J Digit Imaging (2012) 25:497–503
DOI 10.1007/s10278-011-9444-4

Computer-Aided Diagnosis for Detection of Lacunar Infarcts on MR Images: ROC Analysis of Radiologists' Performance

Yoshikazu Uchiyama · Takahiko Asano · Hiroki Kato · Takeshi Hara · Masayuki Kanematsu · Hiroaki Hoshi · Toru Iwama · Hiroshi Fujita

Published online: 4 January 2012

Abstract The purpose of this study was to retrospectively evaluate radiologist performance in detection of lacunar infarcts on T_1- and T_2-weighted images, without and with the use of a computer-aided diagnosis (CAD) scheme. Thirty T_1-weighted and 30 T_2-weighted MR images obtained from 30 patients were used for assessing observer performance. These images were acquired using the fast spin-echo sequence with a 1.5-T MR imaging scanner. The group included 15 patients (age range, 48–83 years; mean age, 67.2 years; 10 men and five women) with a lacunar infarct and 15 patients (age range, 39–76 years; mean age, 64.0 years; eight men and seven women) without lacunar infarcts. Nine radiologists participated in the study. The radiologists initially interpreted the T_1- and T_2-weighted images without and then with the use of CAD, which indicated their confidence levels regarding the presence (or absence) of lacunar infarcts and the most likely position of a lesion on each MR scan. The observers' performance without and with the computer output was evaluated by performing receiver operating characteristic analysis. For the nine radiologists, the mean area under the best-fit binormal receiver operating characteristic curve plotted for unit square values of radiologists who interpreted the images without and with the scheme were 0.891 and 0.937, respectively. The performance of the radiologists improved significantly when they used the computer output (p=0.032). The CAD scheme has potential to improve the accuracy of radiologists' performance in detection of lacunar infarcts.

Keywords Lacunar infarct · Magnetic resonance (MR) · Computer-aided diagnosis (CAD) · Observer study · Receiver operating characteristic (ROC)

Y. Uchiyama (✉)
Department of Computer and Control Engineering,
Oita National College of Technology,
1666 Maki,
Oita City 870-0512, Japan
e-mail: uchiyama@oita-ct.ac.jp

T. Asano · H. Kato · M. Kanematsu · H. Hoshi
Department of Radiology, Graduate School of Medicine,
Gifu University,
Gifu, Japan

T. Hara · H. Fujita
Department of Intelligent Image Information,
Graduate School of Medicine, Gifu University,
Gifu, Japan

T. Iwama
Department of Neurosurgery, Graduate School of Medicine,
Gifu University,
Gifu, Japan

Introduction

Stroke is the third leading cause of death in the USA and is one of the most expensive conditions to treat. Each year, in the USA alone, approximately 780,000 strokes are recorded, of which 180,000 are recurrent [1]. Asymptomatic ischemic cerebrovascular disorders are an early manifestation of cerebrovascular diseases. Aging and hypertension are considered to be independent risk factors for silent lacunar infarcts [2]. Therefore, a health check system for early detection of asymptomatic cerebral and cerebrovascular diseases, which is called *brain dock* or *brain check-up*, is widely used in Japan. In this screening system, asymptomatic lacunar infarcts are often detected on magnetic resonance (MR) images. The detection of asymptomatic lacunar infarcts is important because their presence indicates an increased risk of severe cerebral infarction [3, 4]. However, accurate identification of lacunar infarcts on MR images is often difficult for radiologists because of the difficulty in distinguishing infarcts from other lesions, such as enlarged Virchow-Robin spaces [5]. Therefore, we have developed a computer-aided diagnosis (CAD) scheme for the detection of lacunar infarcts on T_1- and T_2-weighted images in order to assist radiologists who use computer outputs as a guide for image interpretation [6–8].

