Proceedings of CONF-CDS 2026 Symposium: Machine Learning and Neural Network Applications in Engineering
DOI: 10.54254/2755-2721/2026.AST36870

Artificial Intelligence-Assisted Detection of Intracranial Aneurysms on CTA and TOF-MRA: Small Lesions, Error Patterns, and Human–AI Collaboration

Qingyang Shen

Faculty of Engineering, The University of Hong Kong SAR, Hong Kong, China
m18721495571@163.com

Abstract. Accurate detection of intracranial aneurysms is essential for further imaging assessment, specialist referral, and clinical management. Computed tomography angiography and time-of-flight magnetic resonance angiography are commonly used non-invasive methods for cerebrovascular examination. However, small aneurysms may go unnoticed due to their limited size, complex vascular anatomy, and imaging artifacts. Artificial intelligence (AI) can automatically identify suspicious lesions and may improve reader performance. This review summarizes studies published between 2022 and 2026 on AI-assisted detection of intracranial aneurysms using computed tomography angiography and time-of-flight magnetic resonance angiography. It focuses on the detection performance for small lesions, common false-positive and false-negative patterns, and the clinical value of human–AI collaboration. Current evidence suggests that artificial intelligence provides satisfactory overall detection performance. In some studies, it also improved clinicians' detection sensitivity and reading efficiency. However, aneurysms measuring ≤3 mm and some atypical lesions remain difficult to detect. Increasing model sensitivity may also result in more false-positive markings. Incorrect artificial intelligence suggestions may further lead to automation bias. At present, AI is more suitable as a lesion-alert tool and a second reader than as a replacement for clinicians in final diagnosis. Future research should investigate the causes of detection errors in relation to dataset composition, model design, external validation, and human–AI interaction. Diagnostic performance, false-positive burden, and effects on patient management should also be evaluated in real clinical settings.

Keywords: Artificial intelligence, Intracranial aneurysm, Computed tomography angiography, Time-of-flight magnetic resonance angiography, Human–AI collaboration

1. Introduction

An intracranial aneurysm is a localized abnormal dilation of the wall of a cerebral artery. Accurate detection of intracranial aneurysms is essential for further imaging assessment, follow-up management, and evaluation by neurosurgical or neurointerventional specialists. Cerebral arteries, aneurysms, and their connections to surrounding structures can be seen by computed tomography angiography (CTA), which can be completed quickly. Time-of-flight magnetic resonance

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.AST36870

angiography (TOF-MRA) usually does not require contrast agent administration and does not involve ionizing radiation, making it suitable for non-invasive examination and imaging follow-up.

Artificial intelligence (AI), particularly deep learning, can automatically extract vascular and lesion features from CTA and TOF-MRA images. It can then mark suspicious aneurysm regions. A meta-analysis published in 2026 showed that deep learning achieved good overall diagnostic performance for detecting intracranial aneurysms on CTA, MRA, and digital subtraction angiography. However, substantial differences were found among studies in imaging modalities, case composition, validation strategies, and evaluation metrics [1]. A methodological scoping review further noted that many existing studies used retrospective designs. They were also affected by potential case-selection bias, insufficient external validation, inconsistent reference standards, and incomplete reporting of key performance measures [2].

Small intracranial aneurysms remain a major challenge for AI-assisted detection. Smaller lesions are more likely to be confused with vascular bifurcations, localized vascular dilatations, and imaging artifacts. A systematic review and meta-analysis focused on CTA showed that models generally performed less effectively for aneurysms measuring <3 mm than for aneurysms overall [3]. In addition, some studies merely reported the area under the curve (AUC) or total sensitivity. Lesion-level sensitivity, patient-level specificity, and false positives per case were not reported concurrently. Consequently, these studies may not fully reflect the clinical value of the models.

Previous reviews have mainly focused on the overall diagnostic performance of AI models. Less attention has been given to lesion-size subgroups, specific error patterns, and the influence of incorrect AI suggestions on clinicians' decisions. Based on studies published between 2022 and 2026, this review examines the clinical value of AI-assisted intracranial aneurysm detection on CTA and TOF-MRA from three perspectives: small lesions, error patterns, and human–AI collaboration. This analysis may provide useful information for engineers seeking to improve model design, clinicians using AI-assisted systems, and researchers developing multicenter validation studies and human–AI workflows.

