

Machine Learning in Predicting Outcomes of Brain Tumor Resection Surgeries: A Scholarly Review

Arman Raoufian

Northwest High School, 13501 Richter Farm Rd, Germantown, Maryland, 20874, USA; infectlate@gmail.com

ABSTRACT: Being able to predict the outcome of a patient following a brain tumor removal is essential to providing care for those patients. Conventional tools for predicting the outcome of patients who undergo a brain tumor removal are not able to accurately model the complex relationships between the various variables of each patient and their condition. Machine learning, however, has the potential to provide a more powerful framework for predicting the survival of patients after brain tumor removal, the complications that may occur following those operations, and the quality of life of those patients following the operations. Through searching databases like PubMed and Google Scholar for articles published between 2015 and 2025, it becomes evident that based models have higher accuracy in predicting the outcomes of patients after brain tumor operations than conventional models. One systematic review of studies that utilized machine learning to model the outcomes of patients after brain tumor operations discovered that the median area under the curve (AUC) score of the models was 0.83, with a median accuracy of 94.5%. Despite these positive findings, however, various barriers to the adoption of these models into the clinical care of brain tumor patients exist. As a result, models will likely be utilized as a supplement to the clinician's care for those patients rather than as a means of replacing the clinician's judgments of those patients' outcomes.

KEYWORDS: Biomedical Engineering, Biomedical Imaging, Machine Learning, Brain Tumor Resection, Outcome Prediction.

■ Introduction

Brain tumor resection aims to remove as much of the tumor mass as possible while preserving critical functions of the brain, such as speech, movement, and memory. The recovery, survival, and quality of life of patients who undergo brain tumor resection depend on various factors, such as the location of the tumor within the brain, the grade of the tumor's histology, and the age and general health of the patients.¹ Each of these factors can have a substantial effect on the outcome of the resection procedure. Accordingly, it is critical to be able to accurately predict the outcome of brain tumor resection for each patient.

Various scoring systems and statistical models are used to attempt to estimate the outcome of brain tumor resections. These models are helpful in providing an estimate of the outcome of brain tumor resection for each patient. However, these models often evaluate each of the model variables separately and make assumptions regarding the linear relationships between those variables and the outcome of the brain resection procedure.² As a result, these models may not account for the effects of nonlinear relationships between the variables on the outcome of the resection of brain tumors. Therefore, there is a need for improved tools for predicting the outcome of resection of brain tumors.

Machine learning models can help address the shortcomings of conventional models for predicting the outcome of brain tumor resection. Machine learning models learn the relationships between various variables from data without having to prescribe a specific model form to the data.³ By training a machine learning model on historical records of brain tumor

resections, machine learning models can provide individualized predictions regarding the outcome of a brain tumor resection procedure for a newly admitted patient. In this review, the literature on the use of machine learning models for predicting the outcome of brain tumor resection is summarized and evaluated for its ability to produce models that can more accurately predict the outcome of the procedure than the conventional models for predicting its outcome.

■ Methodology

This paper is a narrative literature review with a focus on answering the question of to what extent does machine learning, as it pertains to the field of neurosurgery and brain tumor resection surgery, provides stronger predictions of patient outcomes compared to the traditional models that have been used in neurosurgery to date.

A search was performed on the databases PubMed and Google Scholar, utilizing the keywords: machine learning, brain tumor resection, neurosurgery outcome prediction, survival modeling, and postoperative complications. The search was limited to peer-reviewed journals in the English language between January 2015 and April 2025.

The studies were included if they utilized machine learning to make predictions of patient outcomes in the field of neurosurgery, reported some form of performance metrics for their models, and their outcomes related to the resection of brain tumors. Studies were excluded if they did not report performance metrics for their models, they were not related to neurosurgery, they were not peer reviewed, or they were conference abstracts without a published article with the same content.

