Computer-Assisted Detection of Cerebral Aneurysms in MR Angiography in a Routine Image-Reading Environment: Effects on Diagnosis by Radiologists

S. Miki, N. Hayashi, Y. Masutani, Y. Nomura, T. Yoshikawa, S. Hanaoka, M. Nemoto and K. Ohtomo

AJNR Am J Neuroradiol published online 18 February 2016
http://www.ajnr.org/content/early/2016/02/18/ajnr.A4671
Computer-Assisted Detection of Cerebral Aneurysms in MR Angiography in a Routine Image-Reading Environment: Effects on Diagnosis by Radiologists

S. Miki, N. Hayashi, Y. Masutani, Y. Nomura, T. Yoshikawa, S. Hanaoka, M. Nemoto, and K. Ohtomo

ABSTRACT

BACKGROUND AND PURPOSE: Experiences with computer-assisted detection of cerebral aneurysms in diagnosis by radiologists in real-life clinical environments have not been reported. The purpose of this study was to evaluate the usefulness of computer-assisted detection in a routine reading environment.

MATERIALS AND METHODS: During 39 months in a routine clinical practice environment, 2701 MR angiograms were each read by 2 radiologists by using a computer-assisted detection system. Initial interpretation was independently made without using the detection system, followed by a possible alteration of diagnosis after referring to the lesion candidate output from the system. We used the final consensus of the 2 radiologists as the reference standard. The sensitivity and specificity of radiologists before and after seeing the lesion candidates were evaluated by aneurysm- and patient-based analyses.

RESULTS: The use of the computer-assisted detection system increased the number of detected aneurysms by 9.3% (from 258 to 282). Aneurysm-based analysis revealed that the apparent sensitivity of the radiologists’ diagnoses made without and with the detection system was 64% and 69%, respectively. The detection system presented 82% of the aneurysms. The detection system more frequently benefited radiologists than being detrimental.

CONCLUSIONS: Routine integration of computer-assisted detection with MR angiography for cerebral aneurysms is feasible, and radiologists can detect a number of additional cerebral aneurysms by using the detection system without a substantial decrease in their specificity. The low confidence of radiologists in the system may limit its usefulness.

ABBREVIATIONS: CAD = computer-assisted detection; FP = false-positive; TP = true-positive

Detection of unruptured cerebral aneurysms is a challenging task for radiologists. Unenhanced MRA has been widely accepted as a technique for initial screening because it is noninvasive and requires no contrast agent or ionizing radiation. Consider its role as a screening technique, MRA requires high sensitivity. At present, catheter DSA is still the criterion standard of diagnosis. The limitations of MRA compared with DSA include the limited spatial resolution and artifacts such as motion, susceptibility, and flow. Nevertheless, an increasing number of recent studies suggest that DSA is no longer considered essential for establishing the diagnosis of cerebral aneurysms, particularly when 3T MR imaging scanners are used.

Despite the continuing sophistication of the imaging technique, 1 important cause of the limited sensitivity of MRA is the detection failures of radiologists. Interpretation of both source and reconstructed images is recommended to achieve good sensitivity of MRA, but detecting relatively small lesions is a time-consuming and difficult task.

Computer-assisted detection (CAD) of cerebral aneurysms may play a role in improving the accuracy of aneurysm detection by MRA. Stand-alone performance figures of various CAD algorithms for cerebral aneurysms have been studied mainly by using datasets of known aneurysms, and high sensitivities have been reported.

Previous observer performance studies showed that CAD for cerebral aneurysms raises the sensitivity of radiologists or reduces reading time while maintaining the sensitivity. However, those studies were performed under experimental conditions with a relatively small number of aneurysms. The high prevalence...
of aneurysms (19%–44%14-16) may have resulted in a higher estimate of accuracy owing to observer expectation bias.17,18 In some studies, the mean diameters of aneurysms were relatively large (7.1 mm14 and 5.0 mm16) and radiologists had access to only MIP images; neither of these are found in recent routine screening. Recently, Stepán-Bucksakowska et al15 have investigated the detection performance of radiologists in an experimental environment closer to modern clinical settings. Still, to the best of our knowledge, experience with CAD of cerebral aneurysms in real-life clinical environments has not been reported in the literature. The purpose of our study, therefore, was to evaluate the usefulness of CAD of cerebral aneurysms in diagnosis by radiologists in a routine image-reading environment.

