Artificial Intelligence in the Management of Intracranial Aneurysms: Current Status and Future Perspectives


AJNR Am J Neuroradiol 2020, 41 (3) 373-379
doi: https://doi.org/10.3174/ajnr.A6468
http://www.ajnr.org/content/41/3/373
ABSTRACT

SUMMARY: Intracranial aneurysms with subarachnoid hemorrhage lead to high morbidity and mortality. It is of critical importance to detect aneurysms, identify risk factors of rupture, and predict treatment response of aneurysms to guide clinical interventions. Artificial intelligence has received worldwide attention for its impressive performance in image-based tasks. Artificial intelligence serves as an adjunct to physicians in a series of clinical settings, which substantially improves diagnostic accuracy while reducing physicians’ workload. Computer-assisted diagnosis systems of aneurysms based on MRA and CTA using deep learning have been evaluated, and excellent performances have been reported. Artificial intelligence has also been used in automated morphologic calculation, rupture risk stratification, and outcomes prediction with the implementation of machine learning methods, which have exhibited incremental value. This review summarizes current advances of artificial intelligence in the management of aneurysms, including detection and prediction. The challenges and future directions of clinical implementations of artificial intelligence are briefly discussed.

ABBREVIATIONS: AI = artificial intelligence; AUC = area under the curve; CAD = computer-assisted diagnostics; DL = deep learning; FP = false-positive; ML = machine learning; SVM = support vector machines

Unruptured intracranial aneurysm remains a major public health concern affecting about 3%-7% of the general population.1,2 CTA and MRA are the preferred techniques to identify aneurysms. A steadily increasing number of radiologic examinations are performed for neurologic diagnoses, which require human expertise in image interpretation. However, experienced radiologists are in a relative shortage due to an ever-increasing demand for imaging studies.3 The excess workload imposed on physicians requires them to interpret 1 image every 3–4 seconds within work hours. As a result, uncertainty and inevitable errors when making diagnoses and decisions may arise.4

As the most common cause of nontraumatic SAH, aneurysms are responsible for 85% of SAH cases. The overall annual rupture risk is 0.95% and is associated with a high risk of morbidity and mortality.5 Endovascular and surgical treatments have a 3%-10% risk of stroke or death and may not permanently eliminate the propensity for rupture.6 Therefore, recognition of risk factors and construction of prediction models related to aneurysm initiation, growth, rupture, and assessment of intervention have been popular areas of investigation. However, the management of aneurysms is associated with multiple challenges in all aspects of clinical practice, which cannot be overcome easily by conventional methodology.

Artificial intelligence (AI) describes the development of computer algorithms that simulate human intelligence, which includes learning, reasoning, and self-correction. Brilliant progress has been made that allows machines to automatically interpret complicated data as part of routine clinical care.7 AI is also expected to satisfy the clinician’s need for higher accuracy and better efficacy at all stages of aneurysm management. This review summarizes current applications of AI in the various phases (Fig 1) of intracranial aneurysm management, which facilitates optimized workflows and adds benefit to patient care. In addition, this review discusses the future challenges and directions of AI-based management of aneurysm.
Brief Overview of AI

AI can perform tasks with humanlike intelligence such as pattern recognition, object identification, and problem resolution. AI can make autonomous decisions based on the data collected for training. Similarly, in the realm of medicine, AI can identify a likely diagnosis and select a suitable treatment based on health records or imaging information without any explicit programming. Machine learning (ML) endows AI with the ability to learn and train models to extract and memorize features and related parameters. There are 3 types of ML: supervised (training with specific labels or annotations), unsupervised (training without specific labels and the algorithm clusters data to reveal underlying patterns), and semi- or weakly supervised learning (training with both labeled and unlabeled data to reduce the annotation burden).8

Deep learning (DL) is a subset of machine learning that is end-to-end ML in which it receives input and learns its salient features without explicit examples.9 There are 2 key requirements for ML: 1) Data are relevant and detailed enough for answering the clinical questions; and 2) a computational ML technique is appropriate for the type, amount, and complexity of the available data. Finally, it is necessary to further validate the usefulness of ML models in real-world practice.10