There were over 200 articles found during the initial database search. However, only 22 fulfilled the requirements for inclusion within the review. These in-depth studies will be summarized in this literature review according to the following main topics: traditional models, machine learning models, evidence, and challenges. The findings presented in this review are a synthesis of the current literature on the topic, rather than a systematic review of the literature in the way that would be required for a PRISMA study.

Background: Traditional Approaches to Prognostication in Neurosurgery:

Prognostication is the estimation of the likely outcome of a patient's current medical condition. Because the conditions treated by neurosurgeons are often life-threatening conditions, it is essential to be able to accurately estimate the likely outcome of a patient with such a condition.⁴

Clinical Scoring Systems:

Experienced clinicians formulate an impression of the risk based on the history, physical examination, physiological, and intraoperative findings. While this method incorporates the clinician's judgment at the time of the procedure, the risk of inter-observer variability and bias makes this method less suitable for comparing the risk between institutions.⁴

Clinical Scoring Systems:

Several standardized scales have been established to assess the prognosis of TBI. The Glasgow Coma Scale (GCS) evaluates the level of consciousness of the patient and is the most commonly used scale in the evaluation of traumatic brain injuries.⁵ The Hunt and Hess scale is used to assess the neurological deficits of patients who suffer from subarachnoid hemorrhage.⁶ The Karnofsky Performance Scale Index evaluates the functional status of cancer patients and is inversely correlated with the mortality rate of those patients.⁷ These scales allow for better communication between healthcare team members; however, they all have the disadvantage of reducing the complex clinical presentation of a patient to a series of numbers, they can only assess certain static variables of the patient, and they do not consider the interaction between the various variables that contribute to the patient's outcome.⁸

Able to carry on normal activity and to work; no special care needed	100	Normal no complaints; no evidence of disease
	90	Able to carry on normal activity; minor signs or symptoms of disease
	80	Normal activity with effort; some signs or symptoms of disease
Unable to work; able to live at home and care for most personal needs; varying amounts of assistance needed.	70	Cares for self; unable to carry on normal activity or to do active work.
	60	Requires occasional assistance, but is able to care for most of his personal needs.
	50	Requires considerable assistance and frequent medical care.

Unable to care for self; requires equivalent of institutional or hospital care; disease may be progressing rapidly.	40	Disable; requires special care and assistance.
	30	Severely disabled; hospital admission is indicated although death not imminent.
	20	Very sick; hospital admission necessary; active supportive treatment necessary.
	10	Moribund; fatal processes progressing rapidly.
	0	Dead

Figure 1: Karnofsky Performance Scale. The Karnofsky Performance Scale is used to assess a patient's functional status and ability to perform daily activities. Scores range from 100 (normal, no complaints) to 0 (death), with lower scores indicating decreased functional ability and poorer clinical outcomes. This scale is commonly used in neurosurgery to help evaluate prognosis and guide treatment decisions. Created by Raoufian using Canva.

Imaging-based Prognostication:

Neuroimaging allows for the collection of objective and reproducible prognostic data. An MRI can reveal data about the enhancement of the tumor, the presence of edema, diffusion restriction, and mass effect, all of which are correlated with the grade of the tumor and the difficulty that may be encountered during surgical removal of the tumor.⁹ However, interpretation of the MRI reports relies upon the expertise of the interpreting radiologist to recognize patterns of the tumor on the MRI, and any subtle findings that are not visible to the unaided eye can be easily missed.¹⁰

Annotated Brain MRI Scans Highlighting Tumor Regions for Predictive Modeling

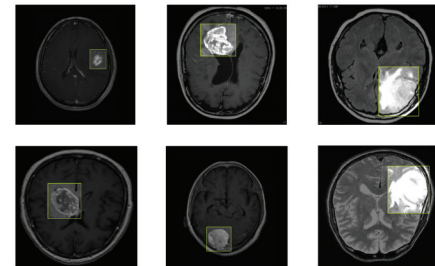


Figure 2: Axial brain MRI images that display the tumor regions outlined by bounding boxes and utilized as the inputs for the machine learning models. The variety in the size, shape, intensity, and location of the tumors in the brain illustrates the heterogeneity of these types of tumors. Each of these features can be extracted from the MRI images to analyze the tumors and predict the outcome of the surgical procedures performed on patients with these types of tumors. Source: Roboflow Brain MRI Dataset (<https://universe.roboflow.com/brain-mri/mri-rskcu>). No changes were made to the data. (CC BY 4.0).

