

# Machine Learning in Neurosurgery: Toward Complex Inputs, Actionable Predictions, and Generalizable Translations

Received 08/27/2023

Review began 09/15/2023

Review ended 12/20/2023

Published 01/09/2024

© Copyright 2024

Schonfeld et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Ethan Schonfeld <sup>1</sup>, Nicole Mordekai <sup>2</sup>, Alex Berg <sup>1</sup>, Thomas Johnstone <sup>1</sup>, Aaryan Shah <sup>3</sup>, Vaibhavi Shah <sup>1</sup>, Ghani Haider <sup>1</sup>, Neelan J. Marianayagam <sup>1</sup>, Anand Veeravagu <sup>1</sup>

1. Neurosurgery, Stanford University School of Medicine, Stanford, USA 2. Medicine, Tel Aviv University, Tel Aviv, ISR  
3. School of Humanities and Sciences, Stanford University, Stanford, USA

**Corresponding author:** Ethan Schonfeld, ethan.schonfeld@stanford.edu

## Abstract

Machine learning can predict neurosurgical diagnosis and outcomes, power imaging analysis, and perform robotic navigation and tumor labeling. State-of-the-art models can reconstruct and generate images, predict surgical events from video, and assist in intraoperative decision-making. In this review, we will detail the neurosurgical applications of machine learning, ranging from simple to advanced models, and their potential to transform patient care. As machine learning techniques, outputs, and methods become increasingly complex, their performance is often more impactful yet increasingly difficult to evaluate. We aim to introduce these advancements to the neurosurgical audience while suggesting major potential roadblocks to their safe and effective translation. Unlike the previous generation of machine learning in neurosurgery, the safe translation of recent advancements will be contingent on neurosurgeons' involvement in model development and validation.

**Categories:** Neurosurgery, Healthcare Technology

**Keywords:** ai and robotics in healthcare, ai and machine learning, outcome prediction, generative adversarial network (gan), generative ai, federated learning, spine, neurosurgery

## Introduction And Background

Artificial intelligence (AI) utilizes computer systems to simulate human cognitive abilities. Machine learning (ML), a domain of AI, enables algorithms to recognize patterns in large, complex datasets to produce predictive outputs at inference time. Biomedical data has become vast and complex, thereby requiring AI to identify clinically meaningful relationships [1]. Hospitals produce large quantities of unstructured data from monitoring devices, imaging, and patient notes, requiring novel methods to help physicians effectively analyze this "big data" to identify predictive relationships. Such novel ML methods applied to certain tasks have outperformed not only classical statistical models but also clinical experts [1].

ML techniques have aided all neurosurgical domains, including early diagnosis, clinical decision-making, patient management, drug discovery, and prognosis prediction. In the case of outcome prediction, there has been an average accuracy of 0.945 and an average area under the curve (AUC) of 0.83, spanning across predictions for functional, spine, neurovascular, brain tumor, and traumatic subdomains [1]. This average performance of ML models was found to be significantly better than logistic regression models (median absolute performance increase of 15% accuracy) [1]. While simple ML models like logistic regression dominate most published applications to neurosurgery, this performance achievement of ML resulted from more complex models [1]. Such complex models (e.g., deep neural networks (DNNs), convolutional neural networks (CNNs)) have allowed for the utilization of complex input types including imaging and real-time surgical video, and thereby the prediction of complex outputs including non-radiographic intraoperative measurements of Cobb angle [2], cerebral artery segmentation in operative field of view [3], and augmented reality guidance for catheter placement for external ventricular drains [4].

With the attention toward AI/ML in neurosurgery significantly increasing, the National Institutes of Health funded 535 projects applying ML to clinical research, representing \$264 million in 2017 [5]. Attitudes in neurosurgery were open to the application of AI in neurosurgery for imaging interpretation, operative planning, autonomous surgery, and hazard/complication prediction [6]. Despite funding, open attitudes, and substantial research efforts, few technological applications have been integrated into patient care. While many algorithms have achieved high predictive power for outcome predictions, few have been externally validated across multiple sites. This is worrisome because unlike fields of autonomous driving and search engine development in which data diversity is high, neurosurgical data used to train ML algorithms is generally from single institutions.

In this comprehensive narrative review, our objective is to acquaint clinical practitioners, who will play a pivotal role in the safe and efficient integration of this technology, with the nuances of ML in the context of neurosurgery. We explore the diverse algorithms and their applications in this field, while also addressing

### How to cite this article

Schonfeld E, Mordekai N, Berg A, et al. (January 09, 2024) Machine Learning in Neurosurgery: Toward Complex Inputs, Actionable Predictions, and Generalizable Translations. Cureus 16(1): e51963. DOI 10.7759/cureus.51963

the technical challenges that may impede their effective implementation in enhancing patient care. A particular emphasis is placed on the critical aspects of data sharing and the necessity of training models using diverse data sets. We discuss these elements in depth, offering both technical insights and clinical perspectives, to underscore their importance in ensuring the safe and successful adoption of ML in neurosurgery.

Definitions of technologies

The AI technologies in this review are described in Table 1.

Technology	Definition
Artificial intelligence (AI)	Simulation of problem-solving and logical thinking by computer systems to perform tasks.
Machine learning (ML)	Subfield of AI that learns patterns in large data without explicit instruction that are often used to make predictions.
Deep learning (DL)	Subfield of ML that uses ANNs to mimic the learning process of the human brain.
Artificial neural networks (ANN)	Systems of artificial neurons trained to learn complex interactions within input data to optimize a cost/objective function.
Supervised learning	Subfield of ML that trains models using labeled datasets to predict labels for similar data.
Unsupervised learning	Subfield of ML that learns patterns in unlabeled data that can be used to label, classify, and represent the data.

**TABLE 1: Technology definitions**

AI: Artificial intelligence; ML: Machine learning; ANN: Artificial neural networks

Review

Input variable types and data sources

With many ML model types and potential inputs, careful attention to input data type should help determine model selection during development. Table 2 covers input data types that are compatible with the different ML models. For the training of such models, Table 3 details publicly available datasets, focusing on those most frequently used and those serving as examples of emerging dataset releases with large potential impacts.

	Input variables	Example use case
Logistic regression	Features (i.e., clinical features)	Prediction of outcomes following aneurysmal subarachnoid hemorrhage using clinical information, neuroimaging features (i.e., clot size), and treatment modality information (Feghali et al., 2022) [7]
Support vector machine (SVM)	Features (i.e., clinical features)	Differentiation of benign and malignant pediatric brain tumors using blood markers (Khayat Kashani et al., 2022) [8]
Artificial neural networks (ANN)	Features (i.e., clinical features)	Prediction of traumatic brain injury outcomes in low-resource settings from easily acquired clinical variables (Adil et al., 2022) [9]
Convolutional neural network (CNN)	Imaging, video, genomics, time series	Intraoperative microscopic tumor detection from label-free stimulated Raman scattering microscopy (Reinecke et al., 2022) [10]
Long short-term memory network (LSTM)	Free text, genomics, time series	Differentiating lumbar disc herniation and lumbar spinal stenosis using free text admission notes (Ren et al., 2022) [11]
Transformer	Free text, imaging, genomics, time series	Classification and histopathological feature labeling of primary brain tumors (Li et al., 2023) [12]
Generative adversarial network (GAN)	Imaging	Generating CT from MRI data (Ranjan et al., 2021) [13]
Foundational model	Broad data	Adding realistic-looking abnormalities to synthetically generated radiology images (Chambon et al., 2022) [14]

