally, preoperative assessment is often performed on a case-specific, reactionary, and episodic basis that leads to fragmentation of the caregivers. This takes valuable operative time, and patient flow is coordinated around the capability of the preoperative diagnostic service. This chronological method of planning has been shown to be inefficient, with one in three patients needing further investigations up to the day of their surgery. Existing practices for preoperative patient evaluations rely only on clinical, diagnostic, and human-report-based methods. The traditional, fragmented processes have additional inefficiencies such as extensive patient waiting times and, in turn, excessive nonoperational time in theaters. This severely limits operative capacity. The preoperative surgical workup helps to minimize postoperative complications and has the chance to maximize the ability to function, reduce the time in recovery, and lower the readmission rates.

Current methods used by surgical teams need to apply definition systems that provide an estimation of the likelihood of various outcomes ranging from patients' attitudes to the outcomes of interest. However, a sequential approach is time-consuming, and most scores currently applied are non-informative or limited by experienced medical teams on human decision-making values. On the other hand, machine learning has been shown, in many cases, to have the capability to process large amounts of data and develop prediction models that can support clinical decision-making. The use of predictive analytics can provide an opportunity to design tools for the surgical team, non-experts in analytics or data management, that could assist in clinical evaluations, discrepancy management, and patient stratification. To this aim, several machine learning models have been proposed and used as assisting tools for effective management. Non-computational models use unsupervised learning. A mixed method based on fuzzy logic and rule-based classifiers of intrinsic dynamic chondrocyte properties was used to stratify it into the progression of osteoarthritis and response to treatment. Integrated wearables, including eye-tracking and machine learning techniques aided in the risk stratification of surgeons for fatigue, distress, and burnout mitigation. Decision support systems, wearable technologies, and semi-supervised learning were integrated to capture valuable patient data that relate to medical care for decision-making, remote patient care, and public health applications. Multiple other examples are available; all demonstrate and confirm the expanding application of machine learning in multifaceted environments.

2.1. Challenges in Traditional Preoperative Planning

Traditional methods of preoperative planning have limitations and face challenges that significantly impact overall surgical outcomes. Such traditional ways of patient assessment before surgery are not standardized and may create variability in results that require subjective clinical judgment. As a result, different intraoperative risk factors are usually not covered by preoperative assessment, which can lead to negative outcomes after surgery. Additionally, clinical judgment and the decision-making process are the main reasons for increased surgical risks, resulting in negative outcomes. Therefore, it is crucial to transition towards the development of an intelligent and automated system.

Nonetheless, using machine learning techniques to develop computer-aided systems can greatly improve and standardize preoperative assessments. In fact, such a system will allow physicians to evaluate risk factors based on real-world data in the shortest time, facilitating the required action. These critical challenges should be promptly addressed in the current preoperative planning and guiding system for better and more comprehensive planning procedures and improvements in the decision-making process. Currently, the traditional methods used in preoperative planning provide limited information in a heterogeneous patient set, leading to invasive, time-consuming, and costly examination procedures. Therefore, the goal in surgical departments is to move towards a systematic and holistic solution.

As such, every patient's safety and medical care require an in-depth and broad analysis that includes all anatomic and functional impairments in an individual presenting for an operative procedure. Moreover, delays in diagnosis and planning may result in valuable on-time and unscheduled admission cancellations due to concomitant therapy, which further increases costs. Significant and serious complications arising before surgery may lead to postoperative complications, morbidity, and mortality.