Thus, standardized, objective, and reproducible, these results, too, are confirmed by the professional as there are certain atypical findings and those visible to the naked eye.¹⁰

Machine Learning Methods in Neurosurgical Outcome Prediction

Machine learning encompasses a variety of algorithms that computers use to learn from examples in the training data instead of relying on predetermined equations to calculate the outcome.¹¹ Medical records of patients who underwent certain neurosurgical procedures in the past are used to train machine learning algorithms to predict the outcome that will be observed for new patients undergoing those same procedures.² Because machine learning algorithms can examine numerous variables simultaneously to find correlations between variables that impact the outcome of surgical procedures, they have

attracted the interest of researchers who wish to incorporate these methods alongside the tools that physicians already utilize today.

Supervised Learning Methods:

The most common form of supervised learning is logistic regression. Using logistic regression, the probability that a patient will have a positive or negative outcome from a surgical procedure is calculated from various features of the patient and the procedure. The method is simple to implement, but it cannot account for the effects of nonlinear relationships between variables.² Decision trees are another form of supervised learning. Decision trees model the outcome of a surgical procedure by assigning thresholds to various features of the patient or the procedure. Each internal node within a decision tree represents some feature of the patient or procedure, and each branch descending from that node contains the possible values for that feature.¹³ Using only single decision trees to model surgical outcomes leads to overfitting, which introduces other methods into the equation.

Support vector machines (SVMs) find the optimal boundary that separates examples in training datasets with different outcomes. SVMs can model nonlinear relationships between features and surgical outcomes.¹⁴ As such, they can offer better discrimination between outcomes than logistic regression, but at the cost of interpretability.

Ensemble Methods:

Ensemble methods such as random forests and gradient boosting algorithms create groups of decision trees and combine the results from each individual tree within the group to find the outcome of the surgical procedure for a patient. Random forests use a group of decision trees that are built independently of each other.¹⁵ Gradient boosting algorithms build decision trees sequentially such that each tree corrects the errors made by the previous tree within the ensemble. The most common form of gradient boosting algorithms is gradient boosting and has been used in a variety of neurosurgical outcome prediction studies.¹⁹

Deep Learning:

Deep learning methods use convolutional neural networks (CNNs) to calculate the outcome of neurosurgical procedures. CNNs can analyze medical images of the patients to determine features related to the outcome of the procedure without having to create features manually.¹⁶ Deep learning methods achieve the highest accuracy with large datasets of medical records and can simultaneously process various types of data, from medical images to medical records.¹⁷

Evidence from Studies

The most comprehensive review of the use of ML models to predict neurosurgical outcomes was that by Senders *et al.* in 2018.² Senders *et al.* reviewed 30 studies using ML methods for outcome prediction after various types of neurosurgical procedures. Across the reviewed studies, ML models

achieved median accuracy rates of 94.5% and AUCs of 0.83 in predicting outcomes after procedures. Furthermore, ML models achieved a median 15% improvement in accuracy and 0.06 improvement in AUC when compared to models based on logistic regression. Additionally, in some studies, ML models outperformed both established outcome indices and expert surgeons in predicting outcomes of patients undergoing procedures. Thus, these results indicate that the advantages of ML models over conventional statistical models for outcome prediction in neurosurgery are consistent across multiple studies.

In the specific domain of outcome prediction following brain tumor resection, van Niftrik *et al.* performed studies comparing various boosting algorithms to conventional statistical models in predicting outcomes in 668 patients undergoing elective intracranial tumor surgeries at a single center.¹⁹ The outcomes that were assessed included the development of complications within the first 24 hours of surgery. Gradient boosting algorithms were the performing models in this group of studies, achieving an accuracy of 0.70 and an AUC of 0.73, while the conventional statistical model achieved accuracy rates of 0.59 and an AUC of 0.64. However, these studies were limited to only one surgical center.