TABLE 2: Machine learning method types and input variables

Dataset	Data collection period	Included labels	Number of subjects
Truven Health Analytics MarketScan research database [15]	2010-2018	Demographic, admission, diagnosis codes, discharge status, procedure codes, length of stay, financial information, drug information	~273 million
National (Nationwide) Inpatient Sample (NIS) discharge database [16]	2012-2022	Demographic, admission, diagnosis codes, procedure codes, financial information, disease severity, hospital information, length of stay	~35 million
MIMIC-III [17]	2008-2014	Demographics, clinical measurement, intervention, billing, pharmacotherapy, clinical laboratory test, medical data	~38,000 distinct adult patients ~53,000 admissions
VerSe [18,19]	2013-2017	CT imaging, annotated vertebrae (segmentation mask, fracture grading)	374 CT scans, 355 patients
OpenSRH [20]		Clinical stimulated Raman histology images, full pathologic annotations, whole slide tumor segmentations, raw and processed optical imaging data	~300 patients ~1300 whole slide optical images
Parkinson's Progression Marker Initiative (PPMI) [21]	2011-ongoing	Demographics, admission, diagnosis codes clinical, imaging, omics, genetic, sensor, and biomarker data	~ 1758 patients
American College of Surgeons National Surgical Quality Improvement Program (NSQIP-P) [22]	2005-ongoing	Demographic variables, preoperative risk factors, intraoperative variables, and 30-day postoperative mortality	~10 million
The Quality Outcomes Database (QOD) [23]	2012-2019	Demographic, admission, diagnosis codes, discharge status, variables specific to lumbar and cervical pathologies	~120,000
eICU Collaborative Research Database [24]	2014-2015	Demographics, vital sign measurements, care plan documentation, severity of illness measures, diagnosis information, treatment information, admission codes	~200,000
The Surveillance, Epidemiology, and End Results Program (SEER) [25]	2002-Ongoing	Patient demographics, primary tumor site, tumor morphology and stage at diagnosis, first course of treatment, and follow-up for vital status	~10 million
University of Pennsylvania Glioblastoma Imaging, Genomics, and Radiomics (UPenn-GBM) [26]	2006-2018	Clinical, demographic, and molecular information MRI, perfusion, and diffusion derivative volumes, computationally derived and manually revised expert annotations of tumor sub-regions, as well as quantitative imaging features corresponding to each of these region	630 patients
Clinformatics™ DataMart [27]	2007-2019	Clinical and demographic information, Date of service, place of service, ICD-9-CM codes, CPT codes, provider type, drug quantity dispensed, days supplied, charges, deductibles, and copayments	~100 million

**TABLE 3: Selected publicly available databases and datasets used to develop machine learning models for neurosurgery and medical applications**

MIMIC-III: Medical Information Mart for Intensive Care-III; VerSe: Large Scale Vertebrae Segmentation Challenge; SRH: Stimulated Raman histology; ICD: International Classification of Diseases; CM: Clinical modification; CPT: Current Procedural Terminology; eICU: Electronic intensive care unit

Outcome prediction

Postoperative Complications

ML modeling has been applicable in predicting neurosurgical postoperative complications. van Niftrik et al.

utilized gradient-boosting ML modeling to predict patients at risk for early complications (AUC = 0.73) following intracranial tumor surgery [28]. Neural networks identified patients with a high risk for cerebrospinal fluid (CSF) leaks following pituitary surgery using clinical and surgical features as input variables, achieving high discrimination (AUC = 0.84) [29]. In spine surgery, 90-day complication following anterior cervical discectomy and fusion has been predicted by DNN (AUC = 0.832) [30], and for both posterior lumbar spinal fusion and adult spinal deformity surgery, specific complications were predicted by artificial neural network (ANN) [31,32].

Quality of Life

ML modeling has been used to predict quality of life based on assessment scores, including the Oswestry Disability Index (ODI) and the Glasgow Outcome Score (GOS). Using ANNs, predictions of GOS, which assess the long-term necessity for rehabilitation, in patients with aneurysmal subarachnoid hemorrhage achieved AUC of 0.85 and 0.96 based on clinical and angiogram data respectively [33]. Azimi et al. used ANNs to predict postsurgical satisfaction over two years for patients with lumbar spinal canal stenosis undergoing lumbar spine surgery [34]. Support vector machine (SVM) predicted postoperative ODI prior to operation for cervical spondylotic myelopathy using preoperative ODI and symptom duration with a coefficient of determination of 0.932 [35].

Recurrence and Mortality

Tumor recurrence has been predicted effectively by ML models. For example, early progression and recurrence outcomes for patients with parasagittal and parafalcine meningiomas were successfully predicted by SVM and random forest (RF), achieving AUC = 0.91 [36]. SVMs were used to predict survival for glioma, achieving similar or higher AUC in comparison to a team of neuroradiologists [37].

Translation of ML in Neurosurgery

A comprehensive analysis of translation efforts in ML applied to neurosurgery is detailed in Table 4.

Study/device	Externally validated	Applied for FDA approval	Translated into clinic	Type of model	Training method	Input	Outcome	Model performance
Valliani et al. [38]	Yes	No	No	Ensemble learning algorithm	Single center	Clinical features and surgical variables	Non-home discharge after thoracolumbar spine surgery	AUC: 0.81 (test) AUC: 0.77 (external)
Senders et al. [39]	Yes	No	No	Accelerated failure time model	Retrospective dataset	Demographic, socioeconomic, clinical, and radiographic features	Prediction of survival in glioblastoma patients	C-index: 0.70 (best time-to-event model), 0.70 (best continuous and binary model)
Liu et al. [40]	Yes	No	No	Decision tree	Retrospective	Demographic, clinical, and aneurysm-specific features as well as Glasgow Coma Score	Long-term outcomes after poor-grade aneurysmal subarachnoid hemorrhage	AUC: 0.88 (test) AUC: 0.94 (external)
Staatjes et al. [41]	Yes	No	No	Logistic generalized additive model	Multicenter retrospective database	Clinical features and labs	Functional impairment after intracranial tumor surgery	AUC: 0.72 (test), AUC: 0.72 (external)
Viz Aneurysm [42]	Yes	Yes	Yes	Neural network	Unknown	CTA imaging	Detection and analysis of cerebral aneurysms  Doses of harmful	93.8% sensitivity 94.2% specificity (external)

