An introduction and overview of machine learning in neurosurgical care

Senders, J.T.; Zaki, M.M.; Karhade, A.V.; Chang, B.; Gormley, W.B.; Broekman, M.L.; ... ; Arnaout, O.

Acta Neurochir (2018) 160:29–38
https://doi.org/10.1007/s00701-017-3385-8

REVIEW ARTICLE - NEUROSURGICAL TECHNIQUES

An introduction and overview of machine learning in neurosurgical care

Joeky T. Senders [1,2] **· Mark M. Zaki** [2] **· Aditya V. Karhade** [2] **· Bliss Chang** [2] **· William B. Gormley** [2] **· Marike L. Broekman** [1,2] **· Timothy R. Smith** [2] **· Omar Arnaout** [2]

Received: 18 October 2017 / Accepted: 29 October 2017 / Published online: 13 November 2017

Abstract

Background Machine learning (ML) is a branch of artificial intelligence that allows computers to learn from large complex datasets without being explicitly programmed. Although ML is already widely manifest in our daily lives in various forms, the considerable potential of ML has yet to find its way into mainstream medical research and day-to-day clinical care. The complex diagnostic and therapeutic modalities used in neurosurgery provide a vast amount of data that is ideally suited for ML models. This systematic review explores ML's potential to assist and improve neurosurgical care.

Method A systematic literature search was performed in the PubMed and Embase databases to identify all potentially relevant studies up to January 1, 2017. All studies were included that evaluated ML models assisting neurosurgical treatment.

Results Of the 6,402 citations identified, 221 studies were selected after subsequent title/abstract and full-text screening. In these studies, ML was used to assist surgical treatment of patients with epilepsy, brain tumors, spinal lesions, neurovascular pathology, Parkinson's disease, traumatic brain injury, and hydrocephalus. Across multiple paradigms, ML was found to be a valuable tool for presurgical planning, intraoperative guidance, neurophysiological monitoring, and neurosurgical outcome prediction.

Conclusions ML has started to find applications aimed at improving neurosurgical care by increasing the efficiency and precision of perioperative decision-making. A thorough validation of specific ML models is essential before implementation in clinical neurosurgical care. To bridge the gap between research and clinical care, practical and ethical issues should be considered parallel to the development of these techniques.

Keywords Artificial intelligence · Machine learning · Neurosurgery

Introduction

In the last few decades, the volume and complexity of biomedical data have grown beyond the physician's ability to extract all meaningful data patterns using conventional statistical methods alone. Hospitals produce every day vast quantities of unstructured data such as imaging data, genomic information, free text, and data streams from monitoring devices. This calls for novel methods to help physicians effectively analyze 'big data' [11, 17, 40].

Machine learning (ML) is a branch of artificial intelligence that enables computer algorithms to learn and improve by studying large datasets, without explicitly being programmed [50]. In recent decades, many other ML algorithms were developed that learn through experience ranging from artificial neural networks, to support vector machines, to random forests, to naive Bayes, and K-nearest neighbors [33]. This branch of artificial intelligence has really taken off in the last few years because computers acquire the computational power to build complex models that can actually process and learn

Joeky T. Senders and Mark M. Zaki shared first author.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00701-017-3385-8) contains supplementary material, which is available to authorized users.

✉ Omar Arnaout
oarnaout@bwh.harvard.edu

1 Department of Neurosurgery, University Medical Center Utrecht, Utrecht University, Heidelberglaan 100, Utrecht 3584 CX, The Netherlands

2 Computational Neurosciences Outcomes Center, Department of Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, 75 Francis Street, Boston, MA 02115, USA

from raw unstructured data [54]. This is a revolutionary change. For example, instead of clinicians who have to come up with meaningful structured features to describe their data, the data can be fed to the algorithm in its raw form, and the algorithm itself will try to learn which features are associated with the outcome of interest. This is especially useful in the bulk of unstructured biomedical data, such as pictures of pathology slices, radiological imaging, free text, and longitudinal trends in monitoring devices. Additionally, ML can read the numeric values that underlie the digital information that is produced by medical technology.

Neurosurgery is particularly a specialty that can benefit from ML. The complex diagnostic and therapeutic modalities used in this specialty provide a vast quantity and variety of complex data, thus presenting an opportune framework for the application of ML models. In a recent systemic review, we compared the performance of ML models with that of clinicians across the spectrum of neurosurgical applications and found that ML models already outperform clinicians most of the time with a median absolute improvement in accuracy and area under the receiving operating curve (AUC) of 13% and 0.14, respectively [58].

Very few ML applications have found, however, their way to actual neurosurgical practice [17]. The current review aims to present a brief introduction into the concepts and theoretical background of ML, as well as an overview of the current literature on ML models directly assisting in direct neurosurgical care. An overview of the current neurosurgical literature on ML provides valuable insight into how ML might transform neurosurgical care in the near and long-term future. Third, we try to identify hurdles in creating, validating, and deploying ML models in the clinical setting. We defined direct neurosurgical care as applications directly supporting or related to the surgical care of neurosurgery patients, including preoperative planning, intra-operative assistance, neurophysiological monitoring, and surgical outcome prediction. Other clinical and scientific applications such as diagnostic classification or clustering of genomic data, although equally important, are outside the scope of this systematic review due to the less surgical nature of these applications.

Machine learning

Within the field of ML, a broad distinction could be made between supervised, unsupervised, and reinforcement learning. Supervised learning algorithms are the most commonly used algorithms and learn from "labeled" training data to produce models that can make predictions on previously unseen data [29, 37, 50, 59]. This learning aspect demonstrates the difference between ML and traditional programming: In traditional programming, a programmer manually writes a set of rules—"the program"—to generate a desired output from a given set of input variables. In ML, the input is provided together with the desired output, and computer algorithms are asked to derive the "rules". The product of this process is, therefore, not the desired output but a model that can predict the desired output in previously unseen data [50, 54]. Both the output and input can take a wide range of possibilities.