Farrokhi *et al.* performed studies that applied supervised machine learning algorithms to predict outcomes following brain stimulation surgeries in 501 patients from multiple surgical centers. Models were developed using 70% of the patients' data and tested on the remaining 30%.²⁰ Gradient boosting machines were among the performing supervised machine learning algorithms for outcome prediction after brain stimulation surgeries. For example, machine learning algorithms achieved AUC values of 0.86 for any surgical complications, 0.91 for complications within the first 12 months after surgery, 0.88 for the need to return to the operating room, and 0.97 for infection after surgery. Factors that were found to impact the outcomes of deep brain stimulation surgeries included the age of the patient, their body mass index, the side on which the surgery was performed, the sex of the patient, and whether they had Parkinson's disease.

In efforts to use MRI scan images to determine the survival of patients with brain tumors, Chato and Latifi assessed the performance of various machine learning algorithms on the BraTS 2017 data set of 163 patients with gliomas and associated MRI scan images and survival times.¹⁶ The features that were used to train the machine learning models included volumetric, statistical, intensity texture, histogram, and deep learning features of the patients' MRI scans. The best results were achieved with machine learning models that used deep learning features of the patients' MRI scans, indicating that features extracted from MRI scans by deep learning methods outperform conventional features used by conventional machine learning classifiers.

In addition to the use of machine learning models trained directly with electronic health record data, Ive *et al.* performed a study in which natural language processing (NLP) models were used to extract information from the electronic health records of 2,268 patients undergoing elective neurosurgical procedures.³ A Random Forest machine learning model that

utilized information extracted by natural language processing models achieved a recall of 0.87 for identifying patients who would experience unplanned admissions to the intensive therapy unit (ITU), with an AUC of 0.99. Furthermore, by using this model, the proportion of patients with unplanned ITU admissions that were missed by the attending neurosurgery experts was reduced from 36% of cases to just 4%.

Badjatia *et al.* performed a review of the literature on the use of machine learning models to predict outcomes following traumatic brain injuries (TBIs).¹² The authors found that machine learning models that incorporated various types of data from the patients following TBIs, including clinical features, neuroimaging data, and autonomic nervous system metrics, can improve the prediction of outcomes following TBIs relative to conventional methods of predicting those outcomes. The review also indicated that many current machine learning models have challenges with issues like data variability, model interpretability, and overfitting. Thus, the authors call for the standardization of these models and the testing of these standardized models outside of the groups with whom they were developed.

Table 1: Summary of Primary Studies: ML Models for Neurosurgical Outcome Prediction

