AJNR

Deep Learning–Based Software Improves Clinicians' Detection Sensitivity of Aneurysms on Brain TOF-MRA

B. Sohn, K.-Y. Park, J. Choi, J.H. Koo, K. Han, B. Joo, S.Y. Won, J. Cha, H.S. Choi and S.-K. Lee

Deep Learning–Based Software Improves Clinicians' Detection Sensitivity of Aneurysms on Brain TOF-MRA

B. Sohn, K.-Y. Park, J. Choi, J.H. Koo, K. Han, B. Joo, S.Y. Won, J. Cha, H.S. Choi, and S.-K. Lee

ABSTRACT

BACKGROUND AND PURPOSE: The detection of cerebral aneurysms on MRA is a challenging task. Recent studies have used deep learning–based software for automated detection of aneurysms on MRA and have reported high performance. The purpose of this study was to evaluate the incremental value of using deep learning–based software for the detection of aneurysms on MRA by 2 radiologists, a neurosurgeon, and a neurologist.

MATERIALS AND METHODS: TOF-MRA examinations of intracranial aneurysms were retrospectively extracted. Four physicians interpreted the MRA blindly. After a washout period, they interpreted MRA again using the software. Sensitivity and specificity per patient, sensitivity per lesion, and the number of false-positives per case were measured. Diagnostic performances, including subgroup analysis of lesions, were compared. Logistic regression with a generalized estimating equation was used.

RESULTS: A total of 332 patients were evaluated; 135 patients had positive findings with 169 lesions. With software assistance, patient-based sensitivity was statistically improved after the washout period (73.5% versus 86.5%, $P < .001$). The neurosurgeon and neurologist showed a significant increase in patient-based sensitivity with software assistance (74.8% versus 85.2%, $P = .03$, and 56.3% versus 84.4%, $P < .001$, respectively), while the number of false-positive cases did not increase significantly (23 versus 30, $P = .20$, and 22 versus 24, $P = .75$, respectively).

CONCLUSIONS: Software-aided reading showed significant incremental value in the sensitivity of clinicians in the detection of aneurysms on MRA without a significant increase in false-positive findings, especially for the neurosurgeon and neurologist. Software-aided reading showed equivocal value for the radiologist.

ABBREVIATIONS: ACA = anterior cerebral artery; CAD = computer-assisted detection

U nruptured intracranial aneurysms remain a major public health concern, and their prevalence is estimated to be 3.2% in healthy adults.[1] The annual incidence of aneurysm rupture is approximately 1% and is associated with a high risk of morbidity and mortality.[2]

TOF-MRA is a widely available technique that shows high sensitivity for the detection of aneurysms.[3] It can be performed as an initial screening because it is noninvasive and requires no contrast agent or radiation.

Detection of cerebral aneurysms on MRA is a challenging task for radiologists, neurosurgeons, and neurologists. Interpretation of both source and MIP reconstructed images is recommended to achieve good sensitivity.[4] However, detecting small lesions is a difficult and time-consuming task. Moreover, there is a relative shortage of experienced radiologists, owing to an increasing demand for imaging studies.[5]

Therefore, computer-assisted detection (CAD) of aneurysms is expected to play a key role in improving detection accuracy. Various CAD software packages for cerebral aneurysms have been investigated and have shown desirable sensitivity.[6-8] Miki et al[9] reported that routine integration of CAD with MRA for aneurysms was feasible and could help radiologists find more aneurysms without any reduction in specificity.

Recently, the use of machine learning methods has led to improvement in image-classification tasks. Along with recent advances in machine learning with the emergence of convolutional neural networks, several studies have evaluated the feasibility of using deep learning–based algorithms for the automated detection of intracranial aneurysms.[10-14] A previous study used a deep learning–based CAD software for the automated detection and localization of intracranial aneurysms on MRA and validated its high sensitivity and specificity via internal and external test sets.[15]

Due to the previously mentioned relative shortage of radiologists and increasing workload, in practice, clinicians sometimes have to meet the patient and review the MRA without an official report from a neuroradiologist. Park et al[11] investigated the performance of clinicians in identifying intracranial aneurysms using CT angiography, assisted by a deep learning model. However, to the best of our knowledge, no study has yet investigated the performance improvement of a clinician in aneurysm detection on MRA using CAD software.