Traditional ML algorithms predefine engineered features that can describe the patterns inherent in regions of interest with explicit parameters based on expert knowledge. Support vector machines (SVM) and random forests are classic ML techniques. These algorithms have been successfully used in AI studies.11,12 Compared with traditional ML algorithms, DL uses an artificial neural network that extracts features from images automatically to create its own filters, called feature maps (independent of human input) and memorizes visual patterns with highest frequency. There are convolution layers, pooling layers, fully connected layers, and normalization layers; the pooling reduces the number of parameters and reduces overfitting. Once a final output is made that is compared with ground truth, the model will reweight the inputs until the best performance is reached, which can be repeated many times (called epochs). For feature-extraction and selection, DL algorithms do not need to predefine features, and the feature representations can be directly learned by navigating the data space. Most DL networks fundamentally build on some basic and similar neural network blocks or layers based on DL attempts to model high-level abstractions.13 Therefore, the extracted complex features are then fed through the last layer of the network for the target, such as classification or prediction. Convolutional neural networks are typical types of DL architecture, which consist of characteristic layers of convolutional operations on the input images to extract abstract features.13

Computer-Assisted Diagnosis of Aneurysms

CTA and TOF-MRA are widely available techniques for detection. However, very small aneurysms are often missed.14 Computer-assisted diagnostics (CAD) acts principally as an internal second
opinion and improves reading accuracy of physicians. Algorithms have been devised for automatic detections based on MRA/CTA studies using conventional-style and DL methods (On-line Table).

Conventional-Style CAD Systems. The conventional-style CAD systems were based on presupplied characteristics or imaging features, such as vessel curvature, thresholding, or a region-growing algorithm. The first CAD system geared toward aneurysms reported in the literature was developed by Arimura et al, in 2004, which consisted of multiple gray-level thresholding techniques. It showed 100% sensitivity with 2.4 false-positives (FPs) per patient based on a leave-one-out-by-patient test method. In subsequent validation studies, Hirai et al and Kakeda et al found that the sensitivity was 100% and 84%, respectively. However, this algorithm was not fully automated, nor could it detect small or fusiform type of aneurysms. In addition, the aneurysms in these studies were not verified by DSA as an outside reference standard. By using DSA as the reference, Yang et al developed a more automated CAD algorithm by combining 2 complementary techniques: 1) automatic intracranial artery segmentation and 2) detection of points of interest from the segmented vessels. They achieved a sensitivity of 95% with up to 9 additional FP detection marks. However, the sensitivity was lower for small aneurysms (<5 mm). The conventional-style CAD schemes relied on a similar rules-based approach in which prior domain knowledge was incorporated into hand-crafted features before using ML techniques as a classifier. The major limitation is that many false-positive results are found in bending or branching portions of vessels. Multiple methods have been proposed such as an ellipsoid convex enhancement filter to selectively enhance aneurysms while reducing FPs, but the number of FPs remains high.

Deep Learning CAD Systems. More recently, DL-based CAD systems have been developed for aneurysm detection on the basis of MRA. For example, Nakao et al used a 2D convolutional neural network to detect aneurysms and reported a sensitivity of >90% in a single-center study. Similarly, an open-source neural network has also been applied by using 2D MIPs or original image data. However, generalization of these studies requires further validation. The work by Ueda et al subsequently improved the sophistication of a DL methodology with a sensitivity of 91% and 93% for the internal and external test datasets, respectively. The external dataset contained images from 4 separate institutions under different environments and MR imaging unit manufacturers, configurations, and field strengths; this feature highlights the rigor and general applicability of the model. However, the study did not have any cases negative for aneurysms. Furthermore, only 74 aneurysms and studies exclusively acquired on Siemens imaging systems were integrated in the external test dataset. Additionally, a high rate of FPs may potentially reduce the enthusiasm of radiologists for the use of this system if distinguishing true-positive aneurysms from FPs becomes too menial.