Study (Author, Year)	ML Model(s) Used	Dataset / Sample	Outcome Predicted	Performance Metrics	Comparison Baseline
Senders <i>et al.</i> , 2018	Systematic review of multiple ML model types (LR, ANN, SVM, RF, gradient boosting)	Review of 30 studies (PubMed/Embase search to Jan 2017)	Survival, recurrence, symptom improvement, adverse events across neurosurgical conditions	Median accuracy 94.5%; median AUC 0.83; ML outperformed LR by median 15% accuracy and 0.06 AUC	Logistic regression; established prognostic indices; clinical experts
van Niftrik <i>et al.</i> , 2019	Gradient boosting algorithms (multiple compared); best model selected from out-of-sample test set	668 patients; single-center prospective registry (University Hospital Zürich, June 2015–May 2017)	Early postoperative complications within 24 h of intracranial tumor surgery (Clavien-Dindo classification)	Accuracy 0.70 (CI 0.59–0.79); AUC 0.73; sensitivity 0.80; specificity 0.67. Conventional model: accuracy 0.59, AUC 0.64	Conventional statistical model (logistic regression)
Farrokhi <i>et al.</i> , 2020	Supervised ML algorithms (multiple compared, including gradient boosting machines); trained on 70%, validated on 30%	501 patients; multicenter DBS surgery registry	Any complication; complication within 12 months; return to OR; infection	AUC: any complication 0.86; within 12 months 0.91; return to OR 0.88; infection 0.97	Logistic regression (used to evaluate risk factors)
Chato & Latifi, 2017	Multiple ML methods (SVM, KNN, linear discriminant, tree, ensemble, logistic regression) plus deep learning features extracted by pre-trained CNN	BraTS 2017 dataset; 163 glioma patients with MRI and survival labels (short-, mid-, long-term)	Overall survival classification (three classes: short-, mid-, long-term survivors)	Best accuracy achieved using deep learning features from pre-trained CNN combined with ML classifier (exact figure not reported in abstract)	Conventional 1 ML classifiers without deep learning features
Ive <i>et al.</i> , 2025	NLP (CogStack-MedCAT) + Random Forest classifier trained on extracted clinical concepts from EHR free text	2,268 elective neurosurgery patients; University College London Hospital	Unplanned ITU (intensive therapy unit) admissions post-surgery	Random Forest recall 0.87 (CI 0.82–0.91) for ITU admissions; AUC 0.99; reduced unplanned ITU cases missed by clinicians from 36% to 4%	Human expert clinical judgment (missed 36% of unplanned ITU cases)
Badjatia <i>et al.</i> , 2025	Narrative review of ML approaches (clinical features, neuroimaging, autonomic nervous system metrics)	Review article — no primary dataset	Review of ML for TBI outcome prognostication	Review conclusion: ML integrating multimodal data enhances early detection of deterioration and improves outcome prediction; no single performance figure reported	Traditional prognostication methods

AUC = Area Under the Receiver Operating Characteristic Curve; BraTS = Brain Tumor Segmentation Challenge; CNN = Convolutional Neural Network; DBS = Deep Brain Stimulation; EHR = Electronic Health Record; ITU = Intensive Therapy Unit; KNN = K-Nearest Neighbors; LR = Logistic Regression; MRI = Magnetic Resonance Imaging; NLP = Natural Language Processing; RF = Random Forest; SVM = Support Vector Machine; TBI = Traumatic Brain Injury.

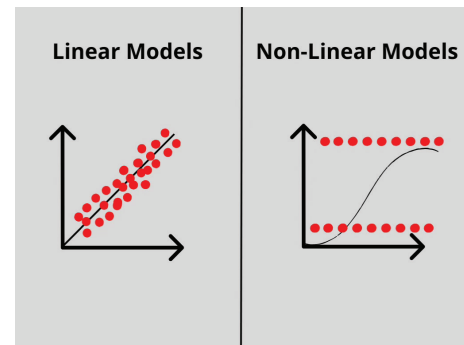


Figure 3: Machine Learning Model Classification Diagram. Static versus dynamic models also contrast logistic regression to ML. Static assumes linear impact between variables, meaning that for the variables' minimum and maximum range, risk can be over- or under-estimated by additive strengths or weaknesses. Nonlinear assumes that demographic, comorbidity, and surgical variables impact one another in a nonlinear approach; at times, certain factors matter less or more depending on other conflicting/intersecting issues, thus distinguishing risk contribution for better accounting of how risk factors impact real-world risk. Created by Raoufian using Canva.

Specific studies in neurosurgery lend credence to this. An extreme gradient boosting model was applied as a precise algorithm to predict complications for intracranial tumor operation patients. Ultimately, from 668 patient data, the model reached a 70% accuracy and an AUC of 0.74, better than conventional statistical models for predictive power.¹⁹ ML has also predicted complications of deep-brain stimulation. Their models were supervised and had high discriminatory performance, with AUC from 0.88 for any complication to 0.97 for infection. These better predictions mean greater flagging of higher risk populations where ML models can more accurately predict preventative measures than conventional means, guiding better preoperative considerations or treatment options.²⁰

These studies show cumulative efforts where MLs and deep learning can analyze the high-dimensional clinical characteristics that are patterned beyond the efforts of standard statistical means, all leading to better postoperative complication prediction and risk of population awareness in neurosurgery for better surgical strategy, patient education, and individualized care.