2.2. Machine Learning Techniques for Preoperative Planning

Data analysis. Preoperative planning, containing risk analysis, is a key process before surgery. In recent years, the application of machine learning for statistical data analysis services has yielded significant results. A machine learning model was built using Electronic Health Record data of 100,299 patients to predict perioperative cardiac arrest. This also shows the important role of modern statistical methods in the preoperative period in prediction modeling based on big data. Moreover, machine learning models are applicable in analyzing massive EHR datasets, holding the potential to be employed in preoperative planning. A Convolutional Neural Networks model has been adopted to train and predict outcomes from more than 150,000 dental images. A clinical decision support system powered by machine learning has been developed to assist orthopedic surgeons in preoperative planning. This platform could predict lumbar pathology and suggest comprehensive treatment options for spine surgery candidates. A retrospective cohort study involving 145 spine surgery candidates showed the effectiveness of this model with up to 84% accuracy. The same support system was also used in hip joints in patients with femoroacetabular impingement. It was possible to predict with 90.5% accuracy whether a patient would undergo hip preservation surgery to slow the onset of hip osteoarthritis in the future. However, integrating patient data with imaging information will provide more accurate reconstruction planning results. An anchor tunable learning approach was utilized to preoperatively simulate patient data with ACL ruptures, identifying which patients with failed conservative treatment would have good postoperative joint stability after single-bundle ACLR surgery. This technique achieved an average area under the receiver operating characteristic curve of 0.768. Although the relatively small sample size of only 60 cases may show the superiority of this machine learning model in preoperative planning, a more comprehensive predictive model trained with big data is still required. In the field of planning tumor resection, some studies have leveraged machine learning models to predict possible post-surgical complications, such as brain tumor-related epilepsy, and their results improved progressively with increasing sample sizes. In conclusion, these machine learning models can predict some outcomes that might be associated with the success of the surgery and provide benefits for the patients. More sophisticated applications in new areas or more common routine preoperative examinations are still being explored. Developments in the use of machine learning models for preoperative planning are also moving in the direction of personalized medical solutions that leverage the technology of training big data. Most preoperative planning applications use patients' preoperative characterization added by mathematical indicators or combined with image features to enhance their prediction.

3. Intraoperative Guidance

In the operating room, machine learning models can be used to enhance surgical outcomes. By integrating visual analytics and robotics, we can increase the situational awareness of the surgical team and provide solutions that are likely to lead to reduced surgical procedure times. Tracking vital structures can make interpretations based on the current state of the surgery and support many different intraoperative tasks such as estimating sign-out times and even predicting patient recovery times. Deep learning models can be trained to track anatomical landmarks in both endoscopic and open surgeries and provide real-time metrics for procedure optimization. Providing these models quantified metrics, designed with surgery in mind, builds trust in these new methods, which are essential for model acceptance in the field. Failure of a model can be critical, especially when defects lead to a drop in performance. Nevertheless, we believe these approaches lead to enhanced perioperative safety by providing thorough and detailed situational recommendations at the right time for the correct decision. Intraoperative guidance presents an opportunity for further development of machine learning techniques, which can lead to significant improvements in surgical outcomes. Our introduction to machine learning preoperative planning and postoperative monitoring tasks provides methods that demonstrate improved patient survival and recovery, and we propose that the intraoperative guidance advances could lead to further patient benefits and decreased length of stay. At the same time, existing workflow problems or previously unknown patient conditions could be discovered while in the operating room, and surgical strategies could be amended. Multiple machine learning models, designed for real-time and near real-time surgical situation inference and definition, will lead to significant benefits for patient recovery and hospital resource estimation. Models that can personalize the intraoperative time depending on the patient's unique surgical history may provide an optimized recovery pathway and potentially lead to decreased recovery times, decreased infection risk, and shorter length of stay.

3.1. Importance of Real-time Guidance

All surgical procedures come with the potential for complications. They may take a huge toll on the patients in the form of mortality, morbidity, extended periods of hospital stay, and healthcare costs. Mortality and the number of surgical procedures can be reduced to some extent if and only if some of the procedures currently performed by human surgeons are replaced by surgical robots. There are some semi-autonomous surgical robots available in the market, and their application range is expanding rapidly. However, the surgical actions of these machines are controlled by the human surgeon. These robots do not have the capability to act on their own or make appropriate decisions at the time of need. If that need is to be addressed, surgical robotic systems must be designed in such a way that they would deliver real-time decision guidance.

The real challenge in obtaining a successful robotic system design that is capable of delivering real-time guidance is that it would require the integration of real-time processing, computer vision, and machine learning algorithms within its own architecture. During the procedure, a wealth of information can be captured in the form of video data and multi-sensor data. However, the high dimensionality and the inherently less structured nature of these clips make extracting the necessary information a harder task. At the same time, surgical data has an inherent class imbalance, requires real-time processing constraints, and has limited access to ground truth labels. The fact that obtaining ground truth label data for surgery is prohibitively expensive and extremely time-consuming should not be overlooked. Digging deep, there are a number of inherent challenges associated with real-time processing, such as the detection and segmentation of salient structures and objects, detecting the tissues and boundaries in high resolution and high-quality images, localizing the hidden object, and tracking the moving stage.