2. AI detection performance for small intracranial aneurysms

What constitutes "small" or "tiny" aneurysms can differ between studies. This review kept the original definitions from the studies and focused on lesions 3 mm or less. The detection performance for the patient level and lesion level are different outcomes, and the number of false positives per case and lesion subgroups should be considered in model evaluation.

2.1. CTA-based detection

Study of AI-based CTA detection has advanced from development of single-center algorithm to multicenter external validation, multi-reader study, and prospective clinical application. There is emerging evidence that AI can be valuable for detection of intracranial aneurysms on CTA. However, model performance is affected by several factors, including lesion size, vascular location, image quality, scanning protocols, and test-set composition.

Lesion size is an important factor affecting the performance of CTA models. Liu et al. developed a three-dimensional convolutional neural network that achieved an overall lesion-level sensitivity of 92.3% in the test set. However, its sensitivity for aneurysms measuring ≤3 mm was only 66.7% [4]. The test set included only 10 aneurysm-positive CTA examinations. Therefore, these results cannot represent the overall performance of the field. Nevertheless, they show that high overall sensitivity may conceal reduced performance for tiny lesions. A systematic review and meta-analysis of CTA

DOI: 10.54254/2755-2721/2026.AST36870

studies reached a similar conclusion. The average detection performance for lesions measuring <3 mm was lower than that for larger lesions [3, 4].

Multicenter validation can provide a more reliable assessment of clinical applicability. Hu et al. used 16,546 head and neck CTA examinations from eight hospitals to develop a model. They subsequently evaluated the model through an external test confirmed by digital subtraction angiography, a multi-reader crossover study, and a prospective clinical study involving 1,562 examinations. AI assistance improved clinicians' diagnostic performance at both patient and lesion levels. These findings suggest that CTA-based aneurysm AI is gradually progressing from technical testing to evaluation within clinical workflows [5].

However, strong performance in development datasets may not be fully reproduced across different hospitals and patient populations. Pettersson et al. conducted prospective internal testing and international external testing. The model achieved a lesion-level sensitivity of 83.5% in the international external test. In addition, the model detected a considerable proportion of lesions missed by radiologists at the external center. This finding supports its potential value as a second reader [6]. The study also shows that model generalizability should be assessed using data from different countries, hospitals, devices, and patient populations.

Studies of human–AI collaboration on CTA further indicate that the practical value of a model cannot be determined solely by its independent performance. Zhuo et al. trained a model using data from 3,829 patients across 11 clinical centers and tested it on 484 patients from three institutions. With AI assistance, both junior and senior radiologists achieved higher patient-level and lesion-level sensitivity. Image-reading and post-processing times were also reduced [7]. However, these results were obtained using a particular software system, dataset, and reader-study design. As of yet, they cannot be directly generalized to every therapeutic context.

Overall, AI has good potential as an assistive tool for detecting intracranial aneurysms on CTA. Nevertheless, aneurysms measuring ≤3 mm remain more likely to be missed. Compared with internal testing alone, multicenter external validation and human–AI reader studies provide a more realistic assessment of clinical value.

2.2. TOF-MRA-based detection

TOF-MRA displays cerebral vessels by using magnetic resonance signals generated by flowing blood. It is appropriate for non-invasive inspection and imaging follow-up and typically doesn't require the administration of contrast agents. Slow blood flow, turbulent flow, vessel overlap, spatial resolution, and motion artifacts may affect TOF-MRA image quality. Research has mainly focused on small-lesion detection, false-positive reduction, and clinical reader assistance.

Terasaki et al. used 559 TOF-MRA cases containing 732 aneurysms from three institutions. They compared two-dimensional, three-dimensional, and multidimensional convolutional neural networks that combined two-dimensional and three-dimensional information. The multidimensional model reached a maximal sensitivity of 89.1% in external testing and produced 4.2 false positives per case. The false positive rate is lower than those of 2D and 3D models [8]. It can be inferred that combining planar and volumetric imaging information may also lead to better detection. However, in each testing set exams still contain incorrect marking lesions requiring further clinical review.