The present study aimed to evaluate the effect of deep learning–based CAD software for the detection of cerebral aneurysms in MRA interpretation by 2 radiologists, a neurosurgeon, and a neurologist. Therefore, our primary end point was to investigate the improvement of accuracy with the assistance of software, including subgroup analysis based on readers and size, volume, and location of the lesions.

MATERIALS AND METHODS

Patient Cohort
This single-center, retrospective study was approved by the institutional review board (Severance Hospital, Seoul, Korea), and the requirement for informed consent was waived. We used the diagnostic cohort of a previous study, which investigated whether a deep learning model can achieve a target performance comparable with that of human radiologists. The required number of examinations was calculated to be 135 for aneurysm-containing examinations and 197 for aneurysm-free examinations using a sample-size calculation formula.[16]

TOF-MRA examinations of intracranial aneurysms were extracted from January 2018 to June 2019. The inclusion criteria were as follows: 1) older than 18 years of age; 2) MRA performed using 1.5T or 3T scanners; and 3) intracranial aneurysms on a radiology report. The number of eligible MRA examinations that met the inclusion criteria was 419. The exclusion criteria were as follows: 1) nonsaccular aneurysm, such as mycotic aneurysm, dissecting aneurysm, or pseudoaneurysm ($n = 8$); 2) giant aneurysm, >25 mm in diameter ($n = 0$); 3) ruptured aneurysm ($n = 0$); 4) aneurysms treated with surgical clipping, coil embolization, or stent insertion ($n = 118$); 5) significant displacement of the intracranial vascular structure due to intracranial hemorrhage or tumor ($n = 1$); and 6) pronounced artifacts ($n = 4$). On the basis of these criteria, 288 aneurysm-containing examinations were eligible for inclusion in the diagnostic cohort.

Reference Preparation
Three neuroradiologists (with 2, 11, and 15 years of experience in neuroradiology) independently reviewed the 288 MRA examinations in consecutive registries and evaluated the number and location of aneurysms, referring to data from other imaging modalities when available, such as CT angiography or DSA. Only examinations in which all 3 neuroradiologists concurred on the number and location of aneurysms were finally included in the diagnostic cohort. Consensus was not reached in 28 aneurysm-containing examinations; therefore, these cases were excluded. All of the excluded cases demonstrated aneurysms of <2 mm in diameter with equivocal, bulging contours, which were possibly a junctional dilation or atherosclerotic luminal irregularity. Finally, 135 aneurysm-containing examinations were prepared for the diagnostic cohort.

For the aneurysm-free examinations in the diagnostic cohort, 300 aneurysm-free MRA examinations based on radiologic reports were randomly extracted. The 3 neuroradiologists reviewed the recruited examinations in chronologic order and reached a consensus that these examinations did not show any discernible aneurysms, significant vascular steno-occlusion, or significant structural abnormalities. Finally, 197 aneurysm-free examinations were selected for the diagnostic cohort. All data were anonymized.

TOF-MRA Data Acquisition
Examinations were acquired on both 1.5T (Achieva, Philips Healthcare) and 3T scanners (Achieva and Ingenia, Philips Healthcare; Trio Tim, Siemens; Discovery MR750, GE Healthcare). The following parameters were used for 3D TOF-MRA: TR, 20–24 ms; TE, 3.4 ms; flip angle, 18°–20°; FOV, 200–250 mm; section thickness, 0.5–0.7 mm; matrix, 384×310–880×637.

CAD Software
The deep learning–based CAD software used in this study was developed by our team. The details of the algorithm have been published elsewhere.[15] To develop the original study model, we randomly extracted 600 patients from our hospital from 2014 to 2016. For validation, 110 patients from another institution were prepared for the external test set. The software was developed for classification using 3D ResNet architecture (https://github.com/kenshohara/3D-ResNets-PyTorch). The patch-wise binary classification algorithm was followed by a pixel-voting algorithm, which presents only boxes that have a higher probability of the presence of an aneurysm than a certain cutoff value to reduce the number of false-positives for aneurysm detection. Finally, the CAD software presented a 1.0×1.0 cm bounding box, which was expected to contain an aneurysm with a certain probability, on final images (Online Supplemental Data). If the algorithm predicted that there would be no aneurysm on examination, no bounding box was presented.