MATERIALS AND METHODS

Subjects

This study was approved by the ethics review board of the University of Tokyo Hospital. The subjects were a successive series of adults who were referred to our institution for their annual whole-body general medical examinations between October 2010 and December 2013. Written informed consent to use their clinical images for research about CAD conducted in our institution was obtained from all the subjects. All the subjects underwent a medical interview by a physician, in which their detailed medical history was taken. The initial inclusion criteria were as follows: 1) first-time visit to our institution, 2) MRA completed without contraindication, and 3) no known history of cerebral aneurysms. 3D time-of-flight unenhanced MRA was performed as part of brain screening, with three 3T MR units (2 Signa HDxt scanners and 1 Discovery MR750 scanner; GE Healthcare, Milwaukee, Wisconsin). The acquisition parameters were as follows: FOV, 240 mm; matrix size, 512×512; pixel spacing, 0.469 mm; section thickness, 1.2 mm; section interval, 0.6 mm (ie, there was a 50% overlap for each section); TR, 25 ms; TE, 2.7 ms for the Signa HDxt and 2.9 ms for the Discovery MR750. Rotational volume-rendered images around the x-, y-, and z-axes were reconstructed by radiologic technologists. Automatically generated MIP images were also transferred.

Imaging Interpretation

Within the data-acquisition period, MR angiograms were interpreted by 26 radiologists as part of daily routine diagnosis. Their years of experience in MRA interpretation in their daily routine ranged from 3 to 21 years. Three of the authors of this article (S.M., N.H., and T.Y.) also participated in the interpretation. Subject information such as age, sex, and current symptoms, if any, was not masked. The image reading method is illustrated in Fig 1. Two radiologists and 1 radiologic technologist were assigned to each subject on a day-to-day basis.

First, the 2 radiologists independently interpreted an MR angiogram without seeing the CAD results. This stage is defined as the “initial diagnosis.” Second, the 2 radiologists independently reviewed the CAD results displayed by a Web-based CAD server, the details of which are described later. Here, the 2 radiologists independently registered their personal “feedback” into the CAD server, to record the location of aneurysms detected and whether they had changed their diagnosis after seeing the CAD results, thus yielding the “post-CAD diagnosis.” The technologist also independently interpreted the images and made his or her personal report. Finally, after the 3 reading reports (1 from each of the radiologists and 1 from the technologist) were made, the 2 radiologists reviewed the 3 personal reading reports and discussed and made a single report by consensus, termed the “final diagnosis.” This review process was helped by a customized, structured reporting system, which automatically showed the 3 reading reports side by side. The diagnostic criterion for aneurysms was a saccular protrusion of ≥2 mm; lesions smaller than this were not included because of the limited spatial resolution of MRA. Fusiform aneurysms were also excluded. The sizes of the aneurysms in millimeter precision were also determined by consensus of the 2 radiologists. Each radiologist was able to interpret the source axial sections, volume-rendered images, and MIP images on computer displays.

CAD Software

The CAD software used in this study was developed by our team. The details of the algorithm are published elsewhere.19 Briefly, after a lesion candidate detection based on curvatures20 and Hessian eigenvalues,21 a classifier ensemble trained by the boosting algorithm22 was used to determine the likelihood of an aneurysm on the basis of 63 feature values of the candidates, such as statistics of voxel values, curvatures, and features derived from Hessian eigenvalues. This software was installed as a plug-in for a Web-based CAD server developed by Nomura et al23 (Fig 2).

The system was configured to always display the top 3 lesion candidates in likelihood per study rather than displaying variable numbers of candidates above a certain threshold of likelihood. One merit for this strategy was that the radiologists could stabilize the interpretation time. This strategy is also robust against the overall likelihood shift due to the inevitable image-quality variation between studies. We have confirmed in our preliminary study (not published) that our CAD system can maintain its sensitivity by using this “show 3” method compared with the variable number method.