■ Discussion

Overall, the studies reviewed show that machine learning outperforms current statistical tools in the neurosurgical field. The ability of machine learning to model the interaction of many different variables at once is beyond the capabilities of existing statistical models. Studies by Senders *et al.* have shown that machine learning models outperform statistical models in 30 different studies and in a variety of neurosurgical conditions. Furthermore, the two primary studies reviewed by van Niftrik *et al.* and Farrokhi *et al.* show that these results can also be obtained in tumor and neurosurgical patient populations.

Within these studies, machine learning models based on gradient boosting are among the performing models. Van Niftrik *et al.* discovered that gradient boosting was the performing model for the prediction of complications from intracranial tumors. Farrokhi *et al.* found that gradient boosting was among the performing models for Deep Brain Stimulation procedures.

Ive *et al.* discovered that by utilizing natural language processing to analyze the free-text clinical notes for patients, it was possible to create a model that contained additional information about the patients beyond the structured variables. This new model was able to reduce the percentage of unplanned ICU admissions that were missed by the clinical staff from 36% to 4%. Thus, it would be beneficial for machine learning models to incorporate both structured and unstructured data.

Because machine learning models were found to have such high performance within the clinical field, it is possible to use these models to provide individualized outcome predictions for each patient. Furthermore, these models could be integrated into the current neurosurgical population as a tool to help surgeons make informed decisions about each patient's neurosurgical procedure. While these models may assist in the making of surgeons, the surgeons will still hold complete control and consideration of each patient and the information that cannot be encompassed within any of these models.

■ Weaknesses and Challenges

Single-Center Designs and Limited Generalizability:

The studies by van Niftrik *et al.* and Farrokhi *et al.* each used designs based on data from single institutions. For van Niftrik *et al.*, the model was tested on patients from a single center, with 668 patients included in that test. However, the authors state that external validation would be needed to determine how the model would perform on patients from other centers. Until such validation is performed, the generalizability of the models is not established.

Interpretability:

One of the main drawbacks of the types of models that performed best in both of these studies (gradient boosting methods and deep neural networks) is that they are often difficult to interpret. For instance, it is difficult for a surgeon to understand the reasons that the model provided a risk score to any given patient with a gradient boosting model. Although methods like SHAP and LIME can provide some insight into the decisions of these models, such approaches have yet to be utilized in studies.

Data Quality and Bias:

Because machine learning models learn from the data that is used to train them, the quality of that data can impact the performance of the models. If the training data does not reflect the diversity of the populations that are to be modeled, such as patients of lower socioeconomic status, patients from rural areas, or patients from countries other than Western Europe, the models may have poor performance with those patients. For instance, Ive *et al.* drew their data entirely from a hospital in London.

Ethical and Privacy Concerns:

In order to train these models, vast amounts of sensitive health information must be collected from patients. Additionally, patients' consent must be obtained for the collection of their health information. Finally, the lack of explanations for

the models' predictions presents difficulties for patients and physicians to contest the predictions of the models ethically. Finally, if the training data for these models is biased in some way, the models may present only discriminatory recommendations for certain groups of patients.

■ Conclusion

Machine learning consistently outperforms conventional statistical models in predicting important clinical outcomes following brain tumor resection neurosurgery. Studies by Senders *et al.* report that machine learning models achieve an average area under the curve (AUC) of 0.83 and accuracy of 94.5% in predicting postoperative outcomes, representing a 15-percentage-point advantage over conventional regression models. Subsequent studies have provided additional confirmations of the performance of machine learning models. For instance, Van Niftrik *et al.* found that gradient boosting models achieved an AUC of 0.73 and an accuracy of 0.70 in predicting the development of intracranial tumor complications following resection procedures. Similarly, Farrokhi *et al.* reported AUC values of up to 0.97 in deep brain stimulation machine learning models for the prediction of complications like infection. Ive *et al.* developed natural language processing (NLP) models to achieve similar results with missed admissions to the ICU following neurosurgical procedures, reducing the percentage of missed admissions from 36% to 4%.