In neurosurgery, output could range from segmentation of brain tumors on the preoperative magnetic resonance imaging (MRI) scan [1] to intraoperative differentiation of tumor and brain tissue [55], and to prediction of future intracranial pressure (ICP) trends on the neurocritical care unit [53]. Relevant input features for these ML applications can be, for example, MRI, Raman spectroscopy, and ICP sensors, respectively. In the learning process, algorithms try to find the optimal combination of input features and weights given to these features in the model, thereby minimizing the difference between the predicted and actual outcome. If the model is overly complex, such as containing too many features relative to the number of cases, the model erroneously incorporates random error or noise as signal; this is referred to as overfitting of the training set which artificially reduces the prediction error in the training set at the cost of a reduced external generalizability. To overcome this problem, an ML model should be tested on data not involved in the learning process, referred to as the test or validation set [17, 37]. Additional techniques, such as cross validation, provide additional robustness to the process.

Unsupervised learning techniques, on the other hand, are able to categorize previously unlabeled data. Since the algorithm analyzes unlabeled input features only, it aims to find the hidden structure or relationships within the data; as such, unsupervised learning algorithms are very useful for association and clustering tasks. These algorithms can be powerful tools for detecting previously unknown patterns in multidimensional data that may not be prima facie detectable by conventional statistical analysis. Unsupervised learning could, for example, be used for clustering and discovering patterns in genomic data for brain tumor patients. It could also function as a preparatory step in identifying key features and providing labeling for a subsequent supervised learning task.

Reinforcement learning focuses on determining the ideal behavior within a specific context based on a simple reward feedback on their actions. A reinforcement learning model monitors the reward it receives from its environment after each step and aims to maximize the total amount of reward. Robotics performing surgical procedures autonomously can be a future application of reinforcement learning [37, 50, 54].

Methods

A systematic literature review was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The PubMed and Embase databases were searched to identify all potentially relevant studies up to January 1, 2017. The search syntax was built with the guidance of a professional librarian using search terms related to ‘artificial intelligence’ and ‘neurosurgery’. The exact search syntax for the PubMed and Embase databases is provided in Supplementary Table S1.

All biomedical studies were included that evaluated ML models assisting in neurosurgical care. Exclusion criteria were conference abstracts, animal models, no surgical patients, no full text, or no English language. The initial search was performed by two independent authors (J.S., A.K.). Data extraction was performed by three independent authors (J.S., M.Z., B.C.). Disagreements were solved by discussion including another author (O.A.).

Data obtained from each study were: (1) name of first author, (2) year of publication, (3) treatment stage, (4) disease condition, (5) ML algorithm, (5) input features used, (6) neurosurgical application evaluated and (7) result/conclusion of this study. Based on these data, we calculated the distribution of all published articles within the domains of neurosurgical treatment stage, disease condition, ML algorithm, input features, and neurosurgical application.

We considered a quantitative synthesis to be inappropriate due to the heterogeneity in neurosurgical applications. A qualitative synthesis of results is provided by means of a narrative approach; however, to summarize the findings in some quantitative form, the median absolute performance was calculated for the most frequently evaluated application by means of the most frequently reported statistical measures of performance.

Results

Of the 6.402 citations initially identified, 221 were selected after consequent title/abstract and full-text screening (Fig. 1). We found an exponential growth in the number of studies evaluating ML models as an assisting tool across multiple paradigms of neurosurgical care, including presurgical planning, intraoperative guidance, neurophysiological monitoring, and neurosurgical outcome prediction (Fig. 2).

ML enhanced neurosurgical care in patients with a wide variety of neurosurgical disorders, including epilepsy, brain tumors, spinal lesions, neurovascular pathology, Parkinson’s diseases, traumatic brain injury, and cerebrovascular abnormalities (Fig. 3a). Algorithms used in order of decreasing frequency were artificial neural networks, support vector machines, fuzzy C-means, Bayesian Learning, random forests, quadratic discriminant analysis, linear discriminant analysis, Gaussian mixture models, logistic regression, k-nearest neighbor, natural language processing, and k-means (Fig. 3b). Magnetic resonance imaging (MRI) data were the most frequently used input features (Fig. 3c), and radiological brain tumor segmentation the most frequently evaluated application (Fig. 3d). Details regarding the treatment stage, disease condition, algorithms, input features, applications, results, and conclusions on each individual study are provided in Supplementary Tables S2-S8.

ML models were used preoperatively for radiological segmentation of a variety of lesions and/or anatomical structures on MRI, computed tomography (CT), positron emission tomography (PET), and/or ultrasound [3, 7, 26, 36, 38]. This included brain tumors, estimated maps for tumor infiltration, critical/target brain structures, epileptic foci, lumbar discs, cervical vertebrae, and arteriovenous malformations. Epileptic foci were also localized or identified by means electrophysiological monitoring [25, 42]. Lastly, ML models were used to determine beam orientation in stereotactic radiosurgery [64].

ML models were used intraoperatively in Parkinson patients to localize the subthalamic nucleus and estimate the volume of activated tissue using features from micro-electrode recordings during deep brain stimulation (DBS) [66, 67]. In brain tumor patients, algorithms differentiated tumor and healthy brain tissue by means of spectroscopy, mass spectrometry, or ultrasound [21, 39, 55]. In addition, ML models were used to align preoperative imaging with the intraoperative guiding system [68], localize eloquent area by means of electrophysiological monitoring [28, 49, 52], increase the specificity of somatosensory evoked potential (SEP) monitoring [24], and compensate for respiratory motion during stereotactic radiosurgery (SRS) [4]. Lastly, one paper investigated computer-assisted surgery in the form of an intelligent trajectory planner, estimating the risk of different surgical strategies [16].

In the context of neurophysiological monitoring, ML models were evaluated for abnormality detection and prediction of future intracranial pressure (ICP) trends, but also for non-invasive ICP assessment using other physiological parameters [48, 53, 57, 65]. ML-based natural language processing detected surgical site infections using free-text clinical notes from the electronic health records [12]. A few studies explored the capability of intracranial devices to predict and prevent seizures or function as a brain-computer interface for severely disabled patients [32, 51, 63].

ML models were used to predict symptom improvement, patient satisfaction, seizure freedom, Glasgow Outcome Score, recurrence, and survival after neurosurgery for a variety of disease conditions, including lumbar disc hernia, lumbar spinal stenosis, cervical myelopathy, Parkinson’s Disease, brain tumor, epilepsy, arteriovenous malformations, hydrocephalus, aneurysmal subarachnoid hemorrhage, and traumatic brain injury [2, 5, 6, 8–10, 20, 23, 30, 60–62]. MRI, clinical, and electrophysiological data were used as input features in the model dependent on the disease condition and application evaluated.