Carra et al. [43]	Yes	No	No	Random forest	Prospective and retrospective databases	ICP and MAP signals	intracranial pressure in patients with severe traumatic brain injury	AUC: 0.94 (external)
Aidoc BriefCase-CSF triage [44]	Yes	Yes	Yes	Neural network	Unknown	CTA images	Detection of cervical spine fractures	(external) Sensitivity - 54.9 Specificity - 94.1
Rapid LVO 1.0 [45]	Yes	Yes	Yes	Density threshold-based traditional machine learning	Unknown	CTA images	Detection of anterior circulation LVOs	Sensitivity - 96.0 Specificity - 98.0
Imaging biometrics (IB) Neuro-Oncology [46]	Yes	Yes	Yes	Artificial Intelligence	Unknown	CTA images	Distinguishing glioblastoma progression	Classification of survival in GBM sensitivity - 80.0 Specificity - 63.0
Staartjes et al. [47]	Yes	No	No	Logistic regression	Retrospective multicenter database	Clinical features and surgical variables	Prediction of mid-term outcomes after lumbar spinal fusion	Functional Impairment AUC:0.75 (test) AUC:0.67 (external) Back pain AUC: 0.71 (test) AUC:0.72 (external) Leg Pain AUC: 0.72 (test) AUC: 0.64 (external)
Thanellas et al. [48]	Yes	No	No	Convolutional neural network	Retrospective single-center dataset	CTA images	Identification and localization of subarachnoid hemorrhage on CT scans	Development: Sensitivity 87.4% and specificity 95.3% External: sensitivity 99.3% and specificity 63.2%
Teng et al. [49]	Yes	No	No	Logistic regression	Retrospective multicenter database	Clinical features lab values and imaging	Predicting high-grade intracranial meningioma	Development: AUC: 0.99 AUC: 0.75 (internal) AUC: 0.842 (external)
Vitrea CT Brain Perfusion [50]	Yes	Yes	Yes	Convolutional neural network	Unknown	CTA images	Detection of intracranial aneurysms	Sensitivity - 90.9-96.3 Specificity - 100.0
Ma et al. [51]	Yes	No	No	Neural network	Retrospective multicenter database	MRI images	Prediction of intramedullary glioma grade	AUC: 0.8431
Biswas et al. [52]	Yes	No	No	Neural network	Retrospective single-center database	Predictor variables (demographic and lab)	Predicting chronic subdural hematoma referral outcome	AUC: 0.951 (test) AUC: 0.896 (external)
Fang et al. [53]	Yes	No	No	Convolutional neural network	Retrospective single-center database	Demographic, clinical, and imaging variables	Glasgow Outcome Scale	AUC (test): 0.93 AUC (external): 93.69
Karhade et al. [54]	Yes	No	No	Logistic regression	Retrospective multicenter database	Clinical labs and tumor-specific variables	Six-week mortality in spinal metastasis	AUC (test): 0.84 AUC (external): 0.81

							cases	
Crabb et al. [55]	Yes	No	No	Logistic regression, random forest, support vector machine, and ensemble learning algorithm	Public database	Age, BMI, clinical variables, and preoperative lab tests	Prediction of unplanned 30-day readmissions after pituitary adenoma resection	AUC (external): 0.76
Warman et al. [56]	Yes	No	No	Ensemble learning algorithm	Prospective single center dataset and public dataset	Patient demographics, presenting vital signs, mechanism of injury, initial Glasgow Coma Scale (GCS)	Predicting in-hospital mortality after traumatic brain injury	AUC (test): 0.91 AUC (external): 0.89
Habets et al. [57]	Yes	No	No	Logistic regression	Retrospective single center	Patient demographic data, disease-specific data, clinical performance scores, and neuropsychological scores	Prediction of motor response after deep brain stimulation	AUC (test): 0.79 AUC (external) 0.79

TABLE 4: Analysis of translation efforts

AUC: Area under the curve; CTA: Computed tomography angiography; ICP: Intracranial pressure; MAP: Mean arterial pressure; CSF: Cerebrospinal fluid; LVO: Large vessel occlusion; IB: Imaging biometrics; GBM: Glioblastoma; CT: Computerized tomography; MRI: Magnetic resonance imaging; BMI: Body mass index

Intraoperative and robotic applications

ML approaches underpin many of the robotic functions that currently drive improvements in patient outcomes in neurosurgery. With these techniques, neurosurgeons retain accuracy over the course of long, technically complex operations, where minuscule reductions in operative times, patient length of stay, revision surgeries, and morbidity translate to large amounts of value added [58]. Spinal robots have improved pedicle screw placement accuracy (odds ratio (OR) 0.44 compared to freehand; OR 0.50 compared to CT navigation) and optimal screw placement [59,60], while intracranial automation has increased the speed and accuracy that leads to deep brain stimulation and stereoelectroencephalography are placed [61]. The exciting capabilities of neurosurgical robotics augmented with ML are image registration for intraoperative neuronavigation and surgical task automation, enhancing operative planning, efficiency, and accuracy.

Neuronavigation

Neuronavigation, the application of technology to localize lesions within the skull or vertebral column, has drastically impacted workflow within various neurosurgical subdisciplines, driving improvements in outcomes from epilepsy surgery to spinal fusion. AI systems have served important roles in traditional neuro-navigation systems, using infrared, electromagnetism, and ultrasonography methods to track position. Currently, novel ML methods are being developed to enhance these traditional systems. A key challenge in intraoperative neuronavigation is high-fidelity image registration, or the action of updating preoperative imaging to reflect real-time changes in the surgical field, while creating navigation plans. For example, resection of a large tumor can significantly alter the patient’s anatomy, such that the preoperative imaging used to create the navigation plan is no longer accurate [62]. Intra-operative registration updates the navigation plan to reflect the amount of tumor removed and the shift of normal tissue toward the resection cavity. Several models have been proposed to optimize registration; Han et. al. have developed an unsupervised, dual-channel deep learning (DL) network that updates preoperative magnetic resonance imaging (MRI) with intraoperative computerized tomography (CT) [63]. Such registration techniques can improve current robot navigation systems, enhancing the speed and accuracy of neurosurgical operations.

Subtask Automation

While neurosurgical procedures require a skilled surgeon, some tasks may be automated, enabling attention to critical areas. During an awake craniotomy, a neurosurgical robot could automatically stimulate a critical area of the eloquent cortex, sensing when the operator has performed the delicate part of the operation. Padoy and Hager have demonstrated such surgeon-robot interactions with Hidden Markov models; they enabled a da Vinci surgical system (Intuitive Surgical, Sunnyvale, United States) to recognize when a surgeon had inserted a suture needle, automatically pull the needle through, and seamlessly transfer control back to the surgeon for the next bite [64]. Similarly, Hu et al. indicated that the RAVEN II Surgical Robot could semi-autonomously perform brain tumor ablation via modeling and implementation of a behavior tree framework [65]. Robotics augmented with ML algorithms can further automate surgical trajectory planning as demonstrated by the minimally invasive RAVEN II Surgical Robot that can plan efficient trajectories to approach and resect residual brain tumor at its margins following debulking [66]. Regarding a stereotactic brain biopsy, neurosurgeons using computer-assisted route planning software were able to create shorter trajectories that were more orthogonal to the skull and carried a lower risk of injuring vasculature as compared to manually planned routes [67].

#### *Non-robotic Intra-operative Applications*

ML has non-robotic intraoperative uses that may improve neurosurgical practice. Wong et al. developed an unsupervised ML model that used microelectrode recording inputs to functionally localize and visualize the subthalamic nucleus (STN) during deep brain stimulation (DBS) procedures [68]. Valsky et al. reported their successfully trained Hidden Markov model that identifies when a DBS lead improperly exited the ventral STN [69]. It is also possible to predict postoperative motor improvement in Parkinson's disease following DBS by applying an RF to intraoperative microelectrode recording data [70]. Beyond functional neurosurgery, Jermyn et al. showed that an RF classifier could be applied to intraoperative Raman spectroscopy data during grade 2-4 glioma resection, identifying more invasive cells with an accuracy of 92% while a surgeon with a bright-field microscope and MRI could identify the same cancer cells with 73% accuracy [71]. More recently, intraoperative high-resolution magic angle spinning nuclear magnetic resonance was used in conjunction with an RF to differentiate tumor cells from healthy controls with a median AUC of 85.6% and area under the precision-recall curve (AUPR) of 93.4, while also being able to distinguish between benign and malignant samples with a median AUC of 87.1% and AUPR of 96.1% [72]. Ritschel et al. showed that SVMs trained on contrast-enhanced ultrasound image data can accurately detect the resection margins of glioblastoma [73]. Li et al. developed an approach where a CNN with a long short-term memory architecture was applied to probe-based confocal laser endomicroscopy images to successfully differentiate between glioblastoma and meningioma [74].