3.2. Types of Machine Learning Models for Intraoperative Guidance

This classification is related to the main goal that the machine learning model must solve: intraoperative support and the type of actions. Action-aware approaches are a category of machine learning methodologies that need to take actions. For the action, there are distinct manners in which the model can implement it, including encapsulated actions, scaling actions, and control actions by applying momentum to improve precision, capturing spatial-temporal properties for trajectory and energy at the part level, collaborative and hierarchical control forecast. The encapsulation of the capability to take a sequence of steps implementing a planned strategy, without requiring further mechanics, is called embodiment. For the intraoperative model, it is crucial to avoid repeated requests for confirmation, which can delay surgery, or incorrect responses, which can harm the patient. In comparison to preoperative models, in the intraoperative case, there is a restriction regarding incorporating outcomes versus taking actions based on plans. Some of the actions are directly related to treating the object, which can be actuated by the robot. Another action is confirming the presence and healthcare status of an entity during surgery, referred to as entity confirmation. Also, during surgical procedures, it is necessary to incorporate the ability to request scans, radiological images, and pictures of the sensitive regions in the surroundings. Types of machine learning methodologies focused on obtaining tools for support during the surgery itself are called action-aware. These action-aware models provide additional information supporting the interventions, resulting in increased decision accuracy. Several categories of machine learning algorithms can help inform these model types. Models for intraoperative surgical guidance fall into different categories with distinct types of data handling and enhanced decisions. Should the intraoperative support be minimally invasive, the model's performance must be real-time, with limited computational cost. A time window must produce actions. Good actions must have limited penalties. The intraoperative support model's performance depends on both future actions and the patient's feedback, such as increased accuracy and reduced costs. For proactivity, model performance is a priority. There is a need to handle variations. Varied data can make a fallback plan for uncertain situations indispensable. Increased costs of bad actions introduce significant risks, but the cost of algorithm complexity is limited and highly depends on the case. Urgency is a priority; a quick explanation is nonessential. The data is reasonably stable and manageable, and the patient is not a part of the loop.

4. Postoperative Monitoring

Monitoring the patient after surgery is an attractive approach to continuously collect evidence about the patient's status in order to estimate outcomes and enhance recovery. It enables the identification of patients who may be experiencing complications, who require immediate intervention, or who may be at risk for negative outcomes. The signs and symptoms typically monitored after surgery include the status of vital signs, pain status and management, general status in terms of nausea and vomiting, urinary status, gastrointestinal status, eating and drinking intake, oral intake versus intravenous intake, fever and signs of infection, and signs or symptoms of pain. In order to quantitatively estimate these parameters, the usual approach is through vital signs and pain monitoring. Serial blood testing and radiological analysis alert the clinicians about the possibility of other specific complications. Unfortunately, the basic requirements of human resources and awareness make an effective general monitoring approach not available, especially in low-income settings.

There is a need for a more scientific and standardized approach to postoperative patient monitoring that takes into consideration the inter-person variability and personalization of guidelines. Continuous data analysis is expected to be vital in order to facilitate timely intervention and avoid deterioration in status. With predictive analytics, there are significant opportunities for identifying complications early and suggesting a path for recovery. Consequently, there is recent interest in leveraging machine learning and related techniques to predict outcomes and guide patient care after surgery. Machine learning is increasingly being utilized in healthcare to achieve improvements in patient care and outcomes. A number of studies have proposed the integration of machine learning models as part of a framework for postoperative patient monitoring to enhance recovery, reduce negative outcomes, and ensure patient safety.

4.1. Current Practices in Postoperative Monitoring

4.1. Postoperative Monitoring Postoperative monitoring is performed in the recovery room, intensive care unit, or surgical ward. Standard monitoring includes the assessment of consciousness, breathing, oxygenation, blood pressure, and body temperature. In addition, continuous electrocardiographic, pulse oximetry, and respiratory rate evaluation should be conducted. Nevertheless, specific types of monitoring and the ways in which postoperative recovery is assessed vary widely among hospitals and, sometimes, even within institutions. Tools such as blood gas analysis, biochemical measures, and various scoring systems have all been shown to effectively provide insight into the quality of recovery. Importantly, monitoring of all forms is based entirely on subjective clinical judgment. Although some monitoring tools may be able to flag potential maloccurrences, such tools observe a single aspect only. Existing guideline statements emphasize that early identification and adequate observation of complications and the recovery phase are based on routinely conducted physical examinations. Prolonged intensive care admissions represent just the tip of the iceberg. Due to cost concerns, monitoring is mainly a manual, nurse-intensive procedure. Software and hardware advancements do exist but are not yet in daily practice. This is also the case for continuous capnography, which should be integrated into patient monitoring.