Study Design and Statistical Analysis
Four physicians (a neurosurgeon, neurologist, neuroradiologist, and radiology resident) interpreted MRA examinations under blinded conditions. After a 1-month washout period, MRA was interpreted again with software assistance. Interpreters could see bounding boxes using the software and freely decide whether to accept or ignore them. We measured the sensitivity and specificity per patient, sensitivity per lesion, the number of false-positives per case, and the total time required for the interpretation. For sensitivity per patient,
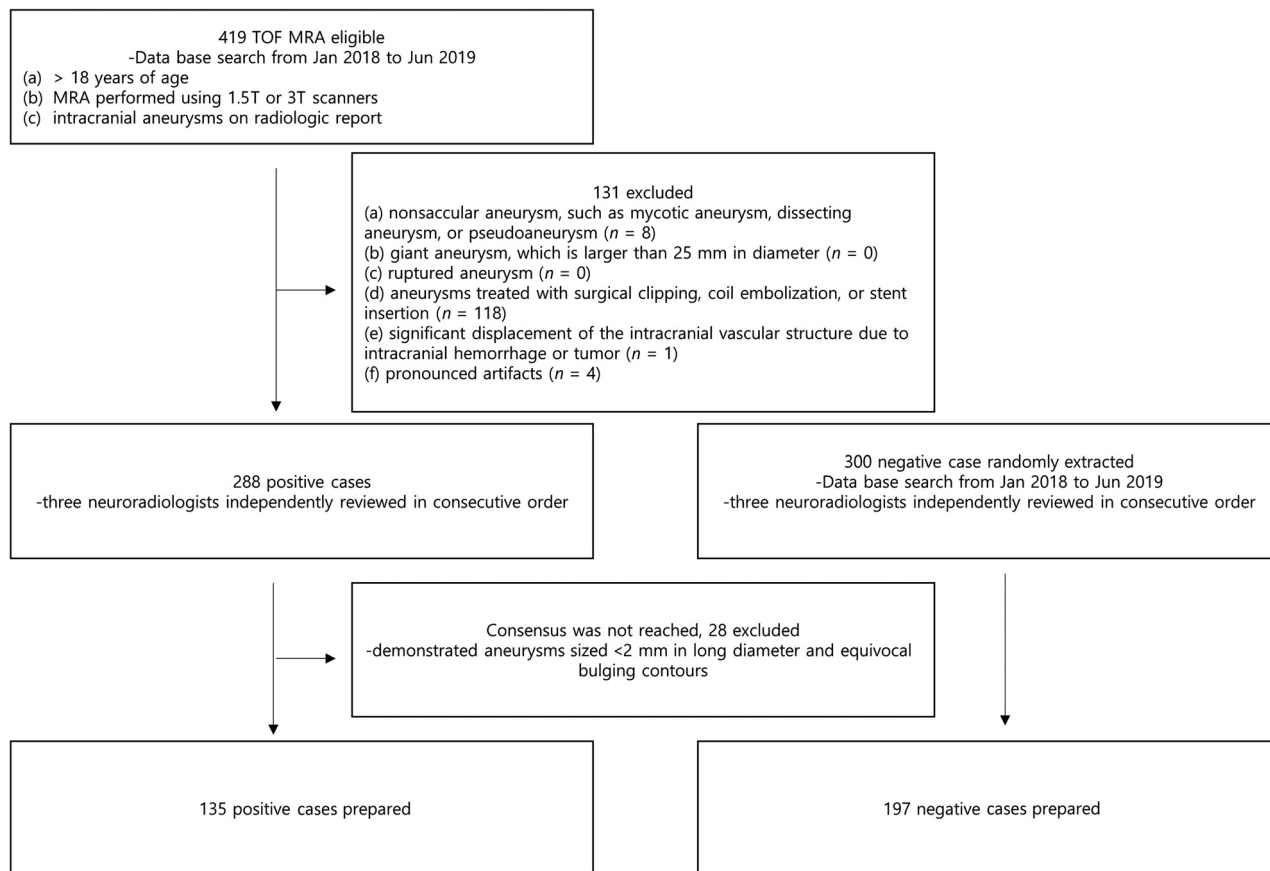
419 TOF MRA eligible
-Data base search from Jan 2018 to Jun 2019
(a) > 18 years of age
(b) MRA performed using 1.5T or 3T scanners
(c) intracranial aneurysms on radiologic report

131 excluded
(a) nonsaccular aneurysm, such as mycotic aneurysm, dissecting aneurysm, or pseudoaneurysm (*n* = 8)
(b) giant aneurysm, which is larger than 25 mm in diameter (*n* = 0)
(c) ruptured aneurysm (*n* = 0)
(d) aneurysms treated with surgical clipping, coil embolization, or stent insertion (*n* = 118)
(e) significant displacement of the intracranial vascular structure due to intracranial hemorrhage or tumor (*n* = 1)
(f) pronounced artifacts (*n* = 4)