Overall, clinically useful applications of ML to neurosurgical robotics and intraoperative decisions are steadily increasing. While ML has benefited domains like DBS lead implantation, neuronavigation, pedicle screw placement, and tumor boundary classification, further innovation is on the horizon.

### **Diagnostic and imaging applications of ML in neurosurgery**

ML can assist in the diagnosis and classification of major neurosurgical conditions. For epilepsy detection, CNNs have achieved performances of >96% accuracy using electroencephalography (EEG) data [75], extending to real-time seizure prediction [76]. Using CT angiography images for intracranial aneurysm detection, CNNs achieved an accuracy of 0.886 on independent internal validation. Aneurysms in the anterior communicating artery, anterior cerebral artery, vertebrobasilar artery, and middle cerebral artery were best predicted, with poor performance for tiny aneurysms (<3 mm) [77]. For the spine, ML has been used to diagnose fracture from CT [78] (accuracy 0.932), to detect foraminal stenosis using generative adversarial networks (GANs) for semantic segmentation [79] (mean average precision 0.837), and to diagnose osteoarthritis by DL from medical record data without any imaging [80] (accuracy 0.768). The use of ML for neurosurgical diagnostics is focused on tumor classification both pre and perioperatively. Such ML applications have predicted meningioma grade [81] (AUC 0.8895), non-invasive tissue classification by SVM from Hyperspectral Imaging [82], intraoperative tumor classification by DNA methylation profile [83] (accuracy 89%), and preoperative tumor classification by CNN from MRI-only data [84] (accuracy 0.9656).

Medical data including medical records, imaging (CT, MRI, EEG, hyperspectral imaging), and intraoperative tissue, serve as inputs to ML models for the diagnosis of neurosurgical conditions. However, ML and feature selection algorithms allow for radiomics: the discovery and development of clinically explainable radiologic features that are useful for prediction. For traumatic brain injury prediction and lesion classification, features including shape, intensity, and texture biomarkers were isolated and used independently, demonstrating the validity of these as radio markers for lesion classification [85]. For Gamma Knife radiosurgery, ML was used to discover features to predict surgical effectiveness, finding higher zone percentage as a radiomarker for achieving local tumor control in brain metastasis [86]. Feature-based ANNs were used to discover radio-markers for preoperative tumor differentiation between skull base chordoma and chondrosarcoma for preoperative planning, extracting the potential radiomarkers from preoperative MRI imaging, and by using only seven of these identified radiomarkers achieved an AUC of 0.93 [87].

### **State-of-the-art applications: GANs and federated learning**

### *Generative Adversarial Networks*

The applications of ML to diagnostics in neurosurgery extend beyond disease prediction, innovating the imaging methods themselves. Using GANs, MRI conversion to CT has been demonstrated in the brain [88-90] and spine [91]. A GAN generates images by training two DL models at the same time: one to generate synthetic images and another to discriminate between real and synthetic images, both competing with each other to simultaneously improve. Synthetic CT scans have demonstrated the ability to validate the downstream application of dosimetric determination for stereotactic brain radiotherapy [92], allowing for the potential conversion of MRI/CT workflow to MRI-only workflows for diagnostics and radiosurgical planning. GANs applied to positron emission tomography imaging improved spatial resolution [93] and dose reduction [94]. Lastly, GANs have allowed for the consideration of multiple imaging modalities, improving Alzheimer's disease diagnosis by synthetically constructing missing imaging information and using a downstream CNN to combine the real and synthetic information to predict the diagnosis [95] (AUC > 0.87).

### *Federated Learning*

A fundamental barrier in ML applications to neurosurgery is the lack of available "big data," mainly due to the cost of storage and the preservation of patient privacy. This preservation prohibits the sharing of medical data across institutions, restricting the ability to build large datasets necessary to train DL models and restricting algorithms to learn from data at one institution [96]. This "domain shift" in turn results in internally validated algorithms failing to generalize and externally validate [97]. One solution, synthetic learning, is to share synthetic data that represents the clinical information of patient data but preserves its privacy; efforts have begun to use GANs to generate synthetic data and to allow sharing across institutions. This approach has been proven to be unbiased and downstream models trained on the synthetic data have achieved high performance on brain tumors [98], nuclei segmentation [98], and spine radiograph abnormality classification [99]. The current most popular method to preserve data privacy while allowing for the training of complex networks across institutions is termed federated learning (FL). In FL, private data never leaves an institution. Instead, a basic ML model is trained on the institution's data, and statistics from the updated model are sent back to a central system. The aggregation of these results across many institutions may send an updated model that better represents patient populations, fine-tuned by the specific institutions before application. Such FL schemes in neurosurgery have demonstrated success, outperforming locally trained models, in intracerebral hemorrhage detection on CT scans [100] (AUC 0.9487) and automatic tumor boundary detection for glioblastoma [101]. The main restriction of FL is in rare data environments, where the initial single institution models cannot learn meaningful information. Future work must develop technical solutions to this problem as many applications of ML in neurosurgery have limited patient volume.

### *Large Language Models*

ChatGPT, a large language model (LLM) trained by OpenAI, has demonstrated the ability to pass neurosurgical boards. Similarly trained LLMs, such as NYUTRON, have been shown to achieve high performance for common neurosurgical outcome prediction tasks such as readmission prediction. At this cutting edge of ML in neurosurgery, motivated by the increasingly "black box" construction of these models and their training, it becomes essential to investigate training strategy, bias analysis, and hallucinations of the models (Table 5).

Paper/model	Input variables	Training data source	Performance measurement	Hallucinations mentioned	Common hallucinations	Translated	Bias investigated
NYUTRON [102,103]	Text in clinical records	NYU EHR records	Mortality prediction, Comorbidity index, readmission prediction	No	No	No	Age, race
GPT-4 ChatGPT and Bard [104]	The 149-question self-assessment neurosurgery examination indications examination	Both supervised and unsupervised learning techniques on a large corpus of Internet text data	Score on neurosurgical examination	Yes	Incorrect image Analysis in test questions	No	N/A

TABLE 5: Large language models in neurosurgery

NYU: New York University; EHR: Electronic health records; GPT: Generative pre-trained transformer

Conclusions

While ML is already impacting patient care, most of the developments discussed in this review have not been translated at a large scale. A primary issue is that these algorithms lack generalizability. While efforts toward standardization in AI research reporting are commendable and have led to significant improvements, the unique challenges in neurosurgery, such as differences in patient cohorts, pathology and outcome definitions, and imaging modality generation differences, highlight the need for tailored approaches to standardization in this field. Data sharing and technical solutions including FL should help transition to "big data" and with that, the increase of ML assistance in direct patient care in neurosurgery. As both algorithms and cases become more complex, incorporating inputs of imaging, video, and genomics, it is essential that neurosurgeons consider the technical elements of the proposed methods. As large language models and ChatGPT begin to be translated to neurosurgery, the translation of these models should be clinician-led and necessitate clinical metrics to carefully validate these tools while developing them to improve the quality of care.