External validation in an independent hospital setting can show restrictions not detectable by the data used for development. Lehnen et al. examined commercial software for AI with TOF-MRA images from 191 subjects. The software achieved an overall sensitivity of 72.6% and a specificity of 87.2%. It performed well for non-thrombosed saccular aneurysms measuring >5 mm but had lower

DOI: 10.54254/2755-2721/2026.AST36870

detection performance for fusiform and thrombosed aneurysms [9]. This suggests that strong performance for regular lesions does not guarantee reliability for atypical cases.

A trade-off exists between reducing false positives and preserving sensitivity. Kuwabara et al. analyzed 10,000 consecutive brain screening examinations. After adjusting algorithms, the false positive ratio decreased from 2.06 to 0.99 per case while sensitivity reduced from 96.5% to 90% [10]. This may imply that reduction in false positive may entail missing true lesions to a certain extent so that we should not look at only one performance indicator.

Performance variation in detection according to the size of aneurysm is shown which emphasizes that lesion size could affect the performance of detection. Li et al. conducted an external clinical validation study involving 510 participants. The AI platform showed overall sensitivity of 91.63% and specificity of 92.20%. Sensitivity for lesions measuring lower than 3 mm was 87.80%, which was relatively low compared with large lesions. AI assistance also increased the average sensitivity of clinicians at different experience levels [11].

In 2026, Kim et al. developed a dual-stream detection framework guided by vascular anatomical landmarks. They used 1,055 TOF-MRA examinations to train and test the model. Lesion-level sensitivity was 0.87 and 0.82 in the two internal test sets, while false positives per case were 1.23 and 1.17, respectively. However, detection performance remained lower for lesions measuring ≤3 mm [12]. The study included images from different manufacturers and magnetic field strengths. Yet, it primarily used internal testing, and further independent external validation is needed.

All things considered, AI has a decent chance of assisting intracranial aneurysm detection on TOF-MRA. However, reported performance varies substantially among studies. Small lesions, atypical morphology, device-related differences, and false-positive burden remain important barriers to clinical application.

3. Detection error patterns

3.1. False-positive findings

A false positive occurs when AI incorrectly marks an imaging structure that is not an aneurysm as an aneurysm. Based on current evidence and relevant imaging characteristics, false-positive findings on CTA may involve arterial, venous, extracranial, and other non-vascular structures. On TOF-MRA, false-positive findings may be related to local vascular morphology, abnormal flow signals, and overlapping structures.

False positives not only affect statistical performance but may also increase clinical workload. Clinicians need to review each AI-generated candidate marker. Many false-positive findings may offset the time saved through AI assistance. Repeated incorrect alerts may also cause alert fatigue and reduce clinicians' trust in the system. If an incorrect marker is not properly identified, it may lead to unnecessary imaging examinations, specialist referrals, or follow-up recommendations.

Current studies have attempted to reduce false positives through algorithm adjustment, integration of two-dimensional and three-dimensional information, and anatomy-based post-processing. Terasaki et al.'s the multidimensional model maintained relatively high sensitivity while reducing false-positive findings on TOF-MRA [8]. Kuwabara et al. adjusted an algorithm using consecutive clinical data. This adjustment substantially reduced false positives per case, although sensitivity also decreased [10].

Kim et al. further incorporated brain-region and arterial–venous anatomical information into the post-processing stage of a CTA model. In external test datasets, anatomy-based post-processing reduced incorrect markings in extracranial regions and venous structures. Some strategies

DOI: 10.54254/2755-2721/2026.AST36870

substantially reduced false positives without removing true-positive lesions [13]. This study suggests that targeted processing based on the anatomical sources of false positives may be more interpretable than simply adjusting a model threshold.

3.2. False-negative findings and performance trade-offs

A false negative occurs when an aneurysm is present on an image but is not marked by the AI system. Lesions measuring ≤3 mm are among the more common types of missed aneurysms [4, 11, 12]. Other factors that may increase the risk of missed detection include fusiform or thrombosed morphology, complex vascular anatomy, certain anatomical locations, poor image quality, and incomplete lesion-level detection in patients with multiple aneurysms. Patterns of missed detection may be related to the composition of training data. If the training dataset mainly contains typical saccular aneurysms, the algorithm may have difficulty identifying fusiform, thrombosed, or other atypical lesions. Lehnen et al. found that commercial AI software performed well for typical and larger saccular aneurysms. However, its sensitivity was substantially lower for fusiform and thrombosed lesions [9].