288 positive cases
-three neuroradiologists independently reviewed in consecutive order

300 negative case randomly extracted
-Data base search from Jan 2018 to Jun 2019
-three neuroradiologists independently reviewed in consecutive order

Consensus was not reached, 28 excluded
-demonstrated aneurysms sized <2 mm in long diameter and equivocal bulging contours

135 positive cases prepared

197 negative cases prepared

FIG 1. Patient-selection flow diagram.

Table 1: Patient demographics

	Positive for Aneurysm	Negative for Aneurysm	P Value
No. of patients	135	197	
Age (range)	62 (52–69)	62 (53–73)	$P = .5^a$
Sex (male ratio)	36 (26.7%)	84 (42.6%)	$P = .0029^b$
No. of lesions	169	0	

^{a} Mann-Whitney U test.
^{b} χ^2 test.

we considered true-positive cases when all true-positive lesions were detected, without any false-negative or false-positive lesions. We compared these diagnostic performances between software versus human readers, software versus humans with CAD assistance, and humans without versus humans with CAD assistance. We performed this analysis on both reader-averaged and reader-individual results. Poisson regression with the generalized estimating equation was used for the number of false-positives per case, and logistic regression with the generalized estimating equation was used for the other cases.

For characteristics analysis, we divided the 169 aneurysms into subgroups according to the diameter, volume, and location. We analyzed the sensitivity per lesion and compared it among the different subgroups. The thresholds of each subgroup in terms of size and volume were 3 and 5 mm, and 10 and 20 mL, respectively. The location was divided into 6 subgroups (Online Supplemental Data). Both reader-averaged and reader-individual comparisons

were performed. The geeglm package from the R statistical and computing software (Version 4.0.2; http://www.r-project.org) was used for the analysis (performed by K.H., statistician).

RESULTS

A total of 332 MRA examinations were collected, of which 135 had positive findings and 197 had negative findings. The 135 examinations with positive findings included 169 aneurysms, ranging from 2.0 to 17 mm in maximum diameter (mean size, 3.98 [SD, 2.11] mm) (Fig 1).

There were 84 (42.6%) and 36 (26.7%) male patients in the aneurysm-negative and aneurysm-positive cohorts, respectively. There was no statistical difference ($P = .5$, by Mann-Whitney U test) between the mean ages of the aneurysm-negative and aneurysm-positive patient cohorts (62 versus 62 years) (Table 1).

Among the 169 aneurysms, 81 lesions were <3 mm, 76 lesions were 3–5 mm, and 12 lesions were ≥5 mm. In terms of volume, 58 lesions were <10 mL, 52 lesions were 10–20 mL, and 59 lesions were ≥20 mL. In terms of location, 5 lesions were located at major branch of the anterior cerebral artery (ACA); 106, at the around-dural ring; 12, at the extradural ring; 23, at the intracranial distal ICA; 19, at the MCA major branch; and 4, at the posterior circulation.