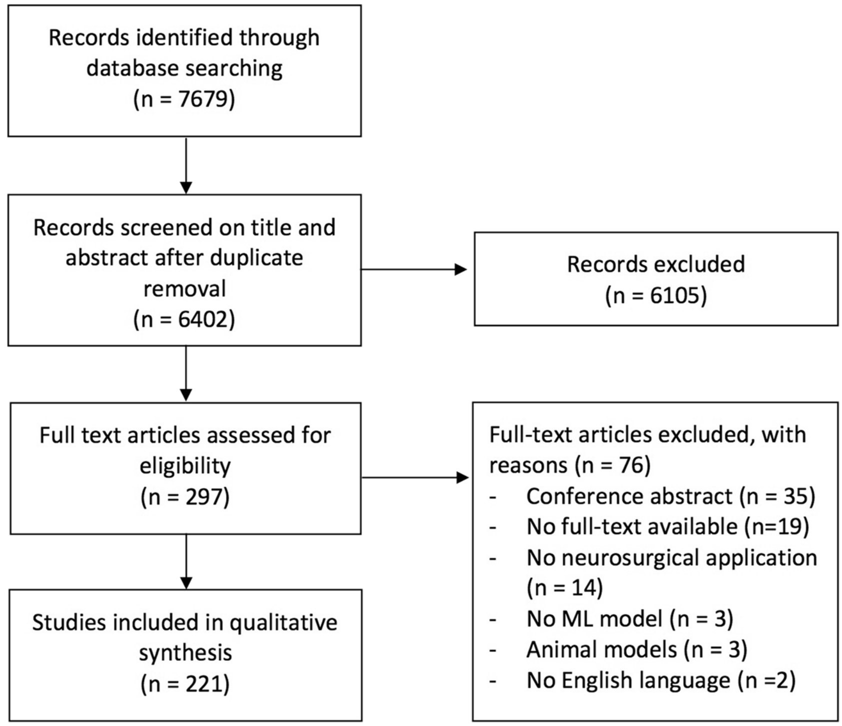

Fig. 1 Flowchart depicting study selection

Quantitative summary on the five most frequently evaluated applications

A total of 58 studies evaluated ML models for brain tumor segmentation, reporting a median accuracy (*n* = 14 studies) of 92% [interquartile range (IQR), 85–95%], AUC (*n* = 5) of 0.93 (IQR, 0.92–0.93), and dice similarity coefficient (*n* = 7) of 88% (IQR, 84-93%). A total of 29 studies evaluated ML models used for radiological segmentation of critical/target brain structures, reporting a median accuracy (*n* = 6) of 94% (IQR, 86-99%)

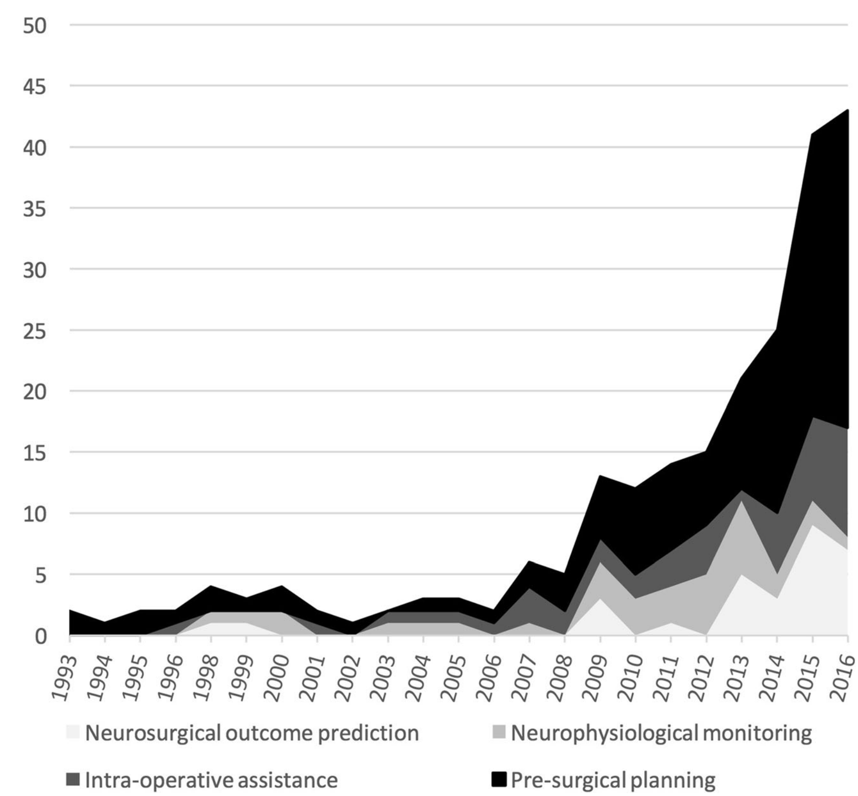

Fig. 2 Number of papers published each year on machine learning in neurosurgery categorized per treatment stage