Additional Information

Disclosures

**Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** Anand Veeravagu declare(s) personal fees from Medtronic, Stryker, Nuvasive, Surgical Theater, Osteocentric, Higgs Boson. Anand Veeravagu declare(s) stock/stock options from ATEC. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

References

- Senders JT, Staples PC, Karhade AV, et al.: Machine learning and neurosurgical outcome prediction: a systematic review. *World Neurosurg.* 2018, 109:476-486.e1. [10.1016/j.wneu.2017.09.149](#)
- Gami P, Qiu K, Kannappan S, et al.: Semiautomated intraoperative measurement of Cobb angle and coronal C7 plumb line using deep learning and computer vision for scoliosis correction: a feasibility study. *J Neurosurg Spine.* 2022, 1-9. [10.3171/2022.4.SPINE22133](#)
- Kim MS, Cha JH, Lee S, Han L, Park W, Ahn JS, Park SC: Deep-learning-based cerebral artery semantic segmentation in neurosurgical operating microscope vision using indocyanine green fluorescence videoangiography. *Front Neurobot.* 2022, 15:735177. [10.3389/fnbot.2021.735177](#)
- Eom S, Sykes D, Rahimpour S, et al.: Deep-learning-based cerebral artery semantic segmentation in neurosurgical operating microscope vision using indocyanine green fluorescence videoangiography. *IEEE Int Sym Mixed Aug Real (ISMAR).* 2022, 355:64. [10.1109/ISMAR55827.2022.00051](#)
- Annapureddy AR, Angraal S, Caraballo C, Grimshaw A, Huang C, Mortazavi BJ, Krumholz HM: The National Institutes of Health funding for clinical research applying machine learning techniques in 2017. *NPJ Digit Med.* 2020, 3:13. [10.1038/s41746-020-0223-9](#)
- Layard Horsfall H, Palmisciano P, Khan DZ, Muirhead W, Koh CH, Stoyanov D, Marcus HJ: Attitudes of the surgical team toward artificial intelligence in neurosurgery: international 2-stage cross-sectional survey. *World Neurosurg.* 2021, 146:e724-30. [10.1016/j.wneu.2020.10.171](#)
- Feghali J, Sattari SA, Wicks EE, et al.: External validation of a neural network model in aneurysmal subarachnoid hemorrhage: a comparison with conventional logistic regression models. *Neurosurgery.* 2022,

- 90:552-61. [10.1227/neu.0000000000001857](https://doi.org/10.1227/neu.0000000000001857)
8. Khayat Kashani HR, Azhari S, Moradi E, Samii F, Mirahmadi MS, Towfiqi A: Predictive value of blood markers in pediatric brain tumors using machine learning. *Pediatr Neurosurg*. 2022, 57:323-32. [10.1159/000525862](https://doi.org/10.1159/000525862)
  9. Adil SM, Elahi C, Patel DN, et al.: Deep learning to predict traumatic brain injury outcomes in the low-resource setting. *World Neurosurg*. 2022, 164:e8-e16. [10.1016/j.wneu.2022.02.097](https://doi.org/10.1016/j.wneu.2022.02.097)
  10. Reinecke D, von Spreckelsen N, Mawrin C, et al.: Novel rapid intraoperative qualitative tumor detection by a residual convolutional neural network using label-free stimulated Raman scattering microscopy. *Acta Neuropathol Commun*. 2022, 10:109. [10.1186/s40478-022-01411-x](https://doi.org/10.1186/s40478-022-01411-x)
  11. Ren G, Yu K, Xie Z, et al.: Differentiation of lumbar disc herniation and lumbar spinal stenosis using natural language processing-based machine learning based on positive symptoms. *Neurosurg Focus*. 2022, 52:E7. [10.3171/2022.1.FOCUS21561](https://doi.org/10.3171/2022.1.FOCUS21561)
  12. Li Z, Cong Y, Chen X, et al.: Vision transformer-based weakly supervised histopathological image analysis of primary brain tumors. *iScience*. 2023, 26:105872. [10.1016/j.isci.2022.105872](https://doi.org/10.1016/j.isci.2022.105872)
  13. Ranjan A, Lalwani D, Misra R: GAN for synthesizing CT from T2-weighted MRI data towards MR-guided radiation treatment. *MAGMA*. 2022, 35:449-57. [10.1007/s10334-021-00974-5](https://doi.org/10.1007/s10334-021-00974-5)
  14. Chambon P, Bluethgen C, Langlotz CP, et al.: Adapting pretrained vision-language foundational models to medical imaging domains. *arXiv*. 2022, [10.48550/arXiv.2210.04133](https://arxiv.org/abs/2210.04133)
  15. Real world evidence | Merative. (2023). Accessed: August 16, 2023: <https://www.merative.com/real-world-evidence>.
  16. HCUP-US NIS overview. (2022). Accessed: August 16, 2023: <https://hcup-us.ahrq.gov/nisoverview.jsp>.
  17. Johnson AE, Pollard TJ, Shen L, et al.: MIMIC-III, a freely accessible critical care database. *Sci Data*. 2016, 3:160035. [10.1038/sdata.2016.35](https://doi.org/10.1038/sdata.2016.35)
  18. Löffler MT, Sekuboyina A, Jacob A, et al.: A vertebral segmentation dataset with fracture grading. *Radiol Artif Intell*. 2020, 2:e190138. [10.1148/ryai.2020190138](https://doi.org/10.1148/ryai.2020190138)
  19. Sekuboyina A, Hussein ME, Bayat A, et al.: Verse: a vertebrae labelling and segmentation benchmark for multi-detector CT images. *Med Image Anal*. 2021, 73:102166. [10.1016/j.media.2021.102166](https://doi.org/10.1016/j.media.2021.102166)
  20. Jiang C, Chowdury A, Hou X, et al.: OpenSRH: optimizing brain tumor surgery using intraoperative stimulated Raman histology. *Adv Neural Inf Process Syst*. 2022, 35:28502-16. [10.48550/arXiv.2206.08439](https://arxiv.org/abs/2206.08439)
  21. The Parkinson progression marker initiative (PPMI). *Prog Neurobiol*. 2011, 95:629-35. [10.1016/j.pneurobio.2011.09.005](https://doi.org/10.1016/j.pneurobio.2011.09.005)
  22. ACS National Surgical Quality Improvement Program. (2023). Accessed: August 16, 2023: <https://www.facs.org/quality-programs/data-and-registries/acs-nsqip/>.
  23. Quality Outcomes Database (QOD) - Annual Report. (2020). Accessed: August 16, 2023: <https://annualreports.aans.org/2019/npa/qod/quality-outcomes-database-qod/>.
  24. Pollard TJ, Johnson AE, Raffa JD, Celi LA, Mark RG, Badawi O: The eICU collaborative research database, a freely available multi-center database for critical care research. *Sci Data*. 2018, 5:180178. [10.1038/sdata.2018.178](https://doi.org/10.1038/sdata.2018.178)
  25. Surveillance, Epidemiology, and End Results Program. (2023). Accessed: August 16, 2023: <https://seer.cancer.gov/>.
  26. Bakas S, Sako C, Akbari H, et al.: The University of Pennsylvania glioblastoma (UPenn-GBM) cohort: advanced MRI, clinical, genomics, & radiomics. *Sci Data*. 2022, 9:453. [10.1038/s41597-022-01560-7](https://doi.org/10.1038/s41597-022-01560-7)
  27. Claims data | Optum. (2023). Accessed: August 16, 2023: <https://www.optum.com/business/life-sciences/real-world-data/claims-data.html>.
  28. van Niftrik CH, van der Wouden F, Staartjes VE, et al.: Machine learning algorithm identifies patients at high risk for early complications after intracranial tumor surgery: registry-based cohort study. *Neurosurgery*. 2019, 85:E756-64. [10.1093/neuros/nyz145](https://doi.org/10.1093/neuros/nyz145)
  29. Staartjes VE, Zattra CM, Akeret K, et al.: Neural network-based identification of patients at high risk for intraoperative cerebrospinal fluid leaks in endoscopic pituitary surgery. *J Neurosurg*. 2019, 133:329-35. [10.3171/2019.4.JNS19477](https://doi.org/10.3171/2019.4.JNS19477)
  30. Rodrigues AJ, Schonfeld E, Varshneya K, Stienen MN, Staartjes VE, Jin MC, Veeravagu A: Comparison of deep learning and classical machine learning algorithms to predict postoperative outcomes for anterior cervical discectomy and fusion procedures with state-of-the-art performance. *Spine (Phila Pa 1976)*. 2022, 47:1637-44. [10.1097/BRS.0000000000004481](https://doi.org/10.1097/BRS.0000000000004481)
  31. Kim JS, Arvind V, Oermann EK, et al.: Predicting surgical complications in patients undergoing elective adult spinal deformity procedures using machine learning. *Spine Deform*. 2018, 6:762-70. [10.1016/j.jspd.2018.03.003](https://doi.org/10.1016/j.jspd.2018.03.003)
  32. Kim JS, Merrill RK, Arvind V, et al.: Examining the ability of artificial neural networks machine learning models to accurately predict complications following posterior lumbar spine fusion. *Spine (Phila Pa 1976)*. 2018, 43:853-60. [10.1097/BRS.0000000000002442](https://doi.org/10.1097/BRS.0000000000002442)
  33. Lo BW, Macdonald RL, Baker A, Levine MA: Clinical outcome prediction in aneurysmal subarachnoid hemorrhage using Bayesian neural networks with fuzzy logic inferences. *Comput Math Methods Med*. 2013, 2013:904860. [10.1155/2013/904860](https://doi.org/10.1155/2013/904860)
  34. Azimi P, Benzel EC, Shahzadi S, Azhari S, Mohammadi HR: Use of artificial neural networks to predict surgical satisfaction in patients with lumbar spinal canal stenosis: clinical article. *J Neurosurg Spine*. 2014, 20:300-5. [10.3171/2013.12.SPINE13674](https://doi.org/10.3171/2013.12.SPINE13674)
  35. Hoffman H, Lee SI, Garst JH, et al.: Use of multivariate linear regression and support vector regression to predict functional outcome after surgery for cervical spondylotic myelopathy. *J Clin Neurosci*. 2015, 22:1444-9. [10.1016/j.jocn.2015.04.002](https://doi.org/10.1016/j.jocn.2015.04.002)
  36. Hsieh HP, Wu DY, Hung KC, Lim SW, Chen TY, Fan-Chiang Y, Ko CC: Machine learning for prediction of recurrence in parasagittal and Parafalcine meningiomas: combined clinical and MRI texture features. *J Pers Med*. 2022, 12:522. [10.3390/jpm12040522](https://doi.org/10.3390/jpm12040522)
  37. Emblem KE, Pinho MC, Zöllner FG, et al.: A generic support vector machine model for preoperative glioma survival associations. *Radiology*. 2015, 275:228-34. [10.1148/radiol.14140770](https://doi.org/10.1148/radiol.14140770)