Despite the strong belief that frequent and consistent patient evaluation will reduce the incidence and severity of complications, improving care and the quality of the outcomes, no further definition as to the exact frequency and content of these evaluations has been made according to the concept of care for the acutely and critically ill patient. It is clear that normalcy in vital signs alone cannot be used as a standard to confirm a normal intraoperative homeostasis or a normal postoperative homeostatic setpoint for a particular patient. Moreover, vital signs are checked intermittently, and postoperative deterioration might occur between these intermittent observations. Failure to recognize postoperative complications, which is a form of delayed response, is associated with increased adverse events and mortality. The health care cost for patient safety is more than $60 billion a year. Moreover, health care professional burnout is closely linked to patient safety because physicians are often the first health care professionals to be informed of a mortality event attributed to delayed response with potential maloccurrence. Finally, the one solution fits all practice should be abandoned. Prerequisites for strategies regarding monitoring should be standardized, and an appropriate methodological approach must be recognized to assess the performance of different systems in terms of responsiveness.

4.2. Advancements with Machine Learning

Advanced postoperative monitoring is perhaps the most surprising advancement stemming from machine learning. Not only can machine learning handle the complex data streams required for real-time decision-making, but it can also leverage large patient databases to develop models that predict postoperative outcomes. Predictive modeling is being used to monitor ambulatory heart failure patients and identify those at risk of complications that would require an inpatient visit. This subsection provides a comprehensive review of the recent developments in postoperative monitoring. Applications: Previous work supports the utility of predictive modeling for situations such as the early detection of acute kidney injury, cognitive decline, hemodynamic instability, hypoxemia, postoperative opioid-induced respiratory depression, and prediction of length of stay. A pilot study illustrates the usefulness of a machine-learning model for the early detection of adverse events during the first postoperative day. Future postoperative monitoring applications may also consider integrating wearable and mobile applications, since continuous monitoring has been possible with the ongoing development of these technologies.

Postoperative Care Systems: The development of postoperative monitoring tools highlights the value generated by such interventions, with multimodal perioperative interventions likely to become a standard part of the modern healthcare system. With support from health systems, these automated monitoring and decision support systems might free healthcare providers to deliver interventions of greater value. Other stakeholder perspectives for this cross-specialty work could include that of the patient, especially with the increased trend in consumer-directed healthcare. Empowering patients with information about their surgical recovery or their appropriate monitoring may activate them to engage as partners in their care delivery, so that the integration of such models into wearable technologies and mobile applications could also enhance patient engagement. This cross-specialty discussion advocates for developing and implementing monitoring systems that are automated and intelligent, either through clinical decision support tools or machine learning, to improve the safety of perioperative care. The challenges regarding studies of early detection include which diagnosis and which population to consider, deciding the timing for diagnosis and how to account for potential latent periods associated with complications, and identifying those who will change management in light of the information and those who will not. More broadly, the automated monitoring process in the operating room and beyond has potential implications for the length of stay. A study examining machine-learning models for the prediction of postoperative pulmonary complications developed a range of predictive models integrating features of anesthesia and of the surgical procedures. A significant negative association was demonstrated between major pulmonary complications and shorter discovered after-surgery functionality, compared with traditional pulmonary complication criteria.

5. Challenges and Future Directions

Ethical, legal, and regulatory considerations are paramount with respect to using AI in medical technologies. There are a mix of challenges, including those centered around privacy, informed consent, security of accessible models, general data protection rights, and avoiding and mitigating bias created by algorithms reliant on population data. One of the main concerns surrounds the use of machine learning algorithms in clinical practice for medical planners, particularly in the decision support or clinical trial phases, because there are many factors that affect the treatment plan once it has been determined with the machine learning model, and it is often harder to get them subjected to a clinical trial. Concerns linked to these practical challenges include practicalities on how to combine a set of models, on how to get both user and patient agreement to do a "combined therapy," and on data protection rights. Added to this, to show that the impact of new ideas is important for cancer cure, there may be a change in this process in England leading to a lack of clarity.