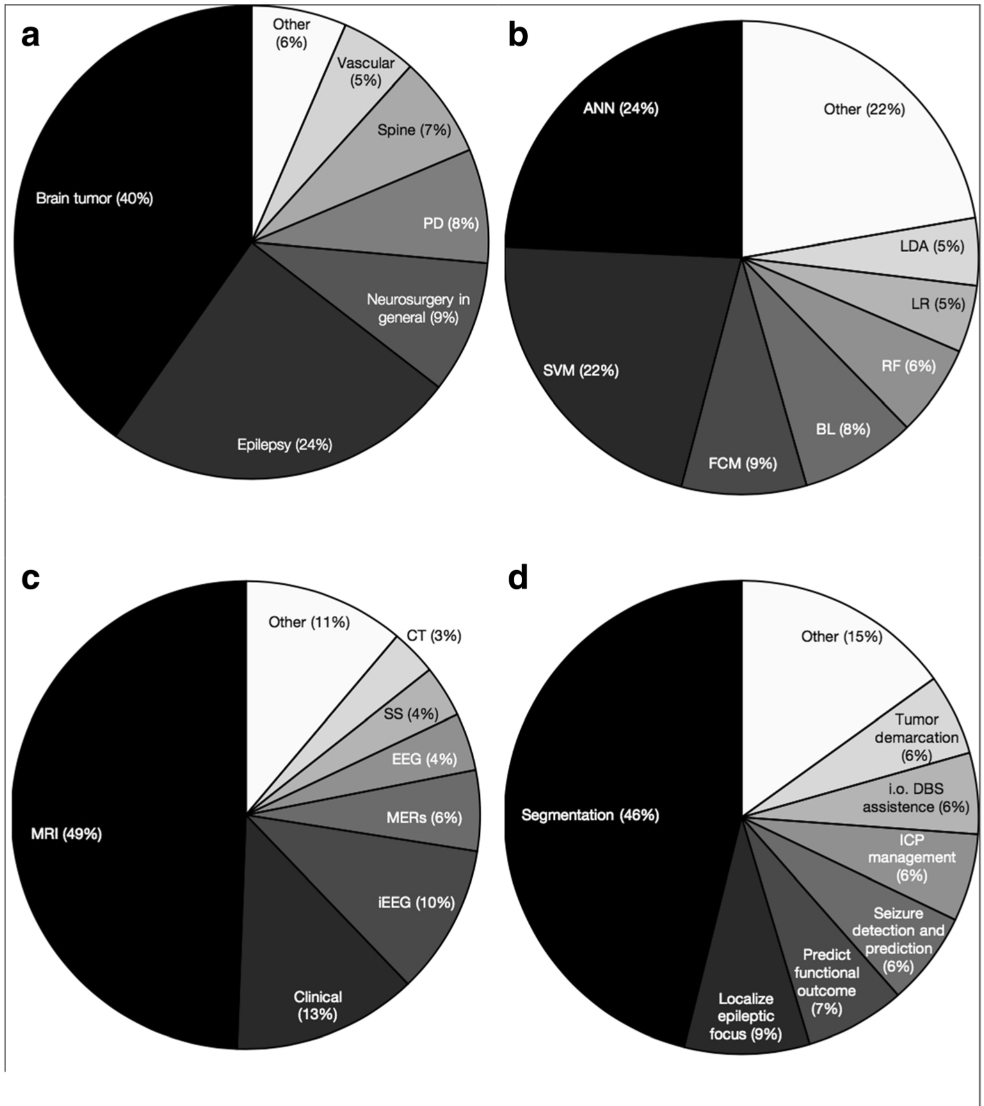

Fig. 3 Pie charts representing the distribution of disease conditions (**a**), machine learning algorithms (**b**), input features (**c**), and general applications in the neurosurgical literature on machine learning (**d**). *ANN* artificial neural networks; *BL* bayesian learning; *CT* computer tomography; *DBS* deep brain stimulation; *EEG* electro-encephalography; *iEEG* intracranial electroencephalography; *ICP* intracranial pressure; *FCM* fuzzy c-means; *MER* micro-electrode recordings; *MRI* magnetic resonance imaging; *LDA* linear discriminant analysis; *LR* logistic regression; *PD* Parkinson's disease; *RF* random forests; *SS* spectroscopy; *SVM* support vector machines

and dice similarity coefficient ($n = 5$) of 91% (IQR, 90-92%). A total of 18 studies evaluated ML models used for the prediction of laterality of the epileptogenic focus, reporting a median accuracy ($n = 13$) of 86% (IQR, 80-94%). A total of 13 studies evaluated ML models used for the detection of seizures by means of intracranial electro-encephalography, reporting a median sensitivity ($n = 9$) of 96% (IQR, 94-98%), false detection rate ($n = 8$) of 0.17 per hour (IQR, 0.08–0.27), and detection delay ($n = 5$) of 11 s (IQR, 3.4–14.8 s). A total of 13 studies evaluated ML models used for intraoperative tumor demarcation reporting a median accuracy ($n = 8$) of 89% (IQR, 83-97%). Results for each individual study are provided in Supplementary Table S2-S8.

Discussion

This paper reviews all available literature on the application of ML in neurosurgical care. As evident by the data, ML has the potential to impact any subspecialty within neurosurgery at any time point of an individual patient's care. Although a wide range of potential applications was observed, ML models were most frequently evaluated for presurgical planning of brain tumor patients using artificial neural networks as algorithms, MRI data as input, and obtaining labeled segments as output.

Machine learning in medicine

A recent study aimed to systematically review the current neurosurgical literature on machine learning too; however, it identified only 37 studies, thereby lacking the vast majority of relevant literature [14]. This study emphasizes nevertheless the potential ML has to improve clinical and surgical management. Other studies have reviewed the application of ML to improve general surgical care and also showed that ML could be a valuable tool assisting preoperative planning [13, 40], intraoperative guidance [40, 41, 43], neurophysiological monitoring [35], and outcome prediction [8].

Canchi et al. [13] reviewed the use of ML models to predict the risk of rupture of abdominal aortic aneurysms and found that ML could be a very valuable tool for surgical planning due to its ability to analyze complex radiological data combined with clinical data. In contrast to the extensive use of radiological input in neurosurgery and vascular surgery, ML models in plastic surgery focus more on interpreting images, probably due to the visual character of input and outcome in plastic surgery [40].

An interesting intraoperative application discussed by Kanevsky et al. [40] is the potential of ML to offer a more objective method for assessing "surgical competency". By means of wearable technology, residents could record their cases and by integrating these intraoperative recordings to the outcomes, ML could identify techniques that lead to certain outcomes. Although this was relatively underexposed in the neurosurgical literature, two reviews focused on the role of ML in intelligent and autonomous surgical robots. These papers explored the usefulness for intraoperative recognition of anatomical landmarks, image guidance and navigation, computer-assisted surgical training, performance metric of surgical competency, and automation of simple surgical procedures [41, 43].

In addition, there's an emerging interest in the application of ML models in critical care medicine [35]. In the intensive care setting, a vast amount of data is produced by a variety of modalities, ranging from cardiac rhythm, to blood pressure, to oxygen saturation, to intracranial pressure monitoring. By combining these sources of physiological parameters, ML can provide personalized risk assessment, reduce false alarm rates, and increase the efficiency of physicians [35].

Since ML can produce very powerful prediction models, ML algorithms are increasingly being evaluated as prediction models for surgical outcome [8]. Predictions produced by ML models are increasing in accuracy and can be tailored to the individual patient due to the expanding volume and dimensions of data collected on each patient.