Our CAD scheme uses both T_1- and T_2-weighted images and accurately detects almost all lacunar infarcts. The current performance level of our CAD scheme for the detection of lacunar infarcts is 96.8% in sensitivity with the number of false positives (FP) of 0.76 per slice [7]. By using our CAD scheme, the radiologist confidence in detecting lacunar infarcts may increase considering the outputs of our CAD scheme because our CAD scheme can accurately detect almost all lacunar infarcts. However, the problem is that a relatively large number of FPs are also detected when using this method. For any CAD scheme, it is important to minimize the number of FPs. However, the majority of FPs of lacunar infarcts detected by the computer may be different from those detected by radiologists, and therefore, radiologists may be able to disregard these obvious FPs identified by the computer. Thus, an important question regarding the output value of our CAD scheme is whether radiologists can improve their performance when the computer output is presented to them. During the last two decades, a large number of observer performance studies have been carried out to answer to this question, and these have clearly shown that radiologist performance can be improved when computer outputs are employed [9–19]. However, the majority of CAD researches in the past focused on three body parts, i.e., chest, breast, and colon [20].

Recently, the concept of using CAD has been extended to the cerebral region. For example, CAD schemes for detection of unruptured intracranial aneurysms have been reported [21–25]. An observer performance study was also conducted, and the results indicated that the receiver operating characteristic (ROC) curve for radiologists in the detection of unruptured intracranial aneurysms in MR angiography improved when computer outputs were available [26]. However, to the best of our knowledge, there have been no previous reports of observer performance studies investigating the effect of a CAD scheme for the detection of lacunar infarcts. Therefore, the purpose of this study was to evaluate the effect of a CAD scheme on radiologist performance in detecting lacunar infarcts on T_1- and T_2-weighted images.

Materials and Methods

Database

Our image database was composed of 1,143 T_1-weighted and 1,143 T_2-weighted MR images obtained from 132 patients (mean age, 63.4 years; age range, 28–83 years). These images were acquired using a 1.5-T MR scanner (Signa Excite Twin Speed 1.5 T; GE Medical Systems, Milwaukee, WI) at Gifu University Hospital (Gifu, Japan). The T_1- and T_2-weighted images were obtained using the fast spin-echo sequence with an effective echo time (TE) of 8–12 ms, and 96–105 ms, respectively, and a repetition time of 300–500 ms and 3,000–3,500 ms, respectively. All the MR images were obtained in the axial plane with a section thickness of 5 mm, and an intersection gap of 2 mm, and a 24×24-cm field of view, which covered the entire brain. The matrix size of the MR images was 512×512, with a spatial resolution of 0.47 mm/pixel. Our institutional review board approved the use of the image database for this study.

All the T_1- and T_2-weighted images in our database were reviewed by two experienced neuroradiologists (T. A. and H. K., having 11 and 9 years of post-training experience in interpreting brain MR images, respectively). The locations of the lacunar infarcts (diameter, 3–15 mm) on T_1- and T_2-weighted images were determined by both neuroradiologists. The two radiologists independently marked the lesions they identified as lacunar infarcts. The lesions marked as lacunar infarct by both the radiologists were considered as “lacunar infarcts” for this study.

Fifteen of the 132 patients with a lacunar infarct were sequentially selected for an observer study. They included 10 men and five women with an age range of 48–83 years (mean, 67.2 years). Since multiple lacunar infarcts are often detected in individual patients, we simplified this study by imposing the criterion that the selected patients should have only one lacunar infarct. This selection criterion would allow easier interpretation of the results of ROC analysis. Further, 15 of the 132 patients without lacunar infarcts were selected as a control group for this study. They included eight men and seven women with an age range of 39–76 years (mean, 64.0 years). The mean age of the two groups was similar as determined by Students t test ($p=0.36$). The diagnosis of “no lacunar infarcts” was confirmed by consensus between the two neuroradiologists.