It is necessary to remember that a model that has been trained on one data set and performed well on a second data set may not perform effectively in a third set. While models should be developed with robustness, generalizability, and transportability in mind, further validation with existing and/or new data and/or adjusting the model architecture to new data is needed. Lastly, clinicians' opinions and the overall clinical workflow have to be considered since some considerations are made from a systemic point of view, and practitioners are those who must use the tools developed to create a workflow that is efficient and "quicker." There are a number of interesting challenges as the models currently applied in the clinical theater are based on aggregate statistics and probabilities that do not consider the individual patient. Mitigating factors and open challenges exist, but the gains are substantial in machine learning as part of the development of surgical methodologies. The logical next step is to test and use these models in real scenarios.

5.1. Ethical and Regulatory Considerations

To use such algorithms in practice, important issues related to ethical principles and regulations need to be addressed. Patient consent, use of data, and algorithm transparency that allows for tracking and understanding its decisions need to be tackled. Additionally, procurement, usage, understanding, and potential biases of an algorithm that uses medical data should be transparent for potential patients, end-users, and stakeholders. If the AI system makes a high level of errors that are critical for the patient outcome, or if the AI system shows high levels of bias, relying entirely on the AI system might be detrimental to patient care and outcomes. ML algorithms can encode biases present in the population, health, and training data, which will then enter into, be propagated by the model, and translate into biased outputs, impacting patient groups adversely. This is particularly critical in the medical field, and if the data includes certain demographic features, the algorithmic application should be carefully evaluated and might not be acceptable. There is no doubt that regulations should support the use of AI across health care. Some treatments cannot remain innovative and conducive. Although the benefits of using ML algorithms in health care are numerous, the number of compliance with these regulatory acts by AI companies is decreasing. Lack of guidelines is a significant feature in the current regulatory legislation. The FDA has only recently issued a draft guidance that advises the FDA to release audited material addressing the risks, approaches, limits, and strategy for mitigating and monitoring hazards. Ethical and regulatory guidelines should be implemented to ensure the responsible use of these AI applications. The development of these guidelines should be from a multidisciplinary leadership team. These medical professionals are also developing and implementing leadership guidelines of the Department of Health and Human Services. Moreover, developers, clinicians, and regulatory bodies should hold a continuing discussion concerning the proofs essential for application, trials, approval, and monitoring for algorithmic devices. This should be based on ethical principles of beneficence, doing good to the patient; and non-maleficence, avoiding harm. A very cautionary principle can lead to innovation paralysis, as surgical AI research would require a bigger and more time-consuming investment to reach equivalence, even though such a standard for 'normal' surgery may be too high. ML training with medical images yields important data and a surprising lack of transparency in machine learning development with proprietary algorithms and limitations in fully explaining their methods. Regulatory applications must have further discussions on what is required for an FDA submission, including a suggested number of iterations needed to achieve the performance and safety benchmarks. It would challenge the ethical principles of beneficence and non-maleficence to expand or not support a frame of accreditation like this. This means that setting a new benchmark for certification research to take a long time, are nondiscretionary and do not damage the function of all other models permissible because they cannot be met at once. Developing robust research and molecular engineering should be the aim of government funding. The process is changing, and teams are expected to report control results annually within the AI grant program. Ethical and regulatory considerations are a new cohesive writing, discussing the challenges in AI development at length before reaching the main text appropriate for benefit proponents. Excessive detail in some sections should be addressed. A conclusion in the minority race and machine learning model dataset is particularly lacking.

5.2. Opportunities for Further Research

Much of the current academic literature regarding machine learning is technical, offering insight into the latest algorithms and methods. However, surgical stakeholders can only justify investment in AI and IT-related technologies if they are stratified on their effectiveness to enhance hospital and surgical outcomes. A gap exists in the literature regarding the application of machine learning algorithms to these three domains. In the context of preliminary research, the PICO strategies included here can offer guidance for future systematic reviews. Given the large number of potential applications, we have suggested a series of secondary questions that are ready for investigation.

Challenges and future prospects for machine learning approaches to enhance surgical outcomes have been presented in the previous subsection. There is much more work to be done, despite the critical need for technique enhancement in all the aforementioned areas. A list of the open challenges broken down into three primary themes is presented here.