A radiologist categorized each lesion candidate as either a “known true-positive (TP),” a “missed TP,” a “false-positive (FP),” or “pending.” Definitions were as follows: “known TP,” a true aneurysm that the radiologist had already recognized before
seeing the CAD results; “missed TP,” a lesion that he or she overlooked before seeing the CAD results; and “FP,” a false-positive candidate (ie, not an aneurysm). The “pending” selection in the final diagnosis indicated that the 2 radiologists did not reach a positive consensus, mainly because the lesion was too small. Such subjects were not referred to experts for further evaluation; thus, we did not include such pending selections as positive aneurysms.

If a lesion detected by the radiologist was not included in the 3 candidates displayed by the CAD system, the radiologist manually recorded the coordinates of the aneurysm by a mouse click (Fig 2B). Thus, by combining all these data, the server collected the following items: 1) all the locations of aneurysms determined by consensus, 2) whether each radiologist successfully detected the lesion before reviewing the CAD results, and 3) whether each positive lesion was successfully included in the CAD results as one of the top 3 candidates. Additionally, the median time for reviewing CAD results and giving feedback was determined by using the server log for the last 2 months of the data-acquisition period, by which time the radiologists were well-accustomed to the system.

**Data Analysis**

Statistical analysis was performed by using R, Version 3.1.2, statistical and computing software (http://www.r-project.org/). The sensitivity and the specificity of the radiologists were calculated by using the final diagnosis as the reference standard. The sensitivity of CAD was calculated as the successful presentation rate of positive lesions in the top 3 candidates. Then, 95% confidence intervals were computed on the basis of binomial distribution. The sensitivities of the radiologists before and after the CAD reference were compared by using the McNemar test. Additionally, the detection performance figures for both CAD and radiologists were compared between small (<3 mm) and large (≥4 mm) aneurysms by using the χ² test.

**RESULTS**

**Characteristics of Patients and Aneurysms**

During the study period, 2804 first-time visitors to our institution completed the annual health checkup examination program. Among them, 39 subjects did not meet the initial inclusion criteria because MR imaging was contraindicated/refused (n = 21) or they had a history of known cerebral aneurysms (n = 18). Moreover, 64 were excluded from further evaluations because the assigned radiologists did not complete the feedback registration. None of the MR angiograms were excluded because of poor image quality, and all of them were successfully processed by the CAD system. Thus, MR angiograms of 2701 subjects (1674 men, 1027 women) were included in this analysis. Subject median age was 54 years (range, 22–90 years). Two hundred three aneurysms from 189 subjects were determined in the final diagnosis (prevalence rate, 7.0%), the details of which are shown in the Table.

**Performance of the CAD System**

Overall, our CAD system successfully presented 166 (82%; 95% CI, 0.76%–0.87%) of the 203 aneurysms as the top 3 lesion candidates. Twenty-six aneurysms were detected but not presented as the top 3 candidates, and 11 aneurysms were not detected by CAD. The performance of the CAD system in relation to aneurysms of the top 3 candidates. Additionally, the median time for reviewing CAD results and giving feedback was determined by using the server log for the last 2 months of the data-acquisition period, by which time the radiologists were well-accustomed to the system.
Characteristics of 203 detected aneurysms in the final diagnosis

<table>
<thead>
<tr>
<th>Feature</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>101</td>
</tr>
<tr>
<td>Female</td>
<td>88</td>
</tr>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>Younger than 50 yr</td>
<td>50</td>
</tr>
<tr>
<td>50 yr and older, younger than 60 yr</td>
<td>54</td>
</tr>
<tr>
<td>60 yr and older</td>
<td>85</td>
</tr>
<tr>
<td>Multiplicity b</td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>177</td>
</tr>
<tr>
<td>Double</td>
<td>10</td>
</tr>
<tr>
<td>Triple</td>
<td>2</td>
</tr>
<tr>
<td>Location b</td>
<td></td>
</tr>
<tr>
<td>ACA territory</td>
<td>54</td>
</tr>
<tr>
<td>MCA territory</td>
<td>31</td>
</tr>
<tr>
<td>ICA territory</td>
<td>107</td>
</tr>
<tr>
<td>PCA territory</td>
<td>5</td>
</tr>
<tr>
<td>Other</td>
<td>6</td>
</tr>
<tr>
<td>Size (mm)b</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>97</td>
</tr>
<tr>
<td>3</td>
<td>61</td>
</tr>
<tr>
<td>4</td>
<td>29</td>
</tr>
<tr>
<td>≥5</td>
<td>16</td>
</tr>
</tbody>
</table>