A reasonable balance is required between sensitivity and false-positive findings. Lower detection thresholds allow the model to mark more suspicious regions and may reduce missed lesions. However, they also increase the clinical review burden. Higher thresholds can reduce false-positive findings but may result in missed small aneurysms. Therefore, AI systems should not report only overall accuracy, AUC, or sensitivity. They should also report lesion-level sensitivity, patient-level specificity, lesion-size subgroup results, false positives per case, and 95% confidence intervals.

4. Human–AI collaboration

4.1. Diagnostic assistance

AI-assisted reading mainly helps clinicians review potentially missed lesions by highlighting suspicious regions. Several studies have shown that AI assistance can improve detection sensitivity for some clinicians in specific datasets and reader experiments. It may also improve inter-reader consistency and reading efficiency. The multicenter study found that AI assistance improved patient-level and lesion-level diagnostic performance on CTA [5]. Zhuo et al. further reported improved sensitivity among both junior and senior radiologists. Image-reading and post-processing times also decreased [7]. Li et al.'s external TOF-MRA validation study also found that clinicians at different experience levels achieved higher detection sensitivity with AI assistance [11].

However, the independent performance of AI may not exceed that of specialist clinicians. Adamchic et al. evaluated commercial AI software using 500 clinical TOF-MRA examinations. The sensitivities of neuroradiologists and AI alone were 92.5% and 72.6%, respectively. Combining their detection results improved the reliability of lesion identification and reduced the average analysis time by approximately 19 seconds [14]. However, the reference standard was based on lesions jointly identified by the AI system and neuroradiologists. It differed from studies using digital subtraction angiography or an independent expert consensus. Therefore, the numerical results should not be directly compared with those of other studies.

These findings support the use of AI as an assistive tool or second reader in medical imaging. AI can help clinicians review potentially overlooked regions. However, its candidate markings must still be assessed against original images, vascular anatomy, and the patient's clinical information.

DOI: 10.54254/2755-2721/2026.AST36870

4.2. Automation bias

Automation bias occurs when clinicians accept incorrect suggestions because they rely excessively on an automated system. It may also occur when clinicians become less alert to genuine lesions because the AI system does not issue a warning. Although AI assistance may improve detection sensitivity, excessive reliance on AI and abandonment of independent reading may cause algorithmic errors to become clinical diagnostic errors.

Nine radiologists with varying degrees of expertise were asked by Kim et al. to evaluate 20 TOF-MRA exams both with and without AI support. Incorrect positive AI markings increased the readers' suspicion that an aneurysm was present at the corresponding location. Less experienced readers were also more likely to recommend more intensive follow-up. Although AI assistance reduced reading time, it also introduced a risk of automation bias [15]. The study had a small sample size, and false-positive AI markers were intentionally included in some cases. Therefore, it indicates that automation bias may occur but cannot estimate its actual prevalence in clinical practice.

Based on the current evidence, a relatively cautious human–AI workflow should begin with an independent assessment of the original images by the clinician. AI-generated candidate markers should then be reviewed. Suspicious regions identified by AI should be verified using the original images and multiplanar reconstructions. Normal vascular structures, veins, and imaging artifacts should be carefully excluded. Clinicians should still complete a systematic review even when AI identifies no abnormality. The final imaging assessment and decision regarding further examination should be made by the clinician according to the patient's individual circumstances. As shown in Figure 1, clinicians should complete an independent image assessment before viewing AI-generated candidate markers. This workflow is a narrative recommendation based on current studies of human–AI collaboration and automation bias.