The Online Supplemental Data present the diagnostic performance of the CAD software, reader-averaged results (with and without CAD assistance), and the statistical comparison of
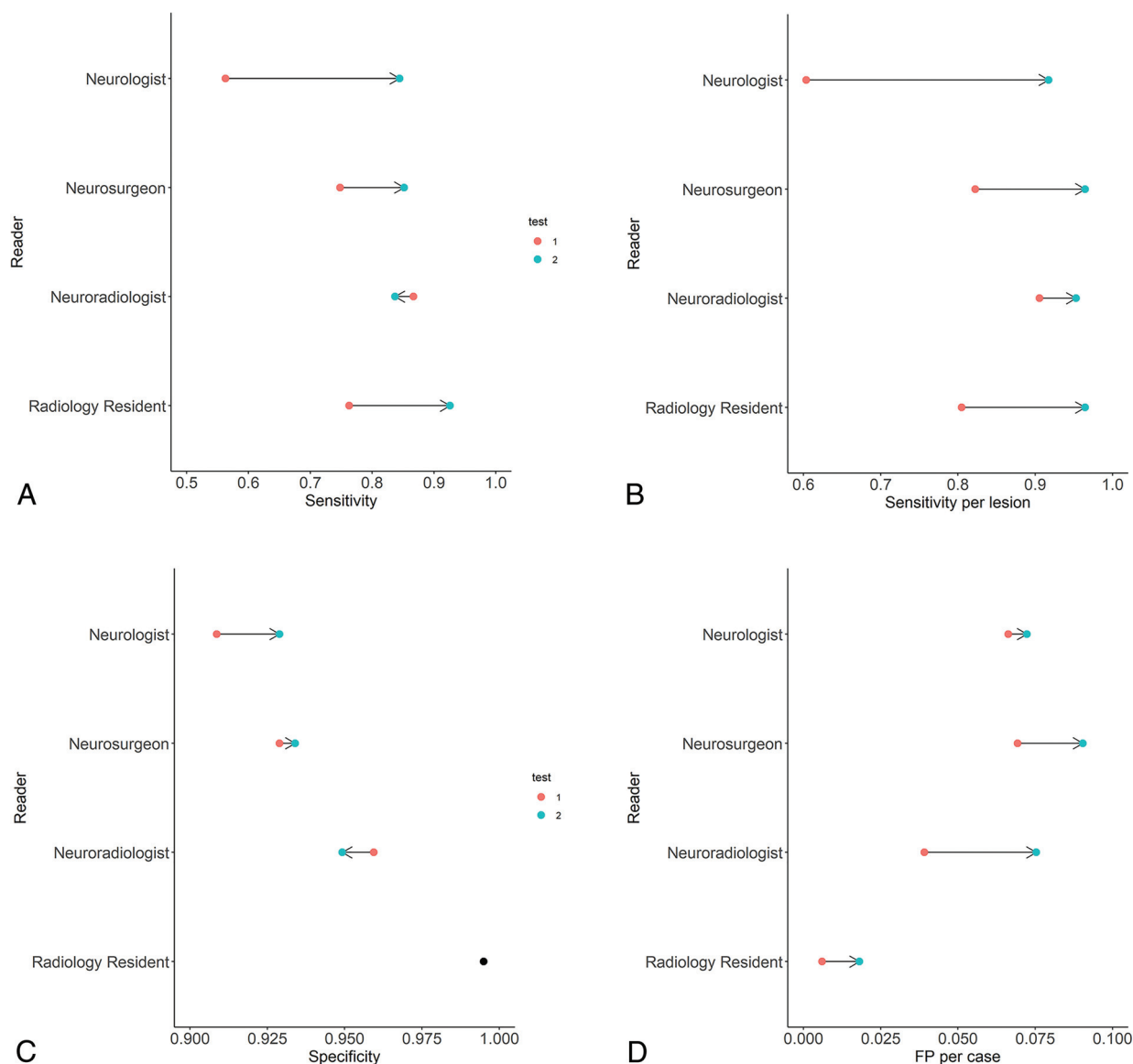
FIG 2. Change in individual diagnostic performance of clinicians. *Arrow lines* depict the change in diagnostic performance for each clinician without (*red dot*) and with (*blue dot*) software augmentation. The radiology resident showed no change in specificity between the 2 occasions (*black dot*). FP indicates false-positive.

diagnostic performance between CAD software and human readers, CAD software versus human readers (with CAD assistance), and human readers (without CAD assistance) versus human readers (with CAD assistance).

Under CAD assistance, both the patient-based sensitivity and sensitivity per lesion were statistically improved after the washout period (Fig 2). There was no difference in patient-based sensitivity between the CAD software and reader-averaged results without CAD assistance. However, when the CAD software was used, higher lesion-based sensitivity was observed compared with reader-averaged results. With CAD assistance, the patient-based sensitivity of readers was higher than that of CAD software. The number of false-positive cases was higher in readers than in the CAD software, and it increased significantly with CAD assistance in terms of reader average. There was no significant difference in specificity in any comparison.

Reader-individual diagnostic performances with/without CAD assistance are presented in Table 2.

For all readers, except the neuroradiologist, patient-based sensitivity and per-lesion sensitivity were increased significantly with CAD assistance (Fig 2). With respect to per-lesion sensitivity, neuroradiologists also showed a tendency toward increased performance with CAD assistance (95.3 versus 90.5, $P = .07$). No significant change in specificity was observed in any individual reader. The number of false-positives detected was significantly high for the neuroradiologist with CAD assistance. Other readers showed no significant changes.

