Review Article

Artificial Intelligence and Machine Learning for Risk Prediction in Surgery

Shamsul Masum[1], Adrian Hopgood[2]*, Jim Khan[3,4]

Artificial Intelligence and Machine Learning

Artificial Intelligence (AI) has been a field of research for more than 70 years, with the goal of mimicking human thought processes in a computer. There were early successes in the subgenre of expert systems, designed to capture knowledge in specialist domains like medicine. These expert systems are part of a broader family of AI known as knowledge-based systems, which contain explicit knowledge expressed in human-readable form [1]. However, the current wave of excitement is largely driven by a different model, namely machine learning (ML). The idea is that by showing a computer algorithm thousands of examples of images or other forms of data, it will learn to associate those examples with their correct classification [1]. A key characteristic of ML is generalization. When presented with an image or data pattern that it has not seen before, the algorithm can classify it reliably, provided that similar examples existed in the training set. Unsurprisingly, many surgeons have limited knowledge of AI and ML. Nevertheless, the fusion of their experiences from the medical domain with those from the computing sciences has led to a significant interest in the developing discipline of health informatics.

AI for Risk Prediction

AI has shown great potential in surgery risk prediction, with various innovations and ideas that could shift regular practice. Jalali et al. have used ML models to predict the mortality risk and length of stay (LoS) following the Norwood surgical procedure [2]. Their study compared ridge logistic regression, decision trees, random forest, gradient boosting, and deep neural networks as ML algorithms. They found that deep neural networks outperformed the other algorithms and they claimed that their model may help clinicians and organizations make decisions. Masum et al. have developed prediction models for patient outcomes following colorectal cancer surgery [3, 4]. They compared different ML algorithms to extract important features and predict the LoS, readmission, and mortality following colorectal surgery. They showed that support vector regressor (SVR) algorithms perform best for the LoS prediction and bidirectional long short-term memory (Bi-LSTM) performs best in predicting readmission and mortality.

Chang et al. have used AI algorithms to predict mortality for patients with congenital heart disease undergoing cardiac surgery [5]. They considered multilayer perceptron, random forest, extra trees, stochastic gradient boosting, Ada boost classification, and bag decision tree for their study. They found that random forest outperformed other algorithms and suggested that it should be considered to predict mortality following cardiac surgery in patients with congenital heart disease. Bihorac et al. introduced 'MySurgeryRisk', an automated predictive tool that can predict the risk of postoperative complications and death after surgery [6]. They used a generalized additive model and random forest, and applied them to preoperative electronic health

records data. Corey et al. introduced 'Pythia', which identifies high-risk surgical patients from electronic health records [7]. They used least absolute shrinkage and selection operator, penalized logistic regressor, random forest models, and extreme gradient boosted decision trees as ML algorithms.

The supervised learning techniques of ML have been used recently for COVID-19 mortality risk and severity prediction. The COVIDSurg cohort study developed a machine learning-based risk score to predict postoperative mortality risk in patients with perioperative SARS-CoV-2 infection [8]. They used logistic regression, decision trees, and random forest as ML algorithms to find the important features and develop the prediction models.

The way forward?

AI is playing an essential role in our daily lives. It has gained trust and reliability in tasks like internet search, wellbeing monitoring, shopping recommendations, and smart-home devices. In most of these domains, failure is inconvenient but not serious. In contrast, how does AI perform in predicting risk in the medical domain, and can it be trusted? Different AI approaches have been used to build prediction models for the risk of disease and other medical outcomes. However, clinical use and trial of these models are limited due to the model limitations and the workload of the clinicians. These models are often criticised for their dependence on reliable datasets, validation issues, and the opacity of a 'black-box' model with little understanding or explanation capability. ML requires large databases and clean data to perform the risk prediction task, but clean data and large databases are rare in the medical arena. Moreover, access to such data is a hurdle for AI researchers as most of these datasets are not publicly available.

One of the limitations of these risk-prediction models is that they are not externally validated. They split the dataset into training and testing sets to evaluate the model, whereas the best evaluation method would be to test it on external or different datasets. Using risk-prediction models as an opaque black box can reduce trust, whereas an explainable AI model would help to build trust. Explainable AI can be achieved by complementing the ML with knowledge-based representations to explain how the predictions were performed, how features interact, the importance of variables, and confidence intervals in predicting the risk of the disease. There is a gap between AI and clinical researchers in understanding their respective areas and how they can help each other. Researchers from both fields need to collaborate more so that clinical researchers know the basics of AI methods and AI researchers get to know the basics of medical risks.

Conclusion

In conclusion, AI has shown promising results and has massive potential in risk prediction. In addition to predicting short-term outcomes, AI can be used to predict survival and prognosis following various forms of surgery. However, a proper framework should consider issues like collaboration, data availability, transparency of the developed models, and randomized clinical trials.

References

1. Hopgood AA. Intelligent Systems for Engineers and Scientists: A Practical Guide to Artificial Intelligence, 4th edition, CRC Press (2022).

2. Jalali Ali, Lonsdale H, Do N, et al. Deep Learning for Improved Risk Prediction in Surgical Outcomes. Scientific Reports 10 (2020): 9289.

3. Masum S, Hopgood A, Stefan S, et al. Data analytics and artificial intelligence in predicting length of stay, readmission, and mortality: a population-based study of surgical management of colorectal cancer. Discover Oncology 13 (2022): 1-17.

4. Masum S, Hopgood A, Stefan S, et al. Data analytics and artificial intelligence to predict length of stay, readmission, and mortality after colorectal cancer surgery. European Journal of Surgical Oncology 47 (2021): e5.

5. Chang Junior J, Binuesa F, Caneo LF, et al. (2020). Improving preoperative risk-of-death prediction in surgery congenital heart defects using artificial intelligence model: a pilot study. PLoS One 15 (2020): e0238199.

6. Bihorac A, Ozrazgat-Baslanti T, Ebadi A, et al. MySurgeryRisk: development and validation of a machine-learning risk algorithm for major complications and death after surgery. Annals of Surgery 269 (2019): 652.

7. Corey KM, Kashyap S, Lorenzi E, et al. (2018). Development and validation of machine learning models to identify high-risk surgical patients using automatically curated electronic health record data (Pythia): A retrospective, single-site study. PLoS Medicine 15 (2018): e1002701.

8. COVIDSurg Collaborative. Machine learning risk prediction of mortality for patients undergoing surgery with perioperative SARS-CoV-2: the COVIDSurg mortality score. British Journal of Surgery 108 (2021): 1274-1292.