Theme 1: Proposal of More Sophisticated Techniques • There is an urgent need to pursue more sophisticated methods aimed at producing tailored recommendations for use by surgeons in planning, operating, and post-care. This is necessary because accurate complex data amalgamation methodology does not guarantee the development of decision-support models that offer value for surgical end-users. We need to explore new techniques in order to develop data integration outputs that are tailored to the three potential end-users.

Theme 2: Examination of Further Applications of the Three Domains of Current Research • Once primary studies are conducted in each domain, we must engage surgeons, alongside other stakeholders, to identify other important questions related to preoperatively, intraoperatively, and postoperative care that machine learning tools could be valuable in addressing.

Theme 3: Evaluation of Clinical Trials, Policy, Ethics, and Legal Issues • We must recognize that this field is still in its relative infancy. Although there is some strong evidence supporting the potential utility of machine learning in these domains, further research is required to fully validate the existing reports. • Of particular importance are two-phase large-scale clinical trials of its utility to not only improve surgical outcomes (or reduce complications) in greater populations of patients but also to evaluate the economic feasibility of such improvement. • We note that it is not enough to demonstrate that a 'treatment' (i.e., machine learning guided intervention) is effective; one must also demonstrate that it is an efficient use of resources. • Such research could represent 'low hanging fruit' that could demonstrably pave the way for accelerated translation of machine learning tools into surgical practice. • Lastly, we recommend that healthcare researchers, scholars, and practitioners in different AI disciplines collaborate and develop multidisciplinary AI research groups conducting research on AI challenges and developing up-to-date AI technology in surgical care.

The final gap in the existing literature that may also serve as an opportunity for further research is the examination of blockchain for surgical data management and enhancing surgical system interoperability. In summary, considering the risks, opportunities, significant outcomes, and challenges surrounding the rapid evolution of AI, stakeholders in surgery need to consult and take advantage of chances, take necessary precautions, and allow technology to be in accordance with the intended values. The overall effect, beyond the rhetoric, would be committing to and preparing for an intentional intervention to enhance surgical outcomes through technology.

6. Future Direction

6.1. Deep learning is an unceasing process and becomes more accurate over time by collecting new data and implementing modern algorithms. Consequently, we expect that 3D (or even 4D) deep learning models would become more useful in the next few years for predictive and personalized patient management. It is evident that predictive models using real-time data and other data sources will lead to improved patient care. We anticipate that the concept of AI will be integrated into surgical training and education, which will help newcomers improve their surgical skills. Further, surgical treatment is shifting towards patient-tailored care, and these AI-based models would facilitate the use of treatment strategies with enhanced efficiency and recovery time.

6.2. There is continuous evolution in the world of AI, so it is important to update and adapt the technology to remain current in the ever-evolving health care industry. In order to be more acceptable to the surgeons and patients, it is expected that the future trend of AI in surgery would integrate more thorough feature analysis and provide more accurate predictions in perioperative and postoperative periods. As a summary, 3D and 4D predictive models using ML and especially DL techniques would become more powerful for personalized and predictive surgical evaluations and planning. Development of such models integrated into real-time data analysis will be of clinical and research impact, providing a tailored recovery protocol for each patient. Model adaptation based upon the most recent patient data and clinical advancement will drive the next generation of surgical practice.

7. Conclusion

There is ample evidence that machine learning has the potential to improve preoperative planning, intraoperative guidance, and postoperative monitoring, thus leading to enhanced surgical outcomes. Different components of an AI-based system have to be integrated into a common network through machine learning techniques, with more focus being given to deep learning methods. Additionally, there is a necessity to make efforts toward creating new AI tools, which can connect the three components of preoperative, intraoperative, and postoperative research into a continuous learning system. Our review has also identified key challenges emerging when moving from research applications to real clinical implementations of such AI tools and the necessity to address these challenges to enable larger-scale research, evaluations, and translation into clinical practice. Importantly, necessary considerations for the integration of AI in surgery will be the topic of future research.

Thus, machine learning is expected to have a transformative effect on planning surgery, guiding intervention, and monitoring postsurgical recovery. Creating new research sketches for furthering the known applications with a mix of regulated and unregulated data and concurrent handling of traditional surgery metrics but without divulging the involvement of AI in these measurements could also be an area of targeted research in the future. These development strategies will permit analyses and decision-making that will lead to surgical practices that offer improved outcomes based on evidence and from the experiences of leading surgeons and surgical institutions. We expect these collaborations will give new perspectives and lead to innovations in the ways surgery is planned and executed in the years to come.