Across the broad spectrum of medicine, similar trends can be observed compared to the realms of neurosurgery. Again, radiology and oncology seem to lead the field in the exploration of applied ML models [11, 18, 27, 44, 46, 47, 54, 56]. This is possibly due to the vast amount of complex data that provide a good framework for the creation of ML methods, but also a demand for automated methods to analyze these data [44]. In the clinical care of oncological patients, radiological data can be combined with clinical [45], genomic [46], and histological data [56] to assist physicians with diagnosis [27, 56], surgical care [39], drug development [47], radiotherapy [11], and prognosis of cancer patients [45].

To our knowledge, we present the first comprehensive review on the application of ML in neurosurgical care, demonstrating similar trends in neurosurgical literature compared with other surgical fields and medicine in general, and we identify some hurdles on the road to implementation of clinical ML models.

Limitations

A few limitations should be mentioned. First, this review does not provide an in-depth quantitative and qualitative analysis of all studies due to the large number of studies and the heterogeneity in clinical conditions, applications, and data-modeling techniques used across all studies. Additionally, it limits itself to applications directly assisting in the perioperative care underexposing the potential of ML models to assist in diagnosis and research. Lastly, the quantitative summary of performance of the most frequently evaluated applications should be interpreted with great caution due to the heterogeneity in specific disease conditions, ML algorithms, input features, and reported statistical measures of performance. However, we think these limitations are proportionate to the strengths of this study and the quantitative summary provides a general idea of the current performance of ML models. This systematic review elaborates on the basic concepts of ML and is the first paper that presents an overview of the neurosurgical literature on ML, providing insight into the near and long-term future direction of ML in neurosurgery.

Implications

Although ML models have been demonstrated to be an asset in a variety of neurosurgical applications, we anticipate that pattern recognition in radiological imaging will be one of the first applications that will find its way into clinical practice. Three papers demonstrated a better performance in radiological tumor segmentation tasks by ML models compared to radiologists [15, 19, 22]. These algorithms can provide results in the order of a few minutes or even seconds, compared to much lengthier evaluations by human inspection [34]. By using each voxel as individual input feature, ML can dramatically increase the amount of information extracted from images compared to visual evaluation alone [44].

This implicitly raises the concern of ML models replacing the physician. We and others [44], however, believe that these tools will only supplement and augment the physician. ML algorithms can take over simple and repetitive tasks, enabling clinicians to focus on the more 'complex' or ambiguous cases. Especially in settings where the physician's time is in high demand, this would result in a more sustainable deployment of clinical expertise and physicians working "at the top of their license". Clinicians cannot blindly follow predictions made by ML models; they should always consider if the models are built on a clinically meaningful foundation of input features and are producing reliable output. Most importantly, even if ML models can perform the analysis with very high accuracy, clinicians still must consider the implications of any given analysis and how it fits with the active clinical scenario. Successful introduction of ML into the clinical arena would require, therefore, an extended knowledge of the creation, validation, and statistical pitfalls of ML models applied in the clinical context.

In both neurosurgery and medicine in general, the progress in the field of artificial intelligence is continuously propelled by the increasingly available large amounts of biomedical data. By analyzing patterns across large heterogeneous data sets, ML provides methodology for understanding how numerous factors influence desired outcomes. Therefore, incorporating ML models to determine optimal therapy in individual patients parallels the growing trend towards precision medicine [35, 46]. On the personalized level, this means giving patients the treatment that fits best to their needs. On the systemic level, timely refraining from unnecessary testing or treatment would result in decreased complication rates and healthcare costs. ML models contributing to precision medicine could also be effective in the research arena. In clinical research, ML models can help identify patients for whom therapy would be most effective, reducing the necessary sample size and economic burden of randomized clinical trials. In basic research, patterns across clinical, histological, radiological, and genetic data can be analyzed to construct a multimodal disease phenotype and provide an integrated understanding of disease development and progression.

Future challenges

Still significant hurdles remain associated with the creation, validation, and deployment of these models in clinical care. The mechanisms underlying ML models can be very difficult to interpret and scrutinize. Conventional statistical methods like linear and logistic regressions, address relationships between input and output features in an a priori intuitive fashion compared to the relatively modern ML algorithms. Although ML models can produce powerful predictions, this abstraction, can result in hesitancy to deploy them. However, introducing a new technique into clinical care based on the simple observation that it's effective without universal understanding of its inner workings, is not a new concept in medicine; many pharmaceutical drugs have been introduced in a similar manner [54]. Furthermore, we suspect that as more ML models are produced, validated, and as such become more familiar, clinicians will gain better insight into their underlying mechanisms and will be more comfortable with deploying them in daily practice.

Entrusting ML models with medical tasks raises the question of liability. As clinicians make the final decision or interpretation, it can be argued that they retain the entire responsibility for the outcome; however, as the algorithms and thus their applications become more complex, liability could shift towards institutions involved in the production, validation, and regulation of these algorithms [44]. Due to the learning aspect, ML produces dynamic models that could improve after they have been put into practice. This calls for regulation on performance standards of implemented ML models. Should manufactures freeze the learning process after production and thereby deliver static models alone, or should institutions be allowed to augment performance by means of setting-specific training data, referred to as 'training on the job'? Furthermore, should individual institutions be able to contribute to the training of algorithms on multi-institutional level, and what institutional requirements should be met in this context?

Lastly, the performance of ML is highly dependent on the quality of input data. Due to privacy consideration and poor registrations, the amount of available data could still be insufficient to train ML models effectively. On the other side, models can be trained merely on the patients for whom data were sufficient, thereby introducing selection bias. In the latter case, ML models can perform poorly in the clinical setting compared to the research setting due to lack of generalizability. This also highlights the importance of creating national HIPAA-compliant data-sharing solutions that allow us to move beyond institutional data and tap the potential of heterogeneous big data sets.