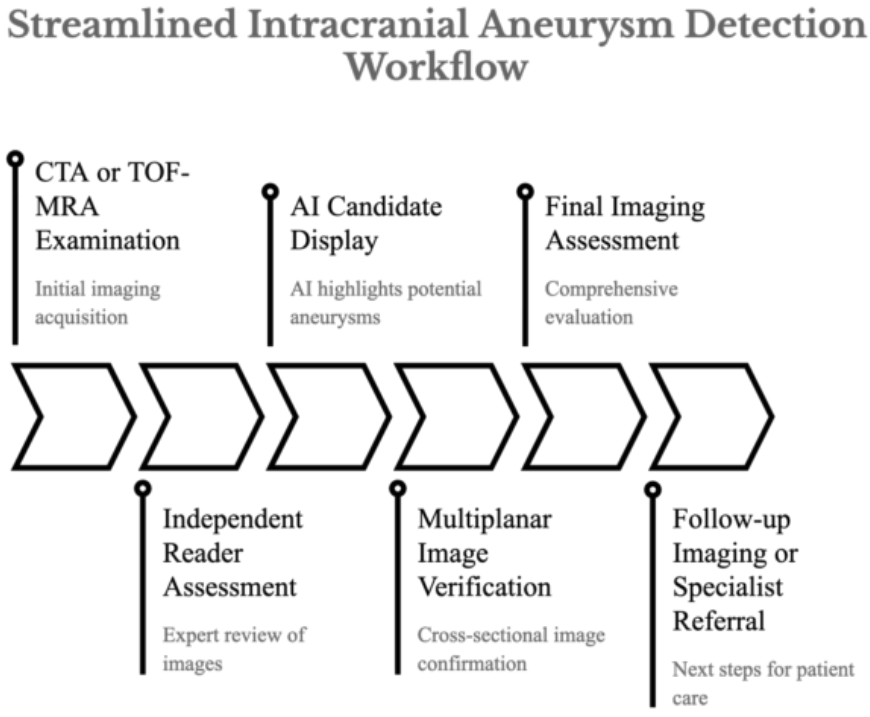

Figure 1. Proposed human–AI workflow for intracranial aneurysm detection [5, 7, 11, 14, 15]

DOI: 10.54254/2755-2721/2026.AST36870

5. Causes of detection errors and improvement strategies

5.1. Causes of detection errors

False-positive and false-negative findings in AI systems do not result only from insufficient model accuracy. Instead, they arise from the combined effects of image quality, lesion characteristics, training data, and model decision-making processes.

First, small aneurysms provide limited imaging information. Aneurysms measuring ≤3 mm occupy relatively few pixels or voxels on CTA or TOF-MRA. Their boundaries may therefore be affected by spatial resolution and partial-volume effects. Information relating to small lesions may be further reduced during image reconstruction, down sampling, or feature compression. As a result, a model that achieves high sensitivity for larger lesions may not detect tiny lesions consistently [3, 4, 11, 12]. CTA and TOF-MRA also use different imaging mechanisms. Vascular anatomy, blood-flow conditions, and imaging resolution may all influence the visualization of small aneurysms [16].

Second, aneurysms may have a high level of morphological similarity to surrounding normal anatomical structures. Vascular bifurcations, tortuous vessels, and some venous structures may also appear as round or saccular regions on local images, leading to false-positive findings. CTA may be affected by extracranial vessels, veins, and non-vascular structures. In comparison, TOF-MRA signals depend on blood flow. Slow flow, turbulent flow, and changes in flow direction may alter vascular appearance [8, 13]. Therefore, local morphological information alone may be insufficient to distinguish true aneurysms from normal structures accurately.

Third, the composition of training data affects the lesion types that a model can learn to identify. If the training dataset is heavily populated by common and large saccular aneurysms and contains few small, fusiform, thrombosed, or special condition lesions, the model is expected to perform well in the case of common cases and not as well for rare cases [9]. Other studies might have excluded low-quality images, previously treated cases, and patients with cerebrovascular diseases other than aneurysms. Therefore, the patient data used for training are not representative of the real-world clinical population [2].

Finally, the model threshold directly affects the type of detection error. A low detection threshold will have many suspicious regions and increase sensitivity at the expense of specificity. A high detection threshold will reduce false-positive findings at the expense of sensitivity or by missing very small or special condition aneurysms. The study by Kuwabara et al., in which false positives decreased while sensitivity also declined, illustrates this performance trade-off [10]. False-positive and false-negative findings therefore cannot be addressed completely in isolation. A reasonable balance should be selected according to the intended clinical use.