CAD Scheme

The scheme used in this study for computerized detection of lacunar infarcts on T_1- and T_2-weighted images included several steps [7]. First, we segmented the cerebral region on the basis of the region-growing technique in order to avoid detecting false findings located outside the cerebral region. For identifying the initial candidates for lacunar infarcts, top-hat transformation and multiple-phase binarization techniques were applied to the T_2-weighted image within the segmented cerebral region. Top-hat transformation was employed for enhancing both isolated lacunar infarcts and lacunar infarcts adjacent to the lateral ventricle. By using

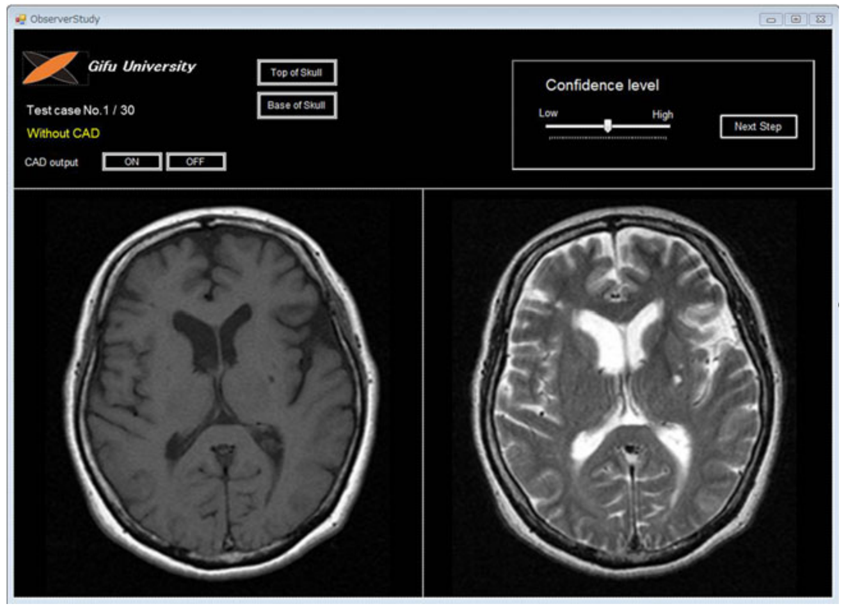

Fig. 1 Example of the computer display used in the current study

top-hat transformation, the extraction of lacunar infarcts adjacent to the lateral ventricle was simplified using a thresholding technique. In multiple-phase binarization, thresholding techniques with several threshold values were applied to the T_2-weighted images following white top-hat transformation. The initial candidates for lacunar infarcts were determined by integrating the gravity centers of all candidates detected by multiple-phase binarization. By using the techniques described above, almost all lacunar infarcts were accurately detected. However, the candidates selected initially also included many FPs. For eliminating these, we determined 12 features, namely, the *x* and *y* locations, signal intensity differences in the T_1- and T_2-weighted images, nodular components (NCs) from a scale of 1 to 4, and nodular and linear components (NLCs) from a scale of 1 to 4. The *x* and *y* locations were defined on the basis of the center of gravity in the candidate regions. The signal intensity differences on the T_1- and T_2-weighted images were determined by the difference between the average pixel value of the lacunar infarct region and the average pixel value of the peripheral region. The NCs and NLCs were obtained using a filter bank technique [27]. Finally, rule-based schemes and a support vector machine with the 12 features were applied to the initial candidates in order to distinguish between lacunar infarcts and FPs. The performance level used in this study for detection of lacunar infarcts indicated that the sensitivity was 96.8% with 0.76 FPs per slice [7].

Observer Performance Study

A total of nine radiologists, including three board-certified radiologists who specialized in neuroradiology and six board-certified radiologists who did not specialize in neuroradiology, took part in the observer performance study. We used a sequential method in the study [15]. T_1- and T_2-weighted images were displayed together at the same transverse location. The observers could manually control the speed or sequence of the slice image display, and they were allowed to change the window level and width on the monitor. Each observer initially read all slice images for T_1- and T_2-weighted images displayed on the LCD monitor without computer output. The observer marked his or her confidence level regarding the likelihood of the presence of a lacunar infarct. After the observer marked the initial level of confidence, the computer outputs were superimposed on the T_1- and T_2-weighted images. Observers again marked their confidence level if they wished to change their initial result. Figure 1 illustrates the display used in the observer study.