To bridge the gap between research and clinical practice, we recommend that future studies focus on the technical aspects of developing ML for clinical applications, but also of creating an ethical and legal framework for the creation, validation, and regulation of ML models in clinical care. An extended knowledge of the mechanisms and statistical pitfalls of ML models is required not only for the clinicians who must interpret these analyses but also preferably by means of an integrated team of clinicians, data scientists, ethicists, and administrators. Furthermore, we recommend standardized and transparent validation methods for ML models 'on site' prior to deployment. ML models should operate parallel to clinical experts and be implemented once their performance and error margin is deemed to be sufficient and acceptable, respectively. Major journals are currently embracing the ideology of open-source coding [31]. This means that authors make their code accessible to others to view, learn from, and apply on their own data. We advocate that any predictive ML model should have their code accessible for anyone to examine, since transparency reduces the black-box aspect of ML models, and statistical pitfalls are detected and solved sooner, resulting in both a safer and more efficient implementation of ML in clinical care.

Conclusion

In neurosurgery, ML can improve clinical decision-making by deriving knowledge from the rapidly increasing quantity of unstructured biomedical data. Although a wide range of applications was observed, pre-surgical planning of brain tumor patients using MRI data as input and obtaining labeled segments as output might be one of the first applications that will find its way to neurosurgical practice. Future research should focus on the creation, validation, and statistical pitfalls of ML models on the one hand but also create an ethical and legal framework for the implementation of these modeling techniques in clinical care.

Compliance with ethical standards

Conflict of interest J.T.S., M.M.Z., O.A., A.V.K., B.C., M.L.B., T.R.S. have nothing to disclose. W.B.G.: Codman, Coviden Proctor, Consultant.

Ethical approval For this type of study formal consent is not required.

References

1. Abbasi S, Tajeripour F (2017) Detection of brain tumor in 3D MRI images using local binary patterns and histogram orientation gradient. Neurocomputing 219:526–535
2. Akbari H, Macyszyn L, Da X, Bilello M, Wolf RL, Martinez-Lage M, Biros G, Alonso-Basanta M, O'Rourke DM, Davatzikos C (2016) Imaging surrogates of infiltration obtained via multiparametric imaging pattern analysis predict subsequent location of recurrence of glioblastoma. Neurosurgery 78:572–580. https://doi.org/10.1227/neu.0000000000001202
3. Akbari H, Macyszyn L, Da X, Wolf RL, Bilello M, Verma R, O'Rourke DM, Davatzikos C (2014) Pattern analysis of dynamic susceptibility contrast-enhanced MR imaging demonstrates peritumoral tissue heterogeneity. Radiology 273:502–510. https://doi.org/10.1148/radiol.14132458
4. Antoni ST, Rinast J, Ma X, Schupp S, Schlaefer A (2016) Online model checking for monitoring surrogate-based respiratory motion tracking in radiation therapy. Int J Comput Assist Radiol Surg 11(11):2085-2096. https://doi.org/10.1007/s11548-016-1423-2
5. Arle JE, Perrine K, Devinsky O, Doyle WK (1999) Neural network analysis of preoperative variables and outcome in epilepsy surgery. J Neurosurg 90:998–1004. https://doi.org/10.3171/jns.1999.90.6.0998
6. Asadi H, Kok HK, Looby S, Brennan P, O'Hare A, Thornton J (2016) Outcomes and complications after endovascular treatment of brain Arteriovenous malformations: a prognostication attempt using artificial intelligence. World Neurosurg 96:562–569.e561
7. Azami ME, Hammers A, Jung J, Costes N, Bouet R, Lartizien C (2016) Detection of lesions underlying intractable epilepsy on T1-weighted MRI as an outlier detection problem. PLoS One 11: e0161498
8. Azimi P, Benzel EC, Shahzadi S, Azhari S, Mohammadi HR (2016) The prediction of successful surgery outcome in lumbar disc herniation based on artificial neural networks. J Neurosurg Sci 60:173–177
9. Azimi P, Benzel EC, Shahzadi S, Azhari S, Mohammadi HR (2014) Use of artificial neural networks to predict surgical satisfaction in patients with lumbar spinal canal stenosis: clinical article. J Neurosurg Spine 20:300–305. https://doi.org/10.3171/2013.12.spine13674
10. Azimi P, Mohammadi HR (2014) Predicting endoscopic third ventriculostomy success in childhood hydrocephalus: an artificial neural network analysis. J Neurosurg Pediatr 13:426–432. https://doi.org/10.3171/2013.12.peds13423
11. Bibault JE, Giraud P, Burgun A (2016) Big data and machine learning in radiation oncology: state of the art and future prospects. Cancer Lett 382:110–117. https://doi.org/10.1016/j.canlet.2016.05.033
12. Campillo-Gimenez B, Garcelon N, Jarno P, Chapplain JM, Cuggia M (2013) Full-text automated detection of surgical site infections secondary to neurosurgery in Rennes, France. Stud Health Technol Inform 192:572–575
13. Canchi T, Kumar SD, Ng EY, Narayanan S (2015) A review of computational methods to predict the risk of rupture of abdominal aortic aneurysms. Biomed Res Int 2015:861627. https://doi.org/10.1155/2015/861627
14. Celtikci E (2017) A systematic review on machine learning in neurosurgery: the future of decision making in patient care. Turk Neurosurg. https://doi.org/10.5137/1019-5149.JTN.20059-17.1
15. Clarke LP, Velthuizen RP, Clark M, Gaviria J, Hall L, Goldgof D, Murtagh R, Phuphanich S, Brem S (1998) MRI measurement of brain tumor response: comparison of visual metric and automatic segmentation. Magn Reson Imaging 16:271–279
16. De Momi E, Ferrigno G (2010) Robotic and artificial intelligence for keyhole neurosurgery: the ROBOCAST project, a multi-modal autonomous path planner. Proc Inst Mech Eng H J Eng Med 224: 715–727
17. Deo RC (2015) Machine learning in medicine. Circulation 132: 1920–1930. https://doi.org/10.1161/CIRCULATIONAHA.115.001593
18. Di Ieva A, Boukadoum M, Lahmiri S, Cusimano MD (2015) Computational analyses of arteriovenous malformations in neuroimaging. J Neuroimaging 25:354–360. https://doi.org/10.1111/jon.12200
19. Dolz J, Betrouni N, Quidet M, Kharroubi D, Leroy HA, Reyns N, Massoptier L, Vermandel M (2016) Stacking denoising autoencoders in a deep network to segment the brainstem on MRI in brain cancer patients: a clinical study. Comput Med Imaging Graph 52:8–18. https://doi.org/10.1016/j.compmedimag.2016.03.003
20. Dumont TM, Rughani AI, Tranmer BI (2011) Prediction of symptomatic cerebral vasospasm after aneurysmal subarachnoid hemorrhage with an artificial neural network: feasibility and comparison with logistic regression models. World Neurosurg 75:57–63; **discussion 25-58**. https://doi.org/10.1016/j.wneu.2010.07.007
21. Eberlin LS, Norton I, Dill AL, Golby AJ, Ligon KL, Santagata S, Graham Cooks R, Agar NYR (2012) Classifying human brain tumors by lipid imaging with mass spectrometry. Cancer Res 72:645–654
22. Emblem KE, Nedregaard B, Hald JK, Nome T, Due-Tonnessen P, Bjornerud A (2009) Automatic glioma characterization from dynamic susceptibility contrast imaging: brain tumor segmentation using knowledge-based fuzzy clustering. J Magn Reson Imaging 30:1–10. https://doi.org/10.1002/jmri.21815
23. Emblem KE, Pinho MC, Zollner FG, Due-Tonnessen P, Hald JK, Schad LR, Meling TR, Rapalino O, Bjornerud A (2015) A generic support vector machine model for preoperative glioma survival associations. Radiology 275:228–234. https://doi.org/10.1148/radiol.14140770
24. Fan B, Li HX, Hu Y (2016) An intelligent decision system for intraoperative somatosensory evoked potential monitoring. IEEE Trans Neural Syst Rehabil Eng 24:300–307. https://doi.org/10.1109/tnsre.2015.2477557
25. Focke NK, Yogarajah M, Symms MR, Gruber O, Paulus W, Duncan JS (2012) Automated MR image classification in temporal lobe epilepsy. NeuroImage 59:356–362. https://doi.org/10.1016/j.neuroimage.2011.07.068
26. Foroni R, Giri MG, Gerosa MA, Nicolato A, Piovan E, Zampieri PG, Pasqualin A, Bortolazzi E, Pasoli A, Marchini G et al (1995) A euristic approach to the volume reconstruction of arteriovenous