5.2. Improvement strategies

Improvements at the data level form the basis for reducing detection errors. Future datasets should include more aneurysms measuring ≤3 mm, as well as fusiform, thrombosed, and unusually located aneurysms. They should also include difficult negative cases, such as vascular bifurcations, venous structures, normal vascular variants, and imaging artifacts. These cases are more likely to generate false-positive findings. Allowing models to learn from both typical lesions and confusing normal structures may reduce their dependence on individual local morphological features. Consecutive clinical cases and multicenter data may also reduce case-selection bias and provide a closer representation of real clinical practice [2, 6].

DOI: 10.54254/2755-2721/2026.AST36870

Model development should place greater emphasis on lesion scale and vascular anatomical information. Multiscale and three-dimensional models can analyze both local lesion details and their relationships with surrounding vessels. This approach may improve the identification of small lesions. The methods by Terasaki et al. and Kim et al. highlight the advantages of integrating two-dimensional and three-dimensional data, utilizing a dual-stream model informed by vascular landmarks [8, 12]. In addition, incorporating brain-region, arterial, and venous segmentation into post-processing may filter out false-positive findings and improve model interpretability [13].

Validation and evaluation methods also require further standardization. Future studies should report patient-level and lesion-level results separately. Aneurysms should be grouped as ≤3 mm, >3–5 mm, and >5 mm. In addition to sensitivity, specificity, and AUC, studies should report false positives per case, 95% confidence intervals, and detection performance for lesions with different locations and morphologies. Before clinical use, models should also undergo external validation across hospitals, devices, and patient populations. This process is necessary to determine whether their performance can be maintained beyond the development dataset [5, 6].

At the level of human–AI interaction, AI systems should provide more than a simple positive or negative result. When displaying AI results, they should display the location of candidate lesions, prediction confidence, and relevant anatomical information [13, 17]. Before viewing the AI results, clinicians should independently review the original CTA or TOF-MRA images. They should also receive appropriate training on false-positive findings, false-negative findings, and automation bias. For low-confidence or atypical lesions, the system may indicate that further clinical review is needed rather than provide a definitive diagnosis. A standardized sequence of independent image review, AI-result presentation, and clinical verification may preserve the benefits of AI while reducing the influence of incorrect suggestions on clinicians' final decisions [14, 15]. Human–AI collaboration studies should record when clinicians view the AI results, whether they accept or reject AI recommendations, and how AI errors and clinicians' responses to these errors affect final decisions. The safety and usability of the system should also be evaluated in real clinical workflows [15, 17].

6. Conclusion

Recent studies have demonstrated that AI has good potential as an assistive tool to detect intracranial aneurysms on CTA and TOF-MRA. Deep learning models can automatically flag suspicious lesions. Some reader studies show that AI assistance also improved the sensitivity of clinicians for detection and reading efficiency. Multicenter validation and prospective clinical studies indicated that AI could support in different hospitals and clinical workflows. However, the high overall accuracy does not mean the model works equally well with all lesions, especially those ≤3 mm in size and atypical aneurysms, such as fusiform and thrombosed aneurysms.

Multiple factors contribute to both false-positives and false-negatives. Small aneurysms contain limited information in imaging data, while normal vascular structures can appear as aneurysms. Also, representation of training data, discrepancy in image devices and scanning parameters, as well as the choice of detection threshold could affect model performance. Lower thresholds may reduce missed lesions, but they also increase incorrect markings with higher sensitivity but also mislabel arteries, veins, and other nonvascular structures, whereas higher thresholds would lower false-positives but miss small aneurysms. Therefore, overall accuracy, AUC, and sensitivity are not sufficient for model evaluation. Model performance among different lesion-size subgroups, false-positives per case, specific patterns of errors and external validation are also important factors to consider.