The time of interpretation was 335 minutes (1.01 minute) for the neurosurgeon, 329 minutes (0.99 minute) for the neurologist, 211 minutes (0.64 minute) for the radiology resident, and 205 minutes (0.62 minute) for the neuroradiologist during the first time period. After CAD assistance, it changed to 260 minutes

Table 2: Reader-individual diagnostic performances with/without CAD assistance

	Reader	Human without CAD Assistance	Human with CAD Assistance	Human without vs with CAD Assistance (P Value)
Sensitivity (%) (95% CI)	Neurologist	56.3 (47.9–64.7)	84.4 (78.3–90.6)	<.001
	Neurosurgeon	74.8 (67.5–82.1)	85.2 (79.2–91.2)	.032
	Radiologist	86.7 (80.9–92.4)	83.7 (77.5–89.9)	.493
	Resident	76.3 (69.1–83.5)	92.6 (88.2–97)	.001
Specificity (%) (95% CI)	Neurologist	90.9 (86.8–94.9)	92.9 (89.3–96.5)	.460
	Neurosurgeon	92.9 (89.3–96.5)	93.4 (89.9–96.9)	.842
	Radiologist	95.9 (93.2–98.7)	94.9 (91.9–98)	.629
	Resident	99.5 (98.5–100)	99.5 (98.5–100)	>.999
Sensitivity per lesion (%) (95% CI)	Neurologist	60.4 (53–67.7)	91.7 (87.2–96.3)	<.001
	Neurosurgeon	82.2 (76–88.5)	96.4 (93.7–99.2)	<.001
	Radiologist	90.5 (85.7–95.3)	95.3 (92.2–98.3)	.070
	Resident	80.5 (74.6–86.4)	96.4 (93.7–99.2)	<.001
No. of false-positives per case (95% CI)	Neurologist	0.066 (0.044–0.099)	0.072 (0.048–0.108)	.751
	Neurosurgeon	0.069 (0.047–0.103)	0.09 (0.062–0.132)	.200
	Radiologist	0.039 (0.022–0.07)	0.075 (0.051–0.112)	.028
	Resident	0.006 (0.002–0.024)	0.018 (0.008–0.04)	.179
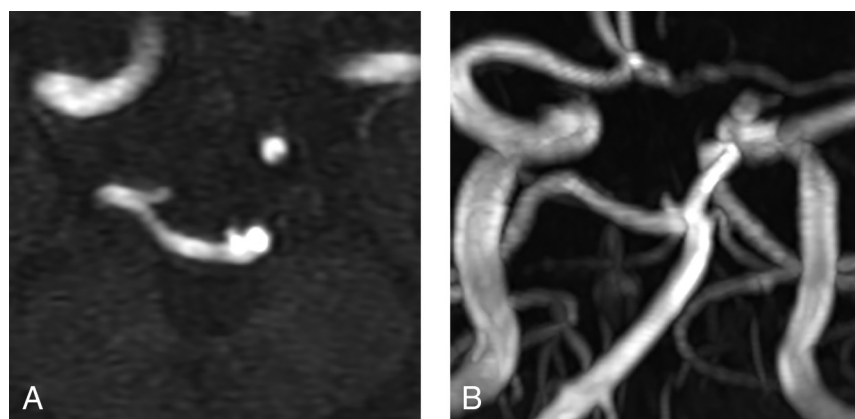
FIG 3. Source (*A*) and MIP image (*B*) from TOF-MRA of 75-year-old female patient. There was a basilar top aneurysm, <3 mm. Human readers detected this lesion, but the computer-assisted detection software did not.

(0.78 minute) for the neurosurgeon, 155 minutes (0.47 minute) for the neurologist, 184 minutes (0.55 minute) for the radiology resident, and 215 minutes (0.65 minute) for the neuroradiologist (entire dataset, with average minutes per individual case in parentheses).

The sensitivity per lesion in subgroups (the 95% confidence interval was estimated by logistic regression with the generalized estimating equation) is presented in the Online Supplemental Data. It shows a comparison of sensitivity per lesion between CAD software versus human readers, CAD software versus human readers with CAD assistance, and human readers with versus without CAD assistance. Reader-individual comparisons were also performed (Online Supplemental Data).

CAD software showed 84% lesion sensitivity for small (<3 mm and 10-mL volume) aneurysms and 100% for larger lesions. CAD software showed higher sensitivity per lesion in every subgroup of diameter and size compared with average human readers.