Despite the demonstrated performance of machine learning models in brain tumor resection procedures, several challenges must be addressed before such models can be widely adopted into clinical practice. One challenge is validating the models in multiple institutions to ensure the generalizability of the results. Secondly, model performance across studies must be standardized to permit comparison. Thirdly, natural language processing and machine learning models must be explained to clinicians to gain their trust. Finally, data representing diverse populations must be collected to ensure equitable implementation of these models.

When developed to the standards described, machine learning models have the potential to improve the management of patients undergoing brain tumor resection procedures.

■ Acknowledgments

I acknowledge Dr. Sava for her informative information regarding the machine learning process and, of course, Prof. Virgel and Prof. Kay for guiding me through this research paper.

■ References

1. IBM. *What Is Machine Learning?* IBM Think. <https://www.ibm.com/think/topics/machine-learning>
2. Senders, J. T.; Staples, P. C.; Karhade, A. V.; Zaki, M. M.; Gormley, W. B.; Broekman, M. L. D.; Smith, T. R.; Arnaut, O. Machine Learning and Neurosurgical Outcome Prediction: A Systematic Review. *World Neurosurgery* 2018, 109, 476–486.e1. <https://doi.org/10.1016/j.wneu.2017.09.149>. PubMed PMID: 28986230
3. Ive, J.; Olukoya, O.; Funnell, J. P.; Booker, J.; Lam, S. H. M.; Reddy, U.; Noor, K.; Dobson, R. J. B.; Luoma, A. M. V.; Marcus, H. J.