CTA images have been studied by combining a neural network segmentation model (the HeadXNet model) to augment diagnostic performance in the detection of aneurysms. In this study, the clinicians showed significant increases in sensitivity, accuracy, and interrater agreement when they were augmented with the model. However, the lack of a reference standard and external data verification, as well as the focus only on nonruptured aneurysms of >3 mm, limited the generalization and further application of the model.

Al in the Prediction of Intracranial Aneurysm Rupture Risk

Automated Morphologic Analysis. Identifying the high-risk morphologic features of saccular aneurysms is very important for rupture-risk stratification and treatment decision. Aneurysm size and shape are regarded as the most important criteria. In clinical practice, the size is routinely measured manually by physicians on 2D/3D projections. However, manual measurements have inherent limitations of subjectivity and inconsistency, which cause intra- and interobserver variations and cannot capture the complex geometric features of aneurysms. Researchers have introduced several computerized procedures to make morphology assessment more objective and consistent. The detection of the neck plane is the key point for multiple subsequent operations. Larrabide et al determined the aneurysm neck based on the topology analysis of the vasculature skeleton and the concepts of deformable cylinders. They can automatically isolate the sac of an aneurysm, reduce interobserver variability, and avoid the bias between the observers. Automatically derived geometric indices were often large, irrespective of segmentation method or operators.

For rapid assessment in the clinical setting, Xiang et al devised an image-based vascular analysis toolkit named AView (https://www.eng.buffalo.edu/Research/Hemo/AView.html) to perform automatic computation of morphologic parameters. AView provided a relatively accurate measurement with an average size error of 0.56% and volume error of 2.1% morphologically. This toolkit also enabled increased consistency in morphologic measurement among operators by 62% in size and 82% in neck diameter measurements, which could help potentially avoid inappropriate clinical decisions. In the real-world scenario, Rajabzadeh-Oghaz et al from the same group, tested the algorithm on 39 aneurysms and found that the computer-assisted 3D approach can lead to a more accurate and consistent determination of aneurysm size and neck diameter. Besides, DL methods such as the convolutional neural network model developed by Stember et al can also be used to automatically analyze morphologic indices. However, there is much work to be done before routine clinical application of these technologies is realized.

Automated Calculation of Hemodynamics. Hemodynamics is currently deemed as an important factor for aneurysm formation and rupture risk. Modern imaging modalities are adequate for the application of computational fluid dynamics modeling. However, complex procedures are time-consuming and demand substantial human interaction, resulting in limited application in real-time clinical practice, which requires automated tools to execute analysis. Seo et al developed a highly automated method to execute a computational process with direct use of a voxelized contrast information from 3D angiograms to construct a level-set-based computational “mask” for a hemodynamic simulation. By testing the method in 7 patient-specific cases against the results of manual evaluation by an experienced neurosurgeon, they found that their proposed algorithm was capable of identifying the lesion and connected vessels for various types of aneurysms. The simulation results presented by the algorithm, which include the values and distribution of wall shear stress, were in line with previous computational studies. However, this study was only hypothesis-generating and requires further refinement and validation. Considering that hemodynamic parameters are critical for the development and rupture risk, such approaches bode rather well for clinical utility due to their automation and diminishing need for human interaction.

AI-Based Prediction of Rupture Risk. An increasing number of unruptured aneurysms are detected with the growing use of advanced imaging techniques. However, we are now confronted with the dilemma of making clinical decisions regarding treatment of unruptured aneurysms, because the risk of treatment-related fatality is relatively high, while the rupture risk is low. Predicting rupture risk of aneurysms is challenging and ML is expected to mitigate this problem. Liu et al adopted 17 parameters as inputs to a 2-layer feed-forward artificial neural network aimed at predicting the rupture status of anterior communicating artery aneurysms and found an excellent performance. However, this study included only 1 single-center population and used an imbalanced number of samples between ruptured (n = 540) and unruptured aneurysms (n = 54). The instability of aneurysms is considered a rupture risk. With this knowledge, Liu et al applied radiomics tools to extract morphologic features to predict stability and found that flatness was the most important parameter; the area under the curve (AUC) in the testing set reached 0.729 when only flatness was used to predict aneurysm stability, implying the usefulness of radiomics-derived morphologic features for aneurysm rupture risk.