DOI: 10.54254/2755-2721/2026.AST36870

There are several limitations in the current evidence. Most studies were retrospective designs. Some datasets contained a higher proportion of aneurysm-positive cases or excluded complex cases than real practice situations which made the model performance in research datasets more favorable than actual clinical practices. The definition of small aneurysms, standards of reference and evaluation measures among studies which hampers direct comparison. This review includes mainly English language PubMed-indexed studies and may miss relevant engineering or non-English studies. Also, there is no meta-analysis due to substantial heterogeneity among included studies.

Further improvements are needed in data, model design, evaluation, and human–AI interaction. Datasets should include more small, atypical, and difficult cases. Model designs should combine multiscale imaging features with vascular anatomical information. Patient-level and lesion-level measures should be reported separately, and lesion-size categories should be standardized. Multicenter external validation is also needed to determine whether model performance can be maintained across different hospitals, imaging devices, and patient populations. In clinical settings, independent clinician assessment should be followed by AI suggestions and human verification. Future studies should also investigate AI's impact on missed diagnoses, unnecessary follow-ups, or referrals, workload, and cost-effectiveness.

Overall, AI is more suitable as a suspicious-lesion alert tool and a second reader for CTA and TOF-MRA than as an independent diagnostic system. It should not directly determine whether a patient requires surgical or endovascular treatment. Its clinical value depends not only on whether it improves sensitivity, but also on whether its errors can be explained and controlled. The model should maintain stable performance across different clinical settings and support clinicians without replacing their independent judgment.

References

[1] Gervais de Liyis, B., Nathania, N., Hartono, O., Alhaq, A. M. G., Sihombing, E. J. N., Hafif, M., Kusdiansah, M., Lang, M. J., Arham, A., Benet, A., & July, J. (2026). Shifting the paradigm in intracranial aneurysm detection with deep learning: A diagnostic accuracy meta-analysis and meta-regression. *Neurosurgical Review, 49*(1), Article 259.

[2] Joo, B. (2025). Methodological challenges in deep learning-based detection of intracranial aneurysms: A scoping review. *Neurointervention, 20*(2), 52–65.

[3] Bizjak, Ž., & Špiclin, Ž. (2023). A systematic review of deep-learning methods for intracranial aneurysm detection in CT angiography. *Biomedicines, 11*(11), Article 2921. https://doi.org/10.3390/biomedicines11112921

[4] Liu, X., Mao, J., Sun, N., Yu, X., Chai, L., Tian, Y., Wang, J., Liang, J., Tao, H., Yuan, L., Lu, J., Wang, Y., Zhang, B., Wu, K., Wang, Y., Chen, M., Wang, Z., & Lu, L. (2023). Deep learning for detection of intracranial aneurysms from computed tomography angiography images. *Journal of Digital Imaging, 36*(1), 114–123.

[5] Hu, B., Shi, Z., Lu, L., Miao, Z., Wang, H., Zhou, Z., Zhang, F., Wang, R., Luo, X., Xu, F., Li, S., Fang, X., Wang, X., Yan, G., Lv, F., Zhang, M., Sun, Q., Cui, G., Liu, Y., . . . China Aneurysm AI Project Group. (2024). A deep-learning model for intracranial aneurysm detection on CT angiography images in China: A stepwise, multicentre, early-stage clinical validation study. *The Lancet Digital Health, 6*(4), e261–e271.

[6] Pettersson, S. D., Filo, J., Liaw, P., Skrzypkowska, P., Klepinowski, T., Szmuda, T., Fodor, T. B., Ramirez-Velandia, F., Zieliński, P., Chang, Y.-M., Taussky, P., & Ogilvy, C. S. (2025). Addressing limited generalizability in artificial intelligence-based brain aneurysm detection for computed tomography angiography: Development of an externally validated artificial intelligence screening platform. *Neurosurgery, 97*(6), 1388–1396.

[7] Zhuo, L., Zhang, Y., Song, Z., Mo, Z., Xing, L., Zhu, F., Meng, H., Chen, L., Qu, G., Jiang, P., Wang, Q., Cheng, R., Mi, X., Liu, L., Hong, N., Cao, X., Wu, D., Wang, J., & Yin, X. (2025). Enhancing radiologists' performance in detecting cerebral aneurysms using a deep learning model: A multicenter study. *Academic Radiology, 32*(3), 1611–1620.