With CAD assistance, human readers showed improved performance in the <3 and 3–5 mm subgroups. Moreover, in every volume subgroup, human readers showed improved performance with CAD assistance.

In terms of location, average human readers showed significant increases in the sensitivity per lesion in every location, except for ACA area lesions. CAD showed higher sensitivity than an average human without CAD assistance in the subgroups of the around-dural ring, intracranial distal ICA, and MCA. In addition, CAD showed higher sensitivity than an average human with CAD assistance in the subgroups of the around-dural ring and intracranial distal ICA. However, with CAD assistance, the average human showed higher sensitivity per lesion in the subgroup of posterior circulation.

We analyzed the detection results for 169 lesions by 4 human readers (*n* = 676) and compared them with the detection results of CAD software by subgroup (Online Supplemental Data).

All lesions that the human readers detected but CAD ignored were small, <3 mm in diameter (Fig 3). Approximately 10% of the lesions that had a diameter of <3 mm that CAD missed could be detected by human readers. In terms of volume, approximately 14% of the lesions of <10 mL that CAD missed could be detected by human readers. However, lesions of >3 mm or 10 mL were never missed by CAD software, though human readers sometimes missed these lesions (Fig 4). The reader-individual comparison with CAD software is presented in the Online Supplemental Data.

DISCUSSION

Our study revealed a significant improvement in diagnostic performance of multiple physicians having different specialties in detecting cerebral aneurysms with the use of deep learning–based software. This type of software, called artificial intelligence, was developed using deep learning and achieved a target diagnostic performance comparable with that of human radiologists. Notably, this software received approval from the Korean

FIG 4. Source (*A*) and MIP (*B*) images from TOF-MRA of a 35-year-old female patient. Some human readers missed this lesion at first, but computer-assisted detection software detected it. *C*, During the second interpretation, human readers accepted the diagnosis of the software.

Ministry of Food and Drug Safety as an artificial intelligence–applied software.

With the assistance of this software, the neurosurgeon and neurologist involved in this study experienced significant improvement in their performance in terms of sensitivity per patient and per lesion without an increase in the number of false-positive cases. The time for interpretation was also reduced. In particular, an improvement in the detection rates for lesions of <5 mm was observed with the assistance of this software.

These results suggest a promising effect of CAD software in a real clinical environment and are compatible with the results of several previous studies. Numerous investigations have shown that deep learning–based artificial intelligence could be helpful in detecting cerebral aneurysms on CT angiography.[11] However, to the best of our knowledge, this is the first study to demonstrate the performance improvement for aneurysm detection on MRA with multiple readers.

However, the effect was unclear in the case of the neuroradiologist with software assistance. Rather, it was observed that both the number of false-positives per case and interpretation time increased. We can speculate that this is a limitation of the CAD software, which should be further improved in future work. The training set for this software was developed on the basis of the neuroradiologist's interpretation of TOF-MRA.[15] Furthermore, the previous results of deep learning–based CAD were comparable with not better than those of the radiologist.

However, it can be concluded that this type of CAD software can be helpful for neurosurgeons and neurologists. In the subgroup analysis, it was observed that the software did not miss any lesions of >3 mm. This result was consistent with that of a previous article that was validated via an external test set.[15] The accuracy of the software may not only offer time saving and safety for clinicians but may also lead to a reduction in psychological burden, because clinically significant aneurysms can be filtered by the software.

Our study is a retrospective, single-center study. We are currently working toward a prospective, multicenter study for the evaluation of the augmenting effect of this software, and it is under review by the institutional review board. Furthermore, we plan to merge this software with our PACS system and apply it real-time in a clinical environment.

Our study has several limitations. First, it was difficult to gather DSA-confirmed positive and negative cases of aneurysms with TOF-MRA. Therefore, a consensus had to be reached among the 3 neuroradiologists after consecutively gathering the MRA examinations. Although DSA is a more powerful diagnostic tool for evaluating cerebral aneurysms, a previous study has shown that in terms of sensitivity, 3T MRA is not inferior to DSA.[17]

Second, our patient cohort showed an unbalanced distribution of lesion locations. The lesions were mostly located on the ICA. The ACA, MCA, and posterior circulation region accounted for a relatively small portion. However, we gathered patients consecutively from our data base; therefore, we believe that this imbalance reflected the natural incidence of cerebral aneurysm location.