38. Valliani AA, Kim NC, Martini ML, et al.: Robust prediction of non-Home discharge after thoracolumbar spine surgery with ensemble machine learning and validation on a nationwide cohort. *World Neurosurg.* 2022, 165:e83-91. [10.1016/j.wneu.2022.05.105](https://doi.org/10.1016/j.wneu.2022.05.105)
39. Senders JT, Staples P, Mehrdash A, et al.: An online calculator for the prediction of survival in glioblastoma patients using classical statistics and machine learning. *Neurosurgery.* 2020, 86:E184-92. [10.1093/neuros/nyz403](https://doi.org/10.1093/neuros/nyz403)
40. Liu J, Xiong Y, Zhong M, Yang Y, Guo X, Tan X, Zhao B: Predicting long-term outcomes after poor-grade aneurysmal subarachnoid hemorrhage using decision tree modeling. *Neurosurgery.* 2020, 87:523-9. [10.1093/neuros/nyaa052](https://doi.org/10.1093/neuros/nyaa052)
41. Staartjes VE, Broggi M, Zattra CM, et al.: Development and external validation of a clinical prediction model for functional impairment after intracranial tumor surgery. *J Neurosurg.* 2020, 134:1743-50. [10.3171/2020.4.JNS20643](https://doi.org/10.3171/2020.4.JNS20643)
42. Viz Aneurysm. (2023). Accessed: August 16, 2023: <https://www.viz.ai/aneurysm>.
43. Carra G, Güiza F, Piper I, Citerio G, Maas A, Depreitere B, Meyfroidt G: Development and external validation of a machine learning model for the early prediction of doses of harmful intracranial pressure in patients with severe traumatic brain injury. *J Neurotrauma.* 2023, 40:514-22. [10.1089/neu.2022.0251](https://doi.org/10.1089/neu.2022.0251)
44. Aidoc briefcase-CSF triage. (2023). Accessed: July 15, 2023: <https://www.aidoc.com/solutions/neuro/>.
45. Inc Ischemaview: rapid LVO: quickly identify LVOs and streamline patient care | RapidAI . (2023). Accessed: August 16, 2023: <https://www.rapidai.com/rapid-lvo...>
46. Neuro-Oncology Solutions - imaging biometrics . (2023). Accessed: August 16, 2023: <https://www.imagingbiometrics.com/neuro-oncology-solutions/...>
47. Staartjes VE, Stumpo V, Ricciardi L, et al.: FUSE-ML: development and external validation of a clinical prediction model for mid-term outcomes after lumbar spinal fusion for degenerative disease. *Eur Spine J.* 2022, 31:2629-38. [10.1007/s00586-022-07135-9](https://doi.org/10.1007/s00586-022-07135-9)
48. Thanellias A, Peura H, Lavinto M, et al.: Development and external validation of a deep learning algorithm to identify and localize subarachnoid hemorrhage on CT scans. *Neurology.* 2023, 100:e1257-66. [10.1212/WNL.000000000000201710](https://doi.org/10.1212/WNL.000000000000201710)
49. Teng H, Yang X, Liu Z, et al.: The performance of different machine learning algorithm and regression models in predicting high-grade intracranial meningioma. *Brain Sci.* 2023, 13:594. [10.3390/brainsci13040594](https://doi.org/10.3390/brainsci13040594)
50. Rava RA, Snyder KV, Mokin M, et al.: Assessment of a Bayesian vitrea CT perfusion analysis to predict final infarct and Penumbra volumes in patients with acute ischemic stroke: a comparison with rapid. *AJNR Am J Neuroradiol.* 2020, 41:206-12. [10.3174/ajnr.A6395](https://doi.org/10.3174/ajnr.A6395)
51. Ma C, Wang L, Song D, et al.: Multimodal-based machine learning strategy for accurate and non-invasive prediction of intramedullary glioma grade and mutation status of molecular markers: a retrospective study. *BMC Med.* 2023, 21:198. [10.1186/s12916-023-02898-4](https://doi.org/10.1186/s12916-023-02898-4)
52. Biswas S, MacArthur JI, Pandit A, et al.: Predicting neurosurgical referral outcomes in patients with chronic subdural hematomas using machine learning algorithms - a multi-center feasibility study. *Surg Neurol Int.* 2023, 14:22. [10.25259/SNI\\_1086\\_2022](https://doi.org/10.25259/SNI_1086_2022)
53. Fang C, Pan Y, Zhao L, Niu Z, Guo Q, Zhao B: A machine learning-based approach to predict prognosis and length of hospital stay in adults and children with traumatic brain injury: retrospective cohort study. *J Med Internet Res.* 2022, 24:e41819. [10.2196/41819](https://doi.org/10.2196/41819)
54. Karhade AV, Fenn B, Groot OQ, et al.: Development and external validation of predictive algorithms for six-week mortality in spinal metastasis using 4,304 patients from five institutions. *Spine J.* 2022, 22:2033-41. [10.1016/j.spinee.2022.07.089](https://doi.org/10.1016/j.spinee.2022.07.089)
55. Crabb BT, Hamrick F, Campbell JM, et al.: Machine learning-based analysis and prediction of unplanned 30-day readmissions after pituitary adenoma resection: a multi-institutional retrospective study with external validation. *Neurosurgery.* 2022, 91:263-71. [10.1227/neu.0000000000001967](https://doi.org/10.1227/neu.0000000000001967)
56. Warman PI, Seas A, Satyadev N, et al.: Machine learning for predicting in-hospital mortality after traumatic brain injury in both high-income and low- and middle-income countries. *Neurosurgery.* 2022, 90:605-12. [10.1227/neu.0000000000001898](https://doi.org/10.1227/neu.0000000000001898)
57. Habets JG, Janssen ML, Duits AA, et al.: Machine learning prediction of motor response after deep brain stimulation in Parkinson's disease-proof of principle in a retrospective cohort. *PeerJ.* 2020, 8:e10317. [10.7717/peerj.10317](https://doi.org/10.7717/peerj.10317)
58. Menger RP, Savardekar AR, Farokhi F, Sin A: A cost-effectiveness analysis of the integration of robotic spine technology in spine surgery. *Neurospine.* 2018, 15:216-24. [10.14245/ns.1836082.041](https://doi.org/10.14245/ns.1836082.041)
59. Naik A, Smith AD, Shaffer A, et al.: Evaluating robotic pedicle screw placement against conventional modalities: a systematic review and network meta-analysis. *Neurosurg Focus.* 2022, 52:E10. [10.3171/2021.10.FOCUS21509](https://doi.org/10.3171/2021.10.FOCUS21509)
60. Ho AL, Muftuoglu Y, Pendharkar AV, Sussman ES, Porter BE, Halpern CH, Grant GA: Robot-guided pediatric stereoelectroencephalography: single-institution experience. *J Neurosurg Pediatr.* 2018, 22:1-8. [10.3171/2018.5.PEDS17718](https://doi.org/10.3171/2018.5.PEDS17718)
61. Neudorfer C, Hunsche S, Hellmich M, El Majdoub F, Maarouf M: Comparative study of robot-assisted versus conventional frame-based deep brain stimulation stereotactic neurosurgery. *Stereotact Funct Neurosurg.* 2018, 96:327-34. [10.1159/000494736](https://doi.org/10.1159/000494736)
62. PhD defense of François-Xavier . (2021). Accessed: August 16, 2023: <https://www.timc.fr/en/phd-defense-francois-xavier-carton-121321>.
63. Han R, Jones CK, Lee J, et al.: Deformable MR-CT image registration using an unsupervised, dual-channel network for neurosurgical guidance. *Med Image Anal.* 2022, 75:102292. [10.1016/j.media.2021.102292](https://doi.org/10.1016/j.media.2021.102292)
64. Padoy N, Hager GD: Human-machine collaborative surgery using learned models . *IEEE Int Conf Robot Autom.* 2011, 5285:10.1109/ICRA.2011.5980250
65. Hu D, Gong Y, Hannaford B, Seibel EJ: Semi-autonomous simulated brain tumor ablation with ravenii surgical robot using behavior tree. *IEEE Int Conf Robot Autom.* 2015, 2015:3868-75. [10.1109/ICRA.2015.7139738](https://doi.org/10.1109/ICRA.2015.7139738)