malformations from biplane angiography. Stereotact Funct Neurosurg 64:134–146
27. Fusco R, Sansone M, Filice S, Carone G, Amato DM, Sansone C, Petrillo A (2016) Pattern recognition approaches for breast cancer DCE-MRI classification: a systematic review. J Med Biol Eng 36: 449–459. https://doi.org/10.1007/s40846-016-0163-7
28. Gazit T, Andelman F, Glikmann-Johnston Y, Gonen T, Solski A, Shapira-Lichter I, Ovadia M, Kipervasser S, Neufeld MY, Fried I et al (2016) Probabilistic machine learning for the evaluation of presurgical language dominance. J Neurosurg 125:1–13. https://doi.org/10.3171/2015.7.jns142568
29. Ghahramani Z (2015) Probabilistic machine learning and artificial intelligence. Nature 521:452–459. https://doi.org/10.1038/nature14541
30. Grigsby J, Kramer RE, Schneiders JL, Gates JR, Brewster Smith W (1998) Predicting outcome of anterior temporal lobectomy using simulated neural networks. Epilepsia 39:61–66
31. Grossman RL, Heath AP, Ferretti V, Varmus HE, Lowy DR, Kibbe WA, Staudt LM (2016) Toward a shared vision for cancer genomic data. N Engl J Med 375:1109–1112. https://doi.org/10.1056/NEJMp1607591
32. Hamzei-Sichani F, Sperling M, Fuertinger S, Sharan A, Simonyan K (2016) Cortical networks high frequency EEG activity patterns in patients undergoing epilepsy surgery. J Neurosurg 124:A1184
33. Hastie T, Tibshirani R, Friedman JH (2009) The elements of statistical learning: data mining, inference, and prediction. Springer, New York
34. Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin PM, Larochelle H (2016) Brain tumor segmentation with deep neural networks. Med Image Anal 35:18–31. https://doi.org/10.1016/j.media.2016.05.004
35. Johnson AE, Ghassemi MM, Nemati S, Niehaus KE, Clifton DA, Clifford GD (2016) Machine learning and decision support in critical care. Proc IEEE Inst Electr Electron Eng 104:444–466. https://doi.org/10.1109/JPROC.2015.2501978
36. Jones TL, Byrnes TJ, Yang G, Howe FA, Bell BA, Barrick TR (2015) Brain tumor classification using the diffusion tensor image segmentation (D-SEG) technique. Neuro-Oncology 17:466–476. https://doi.org/10.1093/neuonc/nou159
37. Jordan MI, Mitchell TM (2015) Machine learning: trends, perspectives, and prospects. Science 349:255–260. https://doi.org/10.1126/science.aaa8415
38. Juan-Albarracín J, Fuster-Garcia E, Manjón JV, Robles M, Aparici F, Martí-Bonmatí L, García-Gómez JM (2015) Automated glioblastoma segmentation based on a multiparametric structured unsupervised classification. PLoS One 10:e0125143
39. Kalkanis SN, Kast RE, Rosenblum ML, Mikkelsen T, Yurgelevic SM, Nelson KM, Raghunathan A, Poisson LM, Auner GW (2014) Raman spectroscopy to distinguish grey matter, necrosis, and glioblastoma multiforme in frozen tissue sections. J Neurooncol 116: 477–485
40. Kanevsky J, Corban J, Gaster R, Kanevsky A, Lin S, Gilardino M (2016) Big data and machine learning in plastic surgery: a new frontier in surgical innovation. Plast Reconstr Surg 137:890e–897e. https://doi.org/10.1097/PRS.0000000000002088
41. Kassahun Y, Yu B, Tibebu AT, Stoyanov D, Giannarou S, Metzen JH, Vander Poorten E (2016) Surgical robotics beyond enhanced dexterity instrumentation: a survey of machine learning techniques and their role in intelligent and autonomous surgical actions. Int J Comput Assist Radiol Surg 11:553–568. https://doi.org/10.1007/s11548-015-1305-z
42. Keihaninejad S, Heckemann RA, Gousias IS, Hajnal JV, Duncan JS, Aljabar P, Rueckert D, Hammers A (2012) Classification and lateralization of temporal lobe epilepsies with and without hippocampal atrophy based on whole-brain automatic MRI segmentation. PLoS One 7:e33096. https://doi.org/10.1371/journal.pone.0033096
43. Kenngott HG, Wagner M, Nickel F, Wekerle AL, Preukschas A, Apitz M, Schulte T, Rempel R, Mietkowski P, Wagner F et al (2015) Computer-assisted abdominal surgery: new technologies. Langenbecks Arch Surg 400:273–281. https://doi.org/10.1007/s00423-015-1289-8
44. Kohli M, Prevedello LM, Filice RW, Geis JR (2017) Implementing machine learning in radiology practice and research. AJR Am J Roentgenol 208:754-760. https://doi.org/10.2214/AJR.16.17224
45. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI (2015) Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J 13:8–17. https://doi.org/10.1016/j.csbj.2014.11.005
46. Lin E, Lane HY (2017) Machine learning and systems genomics approaches for multi-omics data. Biomark Res 5:2. https://doi.org/10.1186/s40364-017-0082-y
47. Madani Tonekaboni SA, Soltan Ghoraie L, Manem VS, Haibe-Kains B (2016) Predictive approaches for drug combination discovery in cancer. Brief Bioinform. https://doi.org/10.1093/bib/bbw104
48. Mariak Z, Swiercz M, Krejza J, Lewko J, Lyson T (2000) Intracranial pressure processing with artificial neural networks: classification of signal properties. Acta Neurochir 142:407–411 **discussion 411-402**
49. Mitchell TJ, Hacker CD, Breshears JD, Szrama NP, Sharma M, Bundy DT, Pahwa M, Corbetta M, Snyder AZ, Shimony JS et al (2013) A novel data-driven approach to preoperative mapping of functional cortex using resting-state functional magnetic resonance imaging. Neurosurgery 73:969–982; **discussion 982-963**. https://doi.org/10.1227/neu.0000000000000141
50. Mitchell TM (1997) Machine learning. McGraw-Hill Science, New York
51. Moghim N, Corne DW (2014) Predicting epileptic seizures in advance. PLoS One 9:e99334. https://doi.org/10.1371/journal.pone.0099334
52. Nowinski WL, Belov D, Benabid AL (2003) An algorithm for rapid calculation of a probabilistic functional atlas of subcortical structures from electrophysiological data collected during functional neurosurgery procedures. NeuroImage 18:143–155
53. Nucci CG, De Bonis P, Mangiola A, Santini P, Sciandrone M, Risi A, Anile C (2016) Intracranial pressure wave morphological classification: automated analysis and clinical validation. Acta Neurochir 158:581–588; **discussion 588**. https://doi.org/10.1007/s00701-015-2672-5
54. Obermeyer Z, Emanuel EJ (2016) Predicting the future—big data, machine learning, and clinical medicine. N Engl J Med 375:1216–1219. https://doi.org/10.1056/NEJMp1606181
55. Orringer D, Ji M, Lewis S, Camelo-Piragua S, Johnson T, Sagher O, Wang A, Maher C, Heth J, Xie X (2015) Visualizing brain tumor infiltration with stimulated Raman scattering microscopy. J Neurosurg 123:A523
56. Saha M, Mukherjee R, Chakraborty C (2016) Computer-aided diagnosis of breast cancer using cytological images: a systematic review. Tissue Cell 48:461–474. https://doi.org/10.1016/j.tice.2016.07.006
57. Schmidt B, Bocklisch SF, Pässler M, Czosnyka M, Schwarze JJ, Klingelhöfer J (2005) Fuzzy pattern classification of hemodynamic data can be used to determine noninvasive intracranial pressure. Acta Neurochir Suppl 95:345–349
58. Senders JT, Arnaout O, Karhade AV, Dasenbrock HH, Gormley WB, Broekman ML, Smith TR (2017) Natural and artificial intelligence in neurosurgery: a systematic review. Neurosurgery. https://doi.org/10.1093/neuros/nyx384
59. Senders JT, Staples PC, Karhade AV, Zaki MM, Gormley WB, Broekman MLD, Smith TR, Arnaout O (2017) Machine learning