Note:—ACA indicates anterior cerebral artery; PCA, posterior cerebral artery.

a Data are number of subjects.

b Data are number of lesions.

rysm size is shown in Fig 3. The CAD system successfully presented 124 of 158 (78%) aneurysms of 2–3 mm and 42 of 45 (93%) aneurysms of ≥4 mm. The sensitivity for the larger aneurysms was significantly higher than that for the smaller aneurysms (P = .040, χ² test). The details of the 3 missed aneurysms of ≥4 mm were as follows: 1) a 4-mm right internal carotid aneurysm, which was detected as the fourth candidate but not presented as a top 3 candidate; 2) a 4-mm left internal carotid artery aneurysm, which was not detected; and 3) a 6-mm right internal carotid artery aneurysm, which was detected as the sixth candidate but not presented.

Performances of the Radiologists

With the final diagnosis as the reference standard, the aneurysm-based overall sensitivity of the 26 radiologists in their initial diagnoses was 64% (258 of the 406 independent interpretations regarding the 203 aneurysms; 95% CI, 0.59–0.68). The sensitivity for small (≤3 mm) aneurysms was 59% (185/316), and that for large aneurysms (≥4 mm) was 81% (73/90), showing a significant difference (P < .001, χ²-test). Other statistical details are shown in Fig 4.

During the study period, the radiologists changed their initial negative diagnosis to a positive one after seeing the CAD results in 26 cases. Among these, the final diagnosis was also positive in 24 (92%). In the remaining 2 cases, aneurysms were “noticed” by 1 radiologist with the aid of the CAD system but were dismissed in the final diagnosis after discussion with the other radiologist. Thus, the CAD system more frequently benefited the radiologists in terms of detection of additional aneurysms than being detrimental in terms of introducing overdiagnosis. The interpretation-based analysis showed that the specificities of the radiologists’ diagnoses before and after seeing the CAD results were 98.9% and 98.8%, respectively.

Of the 24 aneurysms in which the CAD system benefited the radiologists, 6 were in the anterior cerebral artery circulation, 11 were in the internal carotid artery circulation, 5 were in the middle cerebral artery circulation, and 2 were in the posterior circulation. Twenty-one aneurysms were small (≤3 mm), 2 aneurysms were 4 mm, and the remaining aneurysm was 5 mm, protruding medially from the cavernous portion of the left internal carotid artery. The use of the CAD system increased the number of aneurysms detected by 9.3% (from 258 to 282, P < .001 by using the McNemar test), giving an overall sensitivity of 69% in the post-CAD diagnosis (282/406; 95% CI, 0.65–0.74). However, true aneurysms were still not detected even after the radiologists saw the CAD results in 124 interpretations, including 90 interpretations in which the CAD system had presented true aneurysms but the radiologists failed to change their diagnoses. Figure 5 shows a summary of the relationship between the initial diagnoses and post-CAD diagnoses of the radiologists. In 3 cases, neither the 2 radiologists nor the CAD system found the aneurysms but they were detected by the radiologic technologist.

Of the 32 interpretations regarding the 16 aneurysms of ≥5 mm, aneurysms were detected in the initial diagnosis in 29 (91%) cases, detected in the post-CAD diagnosis in 30 (94%) cases, and remained undetected even with the aid of the CAD system in 2 (6%) interpretations (1 with a 5-mm left cavernous aneurysm and 1 with a 5-mm left internal carotid posterior communicating aneurysm protruding caudally).

The median time required for reviewing the CAD results and giving feedback was 16 seconds.

DISCUSSION

In the present study, we evaluated the usefulness of a CAD system for cerebral aneurysms in a routine clinical diagnosis environment. To the best of our knowledge, this is the first study to evaluate the impact of CAD of cerebral aneurysms on radiologists in a large general population. The CAD system showed a sensitivity (82%) comparable with that of radiologists (64%) in the detection of both small and large cerebral aneurysms. The number of detected aneurysms increased by 9.3%, while preserving specificity. The additional time required for checking the CAD results was short. Our findings suggest that there were certain benefits from CAD even for radiologists who interpreted source images from 3T scanners in routine clinical practice.