- AI Assisted Prediction of Unplanned Intensive Care Admissions Using Natural Language Processing in Elective Neurosurgery. *npj Digital Medicine* 2025, 8, 549. <https://doi.org/10.1038/s41746-025-01952-0>. PubMed PMID: 40858789
4. Witsch, J.; Siegerink, B.; Nolte, C. H.; Sprügel, M.; Steiner, T.; Endres, M.; Huttner, H. B. Prognostication after Intracerebral Hemorrhage: A Review. *Neurological Research and Practice* 2021, 3 (1). <https://doi.org/10.1186/s42466-021-00120-5>
 5. Jain, S.; Iverson, L. M. Glasgow Coma Scale. *StatPearls*, National Library of Medicine. <https://www.ncbi.nlm.nih.gov/books/NBK513298/>
 6. Campos, A. Hunt and Hess Grading System (Subarachnoid Haemorrhage). *Radiopaedia*. <https://radiopaedia.org/articles/hunt-and-hess-grading-system-subarachnoid-haemorrhage>
 7. Schag, C. C.; Heinrich, R. L.; Ganz, P. A. Karnofsky Performance Status Revisited: Reliability, Validity, and Guidelines. *Journal of Clinical Oncology* 1984, 2 (3), 187–193. <https://doi.org/10.1200/jco.1984.2.3.187>
 8. Perel, P.; Edwards, P.; Wentz, R.; Roberts, I. Systematic Review of Prognostic Models in Traumatic Brain Injury. *BMC Medical Informatics and Decision Making* 2006, 6 (1). <https://doi.org/10.1186/1472-6947-6-38>
 9. Martucci, M.; Russo, R.; Schimperna, F.; *et al.* Magnetic Resonance Imaging of Primary Adult Brain Tumors: State of the Art and Future Perspectives. *Biomedicines* 2023, 11 (2), 364. <https://doi.org/10.3390/biomedicines11020364>
 10. Waite, S.; Grigorian, A.; Alexander, R. G.; *et al.* Analysis of Perceptual Expertise in Radiology—Current Knowledge and a New Perspective. *Frontiers in Human Neuroscience* 2019, 13, 213. <https://doi.org/10.3389/fnhum.2019.00213>
 11. Kamal, V. K.; Agrawal, D.; Kumar, A.; *et al.* Development, Validation and Clinical Utility of Prognostic Models for Patients with Traumatic Brain Injury in an Indian Setting Using Machine Learning and Traditional Approaches: A Study Protocol. *BMJ Open* 2025, 15 (10), e096275. <https://doi.org/10.1136/bmjopen-2024-096275>
 12. Badjatia, N.; Podell, J.; Felix, R. B.; *et al.* Machine Learning Approaches to Prognostication in Traumatic Brain Injury. *Current Neurology and Neuroscience Reports* 2025, 25 (1), 19. <https://doi.org/10.1007/s11910-025-01405-x>. PubMed PMID: 39969697
 13. Breiman, L. Random Forests. *Machine Learning* 2001, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
 14. Cortes, C.; Vapnik, V. Support-Vector Networks. *Machine Learning* 1995, 20 (3), 273–297. <https://doi.org/10.1007/BF00994018>
 15. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 2016. <https://doi.org/10.1145/2939672.2939785>
 16. Chato, L.; Latifi, S. Machine Learning and Deep Learning Techniques to Predict Overall Survival of Brain Tumor Patients Using MRI Images. *Proceedings of the 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE)*, pp. 9–14. <https://doi.org/10.1109/BIBE.2017.00-86>
 17. Park, Y. W.; Yoo, R.-E.; Shin, I.; *et al.* A Multinational Study of Deep Learning-Based Image Enhancement for Multiparametric Glioma MRI. *Scientific Reports* 2025, 15 (1), 32857. <https://doi.org/10.1038/s41598-025-17993-0>. PubMed PMID: 40998920
 18. Hassan, A. M.; Rajesh, A.; Asaad, M.; *et al.* Artificial Intelligence and Machine Learning in Prediction of Surgical Complications: Current State, Applications, and Implications. *The American Surgeon* 2023, 89 (1), 25–30. <https://doi.org/10.1177/00031348221101488>
 19. van Niftrik, C. H. B.; van der Wouden, F.; Staartjes, V. E.; *et al.* Machine Learning Algorithm Identifies Patients at High Risk for Early Complications after Intracranial Tumor Surgery: Registry-Based Cohort Study. *Neurosurgery* 2019, 85 (4), E756–E764. <https://doi.org/10.1093/neuros/nyz145>. PubMed PMID: 31149726
 20. Farrokhi, F.; Buchlak, Q. D.; Sikora, M.; *et al.* Investigating Risk Factors and Predicting Complications in Deep Brain Stimulation Surgery with Machine Learning Algorithms. *World Neurosurgery* 2020, 134, e325–e338. <https://doi.org/10.1016/j.wneu.2019.10.063>. PubMed PMID: 31634625
 21. Bonney, P. A.; Boettcher, L. B.; Burks, J. D.; *et al.* Review of Imaging Modalities for Monitoring Tumor Recurrence after Resection of Malignant Gliomas. *World Neurosurgery* 2017, 98, 686–696. <https://doi.org/10.1016/j.wneu.2016.12.022>
 22. Buchlak, Q. D.; Esmaili, N.; Leveque, J.-C.; *et al.* Machine Learning Applications to Clinical Decision Support in Neurosurgery: An Artificial Intelligence Augmented Systematic Review. *Neurosurgical Review* 2020, 43 (5), 1235–1253. <https://doi.org/10.1007/s10143-019-01163-8>
 23. Wireko, A. A.; Adebuseye, F. T.; Wellington, J.; *et al.* Recent Outcomes and Challenges of Artificial Intelligence, Machine Learning and Deep Learning Applications in Neurosurgery. *World Neurosurgery: X* 2024, 100301. <https://doi.org/10.1016/j.wnsx.2024.100301>

■ Author

Arman Raoufian is a freshman at Northwest High School in Maryland with a love for neurosurgery and machine learning. He is particularly focused on developing a deeper understanding of how computational methods can support clinical decision-making and improve patient outcomes. He hopes to major in a medical field.