With the introduction of ML methods, it is interesting to understand the distinctive performances of different ML statistical learning approaches. Detmer et al and Silva et al worked at predicting aneurysm rupture status. They trained several ML methods, including SVM and random forests classifiers. Detmer et al found that multilayer perceptron had the best performance with an AUC of 0.826 (95% CI, 0.768, 0.883) in the test set; important variables included aneurysm location, mean surface curvature, and maximum flow velocity. Silva et al found random forests to have the best performance. In their work, aneurysm location and size were the 2 features that contributed most significantly to the efficacy of the model. This difference may contribute to the uniqueness of input variants and candidate ML methods.

AI in the Prediction of Aneurysm Complications

Prediction of Complications of SAH. Delayed cerebral ischemia, vasospasm, and cerebral infarction are among the complications of aneurysm rupture; several studies have explored the applications of ML methods to predict at least one of them. Dumont et al developed a proof-of-concept artificial neural network prediction model of symptomatic cerebral vasospasm and found the artificial neural network–based model had a better predictive value (AUC of 0.960) than 2 multiple logistic regression models (AUC = 0.933 and 0.897) developed by Adams et al and Qureshi et al, respectively. Further validation provided an excellent performance in a markedly distinct geographic population setting of southern Arizona with a prospective use of the artificial neural network predictive model. However, the artificial neural network did not incorporate the timing of symptomatic cerebral vasospasm onset and was not validated in a larger-scale population. Another study had similar findings: ML methods (SVM, random forests, and multilayer perceptron) have a higher performance than logistic regression models in the prediction of delayed cerebral ischemia.

Multiple-task AI has also been studied. Tanioka et al used random forests to construct early prediction models of delayed cerebral ischemia, angiographic vasospasm, and cerebral infarction development with clinical variables and matricellular proteins at post-onset days 1–3. Three such proteins have been reported to be relevant to delayed cerebral ischemia: osteopontin, periostin, and galectin-3. The prediction accuracies of the 3 conditions were 95.1%, 78.1%, and 83.8%, respectively. The random forest models found that osteopontin and galectin-3 were among the top 3 most important variables. These novel studies in the assessment of complications of aneurysms have demonstrated excellent performances using ML methods.

Another interesting application is the use of clinical data and CT perfusion from hospital admissions to predict outcomes of aneurysmal SAH. A random forest model was trained to predict dichotomized mRS (≤2 and >2), and the accuracy was 84.4% in the training folds and 70.9% in the validation folds. However, this study had a small population size and therefore cannot be introduced into clinical practice to practically benefit those who have SAH.

Prediction of Treatment Outcomes. Besides aneurysmal SAH, the outcomes of aneurysm treatment have also been explored by ML. For example, the flow diverter has emerged in recent years as one of the endovascular treatments of aneurysms and is particularly suitable for treating wide-neck and intractable aneurysms with unusual morphologies. However, 25% of treated patients are at high risk for thromboembolism formation and aneurysm rerupture. Palwai et al compared 4 supervised ML algorithms (logistic regression, SVM, K-nearest neighbor, and neural networks) to predict 6-month outcomes of flow diverter–treated aneurysms and found that neural network (AUC = 0.967) performed the best during training; G-SVM (Gaussian-SVM) and neural network had 90%
prediction accuracies in the testing cohort. Although a major concern is the absence of an external validation, it is imperative to develop models that help clinicians choose flow-diverter placement in the appropriate patients.

**Limitations and Challenges of AI on Aneurysms**

Many investigators have endeavored to develop more intelligent and automated methods to improve the process and streamline the management of aneurysms in clinical practice. Complete management of aneurysms includes detection, rupture risk, complications prediction, treatment strategies selection, and risk of recurrence evaluation. AI has been involved in almost all consecutive steps; however, the results are not entirely satisfactory and have some limitations and challenges. First, a systematic review showed no better performance of ML compared with traditional logistic regression in predicting adverse clinical outcomes. Model validation procedures are often not conducted or well-interpreted; thus, a fair comparison in the domain of real-world case studies is still urgently needed. Second, even though DL methods showed promising potential, there remains considerable debate about the time required to train DL. Cost-effectiveness assessment of the high expense of imaging data storage and use of graphics processing units should also be considered. In addition, stakeholders may be responsible for undertaking cost-effectiveness analyses to determine the impact on health care economics, which requires large-scale feasibility studies to evaluate whether unintended consequences may be incurred following the widespread implementation of AI methods. Third, current datasets used for training and validation in aneurysm studies are often small, frequently from only 1 institution, and usually lack external validation within different populations, imaging devices, platforms, and institutions, resulting in algorithm overfitting.
Future Perspectives of AI in Intracranial Aneurysms