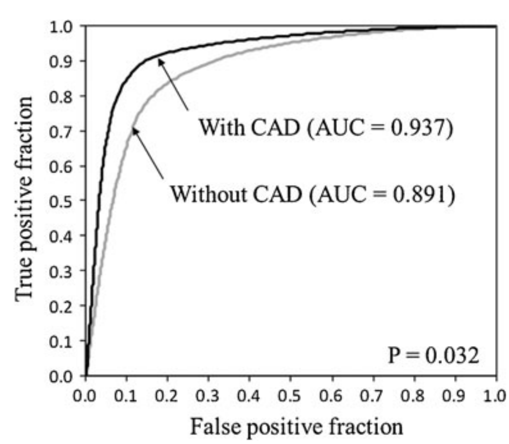

Fig. 2 Average ROC curves obtained from nine radiologists for the detection of lacunar infarcts without and with the use of computer outputs. The average AUC value significantly improved from 0.886 to 0.930 when observers used the computer output (p=0.032)

Table 1 AUC value for radiologists in the detection of lacunar infarcts

Observers	AUC value	
	Without CAD	With CAD
A	0.676	0.702
B	0.878	0.938
C	0.957	0.980
D	0.930	0.984
E	0.957	0.960
F	0797	0.959
G	0.900	0.962
H	0.926	0.965
I	0.995	0.981
Mean for all radiologists	0.891	0.937

The observers were given the following information: (a) The purpose of the study was to evaluate the performance of radiologists in detecting lacunar infarcts without and with the CAD scheme on T_1- and T_2-weighted images. (b) The role of the CAD output as a "second opinion" would be evaluated. (c) The observer study consisted of 30 MRI studies that did not or did contain a lacunar infarct and/or non-lacunar lesions such as enlarged Virchow-Robin spaces. (d) The defined diameter of "lacunar infarcts" was 3–15 mm. (e) Computer performance yielded a sensitivity of 96% and 0.76 FP per slice on average, and this result was not obtained from the 30 cases in this observer study. (f) The observers were instructed to click on the screen using a mouse (1) to indicate on a bar their confidence level regarding the presence (or absence) of a lacunar infarct, and (2) to identify the most likely infarct position in each case. Each observer used a continuous rating scale displayed on the monitor. The observers were blinded to the number of patients with a lacunar infarct. The cases selected for study were presented in the same randomized order to each observer. There was no limit on reading time.

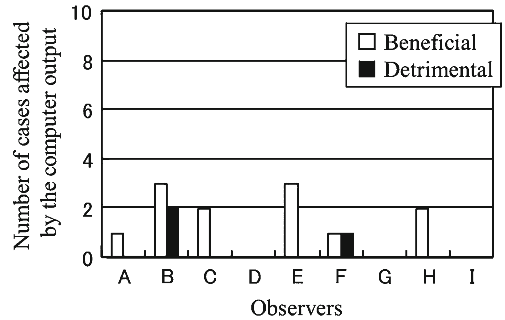

Fig. 4 Graph shows the number of cases (>15%) in which confidence levels were affected by the CAD output in cases with a lacunar infarct

Statistical Analysis

ROC analysis [28] was used for comparison between radiologist performances without and with computer outputs for detection of lacunar infarcts on T_1- and T_2-weighted images. A binormal ROC curve was fitted to each radiologist's confidence rating data from the two reading conditions with quasi-maximum likelihood estimation [29]. The computer program PROPROC [30] was used for obtaining the binormal ROC for each radiologist. The statistical significance of the differences between ROC curves obtained without and with computer outputs was tested using the computer program DBM-MRMC [31, 32], which uses analysis of variance in pseudo-values of the area under the best-fit binormal ROC curve (AUC) calculated from all rating scores of all radiologists. In all analyses, a p value of less than 0.05 was considered to indicate a significant difference.