Third, the total number of false-positive predictions increased for all reviewers with CAD. While this was only significant in the neuroradiologist, considering the low incidence of aneurysms in a real clinical environment, the increase in the false-positive rate could be much higher. In our study population, the incidence of cerebral aneurysms was higher than that in the real world; therefore, the increase in false-positive rates could be underestimated.

Last, we could not apply a crossover interpretation design due to the small number of readers involved in our study. Considering the large number of cases and washout periods, the memory effect would be negligible in our study. However, we hope to have a more organized crossover design and many interpreters in our prospective, multicenter study.

CONCLUSIONS

We found that the assistance of deep learning–based software significantly improved clinicians' sensitivity, without a significant increase in false-positive cases, especially in the case of the neurosurgeon and neurologist.

REFERENCES

1. Vlak MH, Algra A, Brandenburg R, et al. **Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis.** *Lancet Neurol* 2011;10:626–36 CrossRef Medline

2. Juvela S, Korja M. **Intracranial aneurysm parameters for predicting a future subarachnoid hemorrhage: a long-term follow-up study.** *Neurosurgery* 2017;81:432–40 CrossRef Medline

3. Sailer AM, Wagemans BA, Nelemans PJ, et al. **Diagnosing intracranial aneurysms with MR angiography: systematic review and meta-analysis.** *Stroke* 2014;45:119–26 CrossRef Medline

4. Korogi Y, Takahashi M, Mabuchi N, et al. **Intracranial aneurysms: diagnostic accuracy of MR angiography with evaluation of maximum intensity projection and source images.** *Radiology* 1996;199:199–207 CrossRef Medline

5. McDonald RJ, Schwartz KM, Eckel LJ, et al. **The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload.** *Acad Radiol* 2015;22:1191–98 CrossRef Medline

6. Arimura H, Li Q, Korogi Y, et al. **Automated computerized scheme for detection of unruptured intracranial aneurysms in three-dimensional magnetic resonance angiography.** *Acad Radiol* 2004;11:1093–104 CrossRef Medline

7. Hirai T, Korogi Y, Arimura H, et al. **Intracranial aneurysms at MR angiography: effect of computer-aided diagnosis on radiologists' detection performance.** *Radiology* 2005;237:605–10 CrossRef Medline

8. Kakeda S, Korogi Y, Arimura H, et al. **Diagnostic accuracy and reading time to detect intracranial aneurysms on MR angiography using a computer-aided diagnosis system.** *AJR Am J Roentgenol* 2008;190:459–65 CrossRef

9. Miki S, Hayashi N, Masutani Y, et al. **Computer-assisted detection of cerebral aneurysms in MR angiography in a routine image-reading environment: effects on diagnosis by radiologists.** *AJNR Am J Neuroradiol* 2016;37:1038–43 CrossRef Medline

10. Nakao T, Hanaoka S, Nomura Y, et al. **Deep neural network-based computer-assisted detection of cerebral aneurysms in MR angiography.** *J Magn Reson Imaging* 2018;47:948–53 CrossRef Medline

11. Park A, Chute C, Rajpurkar P, et al. **Deep learning-assisted diagnosis of cerebral aneurysms using the HeadXNet model.** *JAMA Netw Open* 2019;2:e195600 CrossRef Medline

12. Sichtermann T, Faron A, Sijben R, et al. **Deep learning-based detection of intracranial aneurysms in 3D TOF-MRA.** *AJNR Am J Neuroradiol* 2019;40:25–32 CrossRef Medline

13. Stember JN, Chang P, Stember DM, et al. **Convolutional neural networks for the detection and measurement of cerebral aneurysms on magnetic resonance angiography.** *J Digit Imaging* 2019;32:808–15 CrossRef Medline

14. Ueda D, Yamamoto A, Nishimori M, et al. **Deep learning for MR angiography: automated detection of cerebral aneurysms.** *Radiology* 2019;290:187–94 CrossRef Medline

15. Joo B, Ahn SS, Yoon PH, et al. **A deep learning algorithm may automate intracranial aneurysm detection on MR angiography with high diagnostic performance.** *Eur Radiol* 2020;30:5785–93 CrossRef

16. Hajian-Tilaki K. **Sample size estimation in diagnostic test studies of biomedical informatics.** *J Biomed Inform* 2014;48:193–204 CrossRef Medline

17. Mine B, Pezzullo M, Roque G, et al. **Detection and characterization of unruptured intracranial aneurysms: comparison of 3T MRA and DSA.** *J Neuroradiol* 2015;42:162–68 CrossRef