AI should have a great potential in the management of aneurysms in the future. Potential imaging-based AI applications (Fig 2) mainly contain 4 aspects: quantification, notification tools, diagnostics, and risk prediction for therapy. For aneurysm management, there are various opportunities for AI, including the following: 1) an automated and reliable prescreening triage system to allow radiologists and emergency physicians to shift their focus to patients at higher risk for adverse events; 2) automated detection and intelligent outcome prediction; 3) subsequent prediction of treatment strategies such as clipping, embolization, or follow-up; and 4) automated and intelligent detection of de novo aneurysm formation or recurrence after treatment and the prediction of rerupture risk.

Researchers should intensify their efforts toward the development of advanced DL algorithms to resolve complex problems, such as extracting information on inflammation of vessel walls from high-resolution vessel wall MR imaging, which has been a nascent area of research in recent years. Additionally, AI tools that can perform multiple tasks, instead of those directed at an isolated problem, are in need. Examples include algorithms that can automatically detect and classify the whole variety of cerebrovascular diseases, including multiple entities like aneurysms, AVMs, and Moyamoya disease. Advanced network structures such as convolutional residual networks, active learning, one-shot learning, and generative adversarial networks may help settle these issues. These areas are newly emerging and require further investigation.

In our perspective, a successful AI tool for aneurysms is expected to be an integrated toolkit that can intelligently and simultaneously navigate multiple tasks, ranging from detection to prediction (Fig 3). In future work, comprehensive multimodality imaging is worth investigating and is expected to learn and exploit feature representation of images more effectively. In addition, human-in-the-loop AI solutions are quickly emerging concepts toward clinical implementation, which can improve the expert-model performance and act as a gatekeeper for clinical decisions.

Most recently published AI studies have not delineated robust validations of clinical performance and generalizability, including in the area of intracranial aneurysms. Kim et al recommended 4 criteria for validating the clinical performance of AI algorithms in real-world practice: 1) the external validation must be obtained; 2) using a diagnostic cohort study; 3) arise from multiple institutions; and 4) be performed in a prospective manner. A successful AI tool for aneurysm management must be completely reliable in distinguishing true-positive cases with high confidence, which requires an exceedingly high number of annotated imaging studies, before we can expect widespread implementation in real-world practice.

CONCLUSIONS

Recent evidence shows that AI, especially deep learning, is rapidly becoming a promising aid in the management of aneurysm. AI is capable of detecting and evaluating rupture risk, triaging clinical therapy strategies, and predicting treatment outcomes. Although we have not quite yet reached the threshold needed for routine clinical application, we believe that AI is capable of solving these issues in a patient-centric manner but will require advancing network structures and more sophisticated validation processes.

Disclosures: Zhao Shi—RELATED: Grant: Key Projects of the National Natural Science Foundation of China [81830057 for LI].* U. Joseph Schoepf—UNRELATED: Other: Dr Schoepf has received institutional research support and/or honoraria for speaking and consulting from Astellas Pharma, Bayer, Bracco, Ekedolmaging, GE Healthcare, Guerbet, HeartFlow, and Siemens.* Xia Li—RELATED: Key Projects of the National Natural Science Foundation of China [81830057 for LI].* Long Jiang—RELATED: Grant: National Key Research and Development Program of China; Key Projects of the National Natural Science Foundation of China. Comments: This work was supported by the National Key Research and Development Program of China [2017YFC0113400 for LI] and Key Projects of the National Natural Science Foundation of China [81830057 for LI].* Money paid to the institution.

REFERENCES