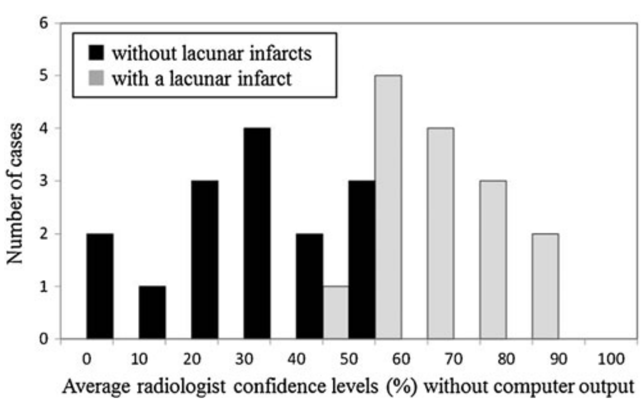

Fig. 3 Histograms of average radiologist confidence levels (%) without the use of computer outputs. Data were obtained from 15 cases with a lacunar infarct and 15 cases without lacunar infarcts

Results

Figure 2 shows ROC curves for the detection of lacunar infarcts on T_1- and T_2-weighted images obtained from all

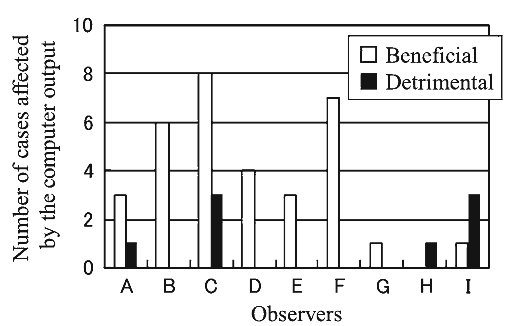

Fig. 5 Graph shows the number of cases (>15%) in which confidence levels were affected by the CAD output in cases without a lacunar infarct

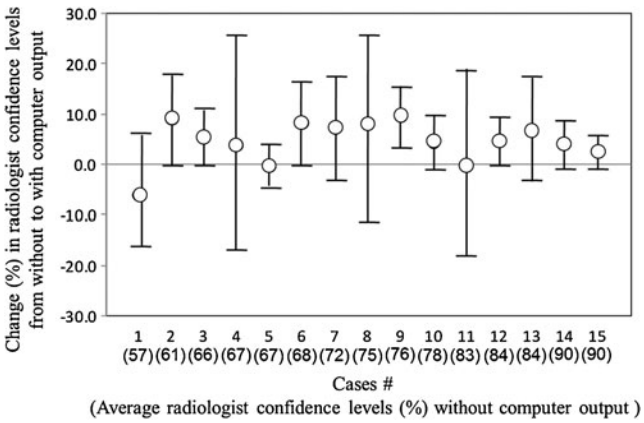

Fig. 6 Graph shows the relationship between average radiologist confidence levels (%) without computer output, and the average change in the radiologist confidence levels without and with the use of computer outputs. Data were obtained from 15 cases with a lacunar infarct

the radiologists without and with computer outputs. The average AUC values of all the radiologists improved from 0.891 (without the computer output) to 0.937 (with the computer output), and this difference was statistically significant (p=0.032). Table 1 shows the AUC values without and with the computer output for each radiologist.

Figure 3 shows the histograms of average radiologist confidence levels (%) without the use of computer outputs. Data were obtained from 15 cases with a lacunar infarct and 15 cases without lacunar infarcts. Since several cases had a subtlety rating, identification of lacunar infarcts was difficult in them.

Figure 4 shows the clinically relevant changes in the confidence ratings of each observer with regard to cases with a lacunar infarct. Clinically relevant change was defined as a case (>15%) affected by CAD output. The average number of cases affected beneficially was 1.33 (8.9%). However, the average number of cases affected detrimentally was 0.33 (2.2%). In the study, each observer indicated (1) their confidence level regarding the presence (or absence) of a lacunar infarct, and (2) the most likely position of the lesion in each case. Using these data, a neuroradiologist (one of the authors) and the first author reviewed the cases to determine the nature of the errors. In two out of the three detrimentally affected cases, our CAD scheme could not accurately detect the lacunar infarct. In one of the three detrimentally affected cases, the observer changed his/her confidence level to a lower value, even though the CAD scheme accurately detected the lacunar infarct. Figure 5 shows clinically relevant changes in the confidence ratings of each observer in cases without a lacunar infarct. The average number of cases affected beneficially and detrimentally were 3.67 (24.4%) and 0.89 (5.9%), respectively. The average number of patients without a lacunar infarct that were affected beneficially was higher than that of patients with a lacunar infarct. These beneficial effects were caused by the fact that observers initially marked lesions that appear similar to lacunar infarcts, such as enlarged Virchow-Robin spaces, with a relatively high confidence level, indicating the presence of lacunar infarcts. However, the confidence level was changed to a lower value after taking into account the absence of lesions in CAD. Therefore, observers were able to make the correct diagnosis using the computer output. On the other hand, eight detrimentally affected cases were caused by FPs detected by our CAD scheme.