66. Hu D, Gong Y, Hannaford B, Seibel EJ: Path planning for semi-automated simulated robotic neurosurgery . Rep U S. 2015;2015:2639-45. [10.1109/IROS.2015.7353737](https://doi.org/10.1109/IROS.2015.7353737)
67. Marcus HJ, Vakharia VN, Sparks R, et al.: Computer-assisted versus manual planning for stereotactic brain biopsy: a retrospective comparative pilot study. Oper Neurosurg (Hagerstown). 2020, 18:417-22. [10.1093/ons/opsz177](https://doi.org/10.1093/ons/opsz177)
68. Wong S, Baltuch GH, Jaggi JL, Danish SF: Functional localization and visualization of the subthalamic nucleus from microelectrode recordings acquired during DBS surgery with unsupervised machine learning. J Neural Eng. 2009, 6:026006. [10.1088/1741-2560/6/2/026006](https://doi.org/10.1088/1741-2560/6/2/026006)
69. Valsky D, Marmor-Levin O, Deffains M, Eitan R, Blackwell KT, Bergman H, Israel Z: Stop! border ahead: automatic detection of subthalamic exit during deep brain stimulation surgery. Mov Disord. 2017, 32:70-9. [10.1002/mds.26806](https://doi.org/10.1002/mds.26806)
70. Kostoglou K, Michmizos KP, Stathis P, Sakas D, Nikita KS, Mitsis GD: Classification and prediction of clinical improvement in deep brain stimulation from intraoperative microelectrode recordings. IEEE Trans Biomed Eng. 2017, 64:1123-30. [10.1109/TBME.2016.2591827](https://doi.org/10.1109/TBME.2016.2591827)
71. Jermyn M, Mok K, Mercier J, et al.: Intraoperative brain cancer detection with Raman spectroscopy in humans. Sci Transl Med. 2015, 7:274ra19. [10.1126/scitranslmed.aaa2384](https://doi.org/10.1126/scitranslmed.aaa2384)
72. Cakmakci D, Karakaslar EO, Ruhland E, et al.: Machine learning assisted intraoperative assessment of brain tumor margins using HRMAS NMR spectroscopy. PLoS Comput Biol. 2020, 16:e1008184. [10.1371/journal.pcbi.1008184](https://doi.org/10.1371/journal.pcbi.1008184)
73. Ritschel K, Pechlivanis I, Winter S: Brain tumor classification on intraoperative contrast-enhanced ultrasound. Int J Comput Assist Radiol Surg. 2015, 10:531-40. [10.1007/s11548-014-1089-6](https://doi.org/10.1007/s11548-014-1089-6)
74. Li Y, Charalampaki P, Liu Y, Yang GZ, Giannarou S: Context aware decision support in neurosurgical oncology based on an efficient classification of endomicroscopic data. Int J Comput Assist Radiol Surg. 2018, 13:1187-99. [10.1007/s11548-018-1806-7](https://doi.org/10.1007/s11548-018-1806-7)
75. Akut R: Wavelet based deep learning approach for epilepsy detection. Health Inf Sci Syst. 2019, 7:8. [10.1007/s13755-019-0069-1](https://doi.org/10.1007/s13755-019-0069-1)
76. Iasemidis LD: Seizure prediction and its applications. Neurosurg Clin N Am. 2011, 22:489. [10.1016/j.nec.2011.07.004](https://doi.org/10.1016/j.nec.2011.07.004)
77. Shi Z, Miao C, Schoepf UJ, et al.: A clinically applicable deep-learning model for detecting intracranial aneurysm in computed tomography angiography images. Nat Commun. 2020, 11:6090. [10.1038/s41467-020-19527-w](https://doi.org/10.1038/s41467-020-19527-w)
78. Burns JE, Yao J, Summers RM: Vertebral body compression fractures and bone density: automated detection and classification on CT images. Radiology. 2017, 284:788-97. [10.1148/radiol.2017162100](https://doi.org/10.1148/radiol.2017162100)
79. Han Z, Wei B, Mercado A, Leung S, Li S: Spine-GAN: semantic segmentation of multiple spinal structures . Med Image Anal. 2018, 50:23-35. [10.1016/j.media.2018.08.005](https://doi.org/10.1016/j.media.2018.08.005)
80. Lim J, Kim J, Cheon S: A deep neural network-based method for early detection of osteoarthritis using statistical data. Int J Environ Res Public Health. 2019, 16:1281. [10.3390/ijerph16071281](https://doi.org/10.3390/ijerph16071281)
81. Hale AT, Stonko DP, Wang L, Strother MK, Chambless LB: Machine learning analyses can differentiate meningioma grade by features on magnetic resonance imaging. Neurosurg Focus. 2018, 45:E4. [10.3171/2018.8.FOCUS18191](https://doi.org/10.3171/2018.8.FOCUS18191)
82. Ruiz L, Martín A, Urbanos G, et al.: Multiclass brain tumor classification using hyperspectral imaging and supervised machine learning. Conf Des Circuits Int Sys. 2020, [10.1109/DCIS51330.2020.9268650](https://doi.org/10.1109/DCIS51330.2020.9268650)
83. Djirackor L, Halldorsson S, Niehusmann P, et al.: Intraoperative DNA methylation classification of brain tumors impacts neurosurgical strategy. Neurooncol Adv. 2021, 3:149. [10.1093/onoajnl/vdab149](https://doi.org/10.1093/onoajnl/vdab149)
84. Badža MM, Barjaktarović MČ: Classification of brain tumors from MRI images using a convolutional neural network. Appl Sci. 2020, 10:1999. [10.3390/app10061999](https://doi.org/10.3390/app10061999)
85. Rosa E de la, Sima DM, Vyvere TV, et al.: A radiomics approach to traumatic brain injury prediction in CT scans. IEEE 16th Int Symp Biomed Imaging. 2019, [10.1109/ISBI.2019.8759229](https://doi.org/10.1109/ISBI.2019.8759229)
86. Huang CY, Lee CC, Yang HC, et al.: Radiomics as prognostic factor in brain metastases treated with Gamma Knife radiosurgery. J Neurooncol. 2020, 146:439-49. [10.1007/s11060-019-03343-4](https://doi.org/10.1007/s11060-019-03343-4)
87. Yamazawa E, Takahashi S, Shin M, et al.: MRI-based radiomics differentiates skull base chordoma and chondrosarcoma: a preliminary study. Cancers (Basel). 2022, 14:3264. [10.3390/cancers14133264](https://doi.org/10.3390/cancers14133264)
88. Lei Y, Harms J, Wang T, et al.: MRI-only based synthetic CT generation using dense cycle consistent generative adversarial networks. Med Phys. 2019, 46:3565-81. [10.1002/mp.13617](https://doi.org/10.1002/mp.13617)
89. Kazemifar S, McGuire S, Timmerman R, et al.: MRI-only brain radiotherapy: assessing the dosimetric accuracy of synthetic CT images generated using a deep learning approach. Radiother Oncol. 2019, 136:56-63. [10.1016/j.radonc.2019.03.026](https://doi.org/10.1016/j.radonc.2019.03.026)
90. Liu Y, Chen A, Shi H, et al.: CT synthesis from MRI using multi-cycle GAN for head-and-neck radiation therapy. Comput Med Imaging Graph. 2021, 91:101953. [10.1016/j.compmedimag.2021.101953](https://doi.org/10.1016/j.compmedimag.2021.101953)
91. Oulbacha R, Kadoury S: MRI to CT synthesis of the lumbar spine from a pseudo-3D cycle GAN . IEEE 17th Int Symp Biomed Imaging. 2020, [10.1109/ISBI45749.2020.9098421](https://doi.org/10.1109/ISBI45749.2020.9098421)
92. Bourbonne V, Jaouen V, Hognon C, et al.: Dosimetric validation of a GAN-based pseudo-CT generation for MRI-only stereotactic brain radiotherapy. Cancers (Basel). 2021, 13:[10.3390/cancers13051082](https://doi.org/10.3390/cancers13051082)
93. Song TA, Chowdhury SR, Yang F, Dutta J: PET image super-resolution using generative adversarial networks. Neural Netw. 2020, 125:83-91. [10.1016/j.neunet.2020.01.029](https://doi.org/10.1016/j.neunet.2020.01.029)
94. Fei Y, Zu C, Jiao Z, et al.: Classification-aided high-quality PET image synthesis via bidirectional contrastive GAN with shared information maximization. Med Image Comput Comput Assist Interv. 2022, 13436:[10.1007/978-3-031-16446-0\\_50](https://doi.org/10.1007/978-3-031-16446-0_50)
95. Lin W, Lin W, Chen G, et al.: Bidirectional mapping of brain MRI and PET with 3D reversible Gan for the diagnosis of Alzheimer's disease. Front Neurosci. 2021, 15:646013. [10.3389/fnins.2021.646013](https://doi.org/10.3389/fnins.2021.646013)
96. Gianfrancesco MA, Tamang S, Yazdany J, Schmajuk G: Potential biases in machine learning algorithms using electronic health record data. JAMA Intern Med. 2018, 178:1544-7. [10.1001/jamainternmed.2018.3763](https://doi.org/10.1001/jamainternmed.2018.3763)
97. Azad TD, Ehresman J, Ahmed AK, et al.: Fostering reproducibility and generalizability in machine learning for clinical prediction modeling in spine surgery. Spine J. 2021, 21:1610-16. [10.1016/j.spinee.2020.10.006](https://doi.org/10.1016/j.spinee.2020.10.006)