and neurosurgical outcome prediction: a systematic review. World Neurosurg. https://doi.org/10.1016/j.wneu.2017.09.149
60. Shamim MS, Enam SA, Qidwai U (2009) Fuzzy logic in neurosurgery: predicting poor outcomes after lumbar disk surgery in 501 consecutive patients. Surg Neurol 72:565–572; **discussion 572**. https://doi.org/10.1016/j.surneu.2009.07.012
61. Shamir R, Duchin Y, Kim JJK, Marmor O, Bergman H, Vitek JL, Sapiro G, Bick AS, Eliyahu R, Eitan R et al (2016) MER validation of a new targeting approach for STN-DBS surgery based on machine-learning and 7T-MRI database (10661). Neuromodulation 19:e67
62. Shi HY, Hwang SL, Lee KT, Lin CL (2013) In-hospital mortality after traumatic brain injury surgery: a nationwide population-based comparison of mortality predictors used in artificial neural network and logistic regression models. J Neurosurg 118:746–752. https://doi.org/10.3171/2013.1.jns121130
63. Shih JJ, Krusienski DJ (2009) Electrocorticography in a brain-computer interface (BCI) paradigm. Epilepsia 50:327
64. Skrobala A (2012) Beam orientation in stereotactic radiosurgery using artificial neural network. Radiother Oncol 103:S220
65. Swiercz M, Mariak Z, Krejza J, Lewko J, Szydlik P (2000) Intracranial pressure processing with artificial neural networks: prediction of ICP trends. Acta Neurochir 142:401–406
66. Taghva A (2010) An automated navigation system for deep brain stimulator placement using hidden Markov models. Neurosurgery 66:108–117; **discussion 117**. https://doi.org/10.1227/01.NEU.0000365369.48392.E8
67. Taghva A (2011) Hidden semi-Markov models in the computerized decoding of microelectrode recording data for deep brain stimulator placement. World Neurosurg 75:758–763.E754
68. Wang J, You X, Wu W, Guillen MR, Cabrerizo M, Sullivan J, Donner E, Bjornson B, Gaillard WD, Adjouadi M (2014) Classification of fMRI patterns—a study of the language network segregation in pediatric localization related epilepsy. Hum Brain Mapp 35:1446–1460. https://doi.org/10.1002/hbm.22265