Hirai et al14 reported that under experimental conditions, radiologists benefited (increase of ≥ 20 units by a using 0–100 confidence rating scale) from the CAD results in 10% of positive aneurysms. Compared with the study by Hirai et al in which only MIP images from a 1.5T scanner were presented to the radiologists, our study used both source and reconstructed images obtained by using 3T scanners. This difference means that even without CAD, a relatively high sensitivity of radiologists was expected, which might limit the benefit obtained from our CAD system. In addition, the benefit from CAD may be underestimated in this study compared with the previous study because our data-collection method focused only on the pure detection failures of radiologists. In reviewing the CAD results, radiologists had the opportunity to reconsider the lesion candidates they had already noticed, possibly changing their level of confidence in the
However, collecting such confidence level data in a routine environment was considered impractical.

This study demonstrated that the specificity of radiologists was not considerably decreased by CAD. This finding implies that FP lesion candidates presented by the CAD system can be easily dismissed by radiologists. This finding is also consistent with that in the previous study. On the other hand, in our series, radiologists dismissed many lesion candidates presented by the CAD system that were true-positives thereafter in the final diagnosis—that is, despite the high sensitivity of our CAD, radiologists were far less affected by the CAD results than by human opinions. In at least some cases, obvious aneurysms were overlooked even with the aid of the CAD system. In our opinion, this low confidence in the CAD system is partly because our CAD system provided little qualitative information about its results and it produced many false-positive marks. In addition, the CAD system presented its results on only axial sections, which might have made the referencing process difficult. Further investigation is needed to develop more efficient methods of displaying results to realize the full potential of CAD.

The prevalence rate of cerebral aneurysms in our reference standard (7.0%) is apparently higher than previously reported figures from angiography studies (3.0%–6.0%) but not as high as the prevalence rate (8.4%) reported by Igase et al., in which only a 3T MR imaging scanner was used to detect cerebral aneurysms in the Japanese population. They suggested that the excellent resolution of 3T MR imaging and the appropriate use of the volume-rendered technique contributed to the detection of additional aneurysms that could not be detected by other modalities. As we previously noted, owing to the recent advancements in MRA technology, invasive procedures with the sole purpose of establishing the diagnosis of cerebral aneurysms are becoming less justified. Although the true nature of the discrepancy in prevalence rates between DSA and MRA is still inconclusive, we believe that the use of 3T scanners and the reading of both source and volume-rendered images provide the highest practically possible accuracy in our reference standard diagnosis.

Our study had some limitations. First, the sensitivity of the radiologists in our method should be regarded as a rough estimate. The reference standard diagnosis was not independent of the diagnosis of the observers being tested, and there was considerable interobserver disagreement between the 2 radiologists. Second, despite the 39-month study pe-
period, the number of positive aneurysms was small, owing to the relatively low prevalence of the disease. In particular, the experience regarding relatively larger aneurysms (≥4 mm) was not sufficient. We are still using this system in daily practice, and further knowledge should be accumulated in the future.

CONCLUSIONS

The computer-assisted diagnosis of cerebral aneurysms is feasible, and radiologists can detect more cerebral aneurysms by using the CAD system without a substantial decrease in specificity. Radiologists are less likely affected by true-positive CAD results compared with the opinion of a different radiologist in double-reading settings.

Disclosures: Soichiro Miki—UNRELATED: Payment for Lectures (including service on Speakers Bureau): AIZ, Bayer Yakuhin; Payment for Manuscript Preparation: Inner Vision, Medical Eye. Naoto Hayashi—RELATED: Grant: HIMEDIC,* Siemens Healthcare KK.† GE Healthcare Japan.†† Yokohiro Nomura—UNRELATED: Grants/Grants Pending: HIMEDIC,* Siemens Healthcare KK.†† Shouhei Hanaoka—UNRELATED: Research (Nos. 23700573, 25750182, 15K01325).† Takeharu Yoshikawa—UNRELATED: Payment for Lectures (including service on Speakers Bureau): Bayer Yakuhin, Daichi Sankyo Healthcare, Fuji Pharma.†* Money paid to the institution. †The Department of Computational Diagnostic Radiology and Preventive Medicine is an endowed division supported by HIMEDIC, Siemens Healthcare KK, and GE Healthcare Japan. However, these companies did not play any direct role in running this specific study, including the study design.

REFERENCES