98. Chang Q, Qu H, Zhang Y, et al.: Synthetic learning: learn from distributed asynchronized discriminator GAN without sharing medical image data. IEEE/CVF Conf Comput Vis Pattern Recogn. 2020, [10.48550/arXiv.2006.00080](https://arxiv.org/abs/2006.00080)
99. Schonfeld E, Veeravagu A: Demonstrating the successful application of synthetic learning in spine surgery for training multi-center models with increased patient privacy. Sci Rep. 2023, 13:12481. [10.1038/s41598-023-39458-y](https://doi.org/10.1038/s41598-023-39458-y)
100. Cheung AT, Nasir-Moin M, Fred Kwon YJ, et al.: Methods and impact for using federated learning to collaborate on clinical research. Neurosurgery. 2023, 92:431-8. [10.1227/neu.0000000000002198](https://doi.org/10.1227/neu.0000000000002198)
101. Pati S, Baid U, Edwards B, et al.: Federated learning enables big data for rare cancer boundary detection . Nat Commun. 2022, 13:7346. [10.1038/s41467-022-33407-5](https://doi.org/10.1038/s41467-022-33407-5)
102. Jiang LY, Liu XC, Nejatian NP, et al.: Health system-scale language models are all-purpose prediction engines. Nature. 2023, 619:357-62. [10.1038/s41586-023-06160-y](https://doi.org/10.1038/s41586-023-06160-y)
103. Oncology compass | Congress updates and practice changing publications for oncologist . (2023). Accessed: August 16, 2023: <https://oncologycompass.ch/blog/post/ai-powered-language-model-predicts-readmissions...>
104. Ali R, Tang OY, Connolly ID, et al.: Performance of ChatGPT, GPT-4, and Google Bard on a neurosurgery oral boards preparation Question Bank. Neurosurgery. 2023, [10.1227/neu.0000000000002551](https://doi.org/10.1227/neu.0000000000002551)