[8] Terasaki, Y., Yokota, H., Tashiro, K., Maejima, T., Takeuchi, T., Kurosawa, R., Yamauchi, S., Takada, A., Mukai, H., Ohira, K., Ota, J., Horikoshi, T., Mori, Y., Uno, T., & Suyari, H. (2022). Multidimensional deep learning reduces false-positives in the automated detection of cerebral aneurysms on time-of-flight magnetic resonance angiography: A multi-center study. *Frontiers in Neurology, 12*, Article 742126.

DOI: 10.54254/2755-2721/2026.AST36870

[9] Lehnen, N. C., Haase, R., Schmeel, F. C., Vatter, H., Dorn, F., Radbruch, A., & Paech, D. (2022). Automated detection of cerebral aneurysms on TOF-MRA using a deep learning approach: An external validation study. *American Journal of Neuroradiology, 43*(12), 1700–1705.

[10] Kuwabara, M., Ikawa, F., Sakamoto, S., Okazaki, T., Ishii, D., Hosogai, M., Maeda, Y., Chiku, M., Kitamura, N., Choppin, A., Takamiya, D., Shimahara, Y., Nakayama, T., Kurisu, K., & Horie, N. (2023). Effectiveness of tuning an artificial intelligence algorithm for cerebral aneurysm diagnosis: A study of 10,000 consecutive cases. *Scientific Reports, 13*(1), Article 16202.

[11] Li, Y., Zhang, H., Sun, Y., Fan, Q., Wang, L., Ji, C., Gu, H., Chen, B., Zhao, S., Wang, D., Yu, P., Li, J., Yang, S., Zhang, C., & Wang, X. (2024). Deep learning-based platform performs high detection sensitivity of intracranial aneurysms in 3D brain TOF-MRA: An external clinical validation study. *International Journal of Medical Informatics, 188*, Article 105487.

[12] Kim, D., Park, J., Yang, H., Kim, G.-Y., Lee, J., Kim, D., Han, H. J., Park, K. Y., & Lee, M. (2026). A landmark-guided dual-stream synergistic framework for automated intracranial aneurysm detection in magnetic resonance angiography. *Journal of Imaging Informatics in Medicine*. Advance online publication.

[13] Kim, J., Ceballos-Arroyo, A., Lin, C.-H., Liu, P., Jiang, H., Yadav, S., Wan, Q., Qin, L., & Young, G. S. (2025). Automated, anatomy-based, heuristic post-processing reduces false positives and improves interpretability of deep learning intracranial aneurysm detection models. *Scientific Reports, 16*(1), Article 3218.

[14] Adamchic, I., Kantelhardt, S. R., Wagner, H.-J., & Burbelko, M. (2024). Artificial intelligence can help detecting incidental intracranial aneurysm on routine brain MRI using TOF MRA data sets and improve the time required for analysis of these images. *Neuroradiology, 66*(12), 2195–2204.

[15] Kim, S. H., Schramm, S., Riedel, E. O., Schmitzer, L., Rosenkranz, E., Kertels, O., Bodden, J., Paprottka, K., Sepp, D., Renz, M., Kirschke, J., Baum, T., Maegerlein, C., Boeckh-Behrens, T., Zimmer, C., Wiestler, B., & Hedderich, D. M. (2025). Automation bias in AI-assisted detection of cerebral aneurysms on time-of-flight MR angiography. *La Radiologia Medica, 130*(4), 555–566.

[16] Maupu, C., Lebas, H., & Boulaftali, Y. (2022). Imaging modalities for intracranial aneurysm: More than meets the eye. *Frontiers in Cardiovascular Medicine, 9*, Article 793072. https://doi.org/10.3389/fcvm.2022.793072

[17] Vasey, B., Nagendran, M., Campbell, B., Clifton, D. A., Collins, G. S., Denaxas, S., Denniston, A. K., Faes, L., Geerts, B., Ibrahim, M., Liu, X., Mateen, B. A., Mathur, P., McCradden, M. D., Morgan, L., Ordish, J., Rogers, C., Saria, S., Ting, D. S. W., . . . DECIDE-AI Expert Group. (2022). Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *BMJ, 377*, Article e070904.