Discussion

Our results indicate that the performance of eight out of the nine radiologists improved when they used the computer output. However, the performance of one observer did not improve when the computer output was employed. It should be noted that the AUC value of this observer (radiologist I) without the computer output was quite high. The reason for the decrease in the AUC value upon use of the computer output was that this observer showed higher sensitivity than our CAD scheme in detecting lacunar infarcts. Therefore,

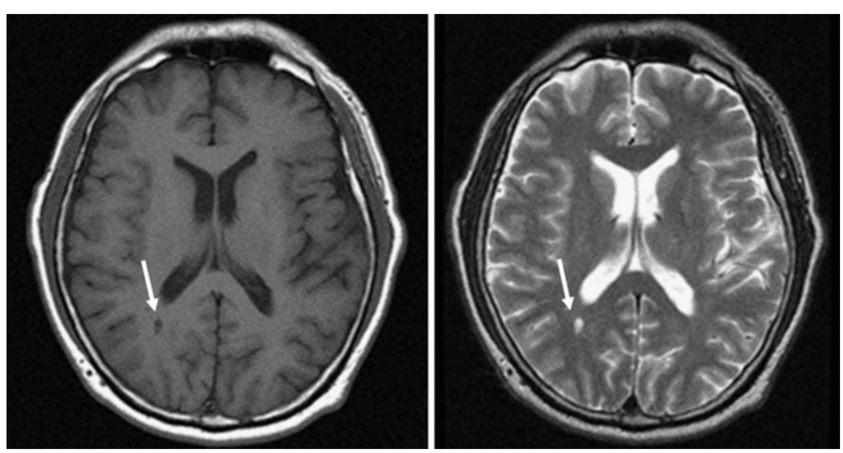

Fig. 7 T_1- and T_2-weighted images from case no. 1 in Fig. 6. Our CAD scheme could not accurately detect this lacunar infarct

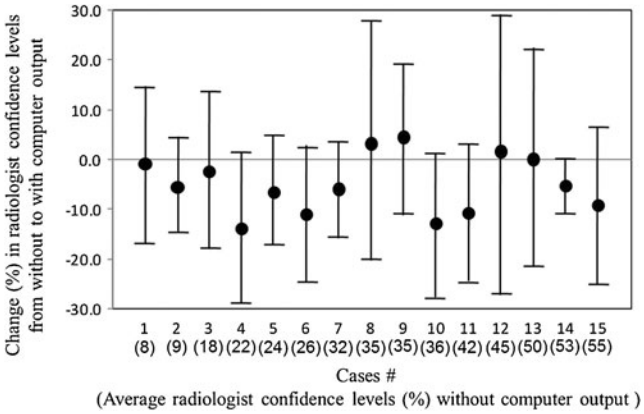

Fig. 8 Graph shows the relationship between average radiologist confidence levels (%) without computer output, and the average change in the radiologist confidence levels without and with the use of computer outputs. Data were obtained from 15 cases without lacunar infarcts

the radiologists' confidence in detecting lacunar infarcts could have decreased when the outputs of our CAD scheme were considered. As shown in Fig. 4, the detrimentally affected cases were caused by FPs in our CAD scheme. If this observer was to use an improved CAD scheme with a lower FP rate for the detection of lacunar infarcts, the AUC value might improve. Therefore, in the future, we need to improve our method for further elimination of FPs.

It is also important to know which cases are affected beneficially or detrimentally without and with the computer output. We investigated the changes in the radiologist confidence levels without and with the use of computer outputs (Fig. 6). These values were obtained from 15 cases with a lacunar infarct. Almost all the cases were affected beneficially when using the computer outputs. However, one case (case no. 1) was affected detrimentally because, in this case, our CAD scheme could not detect the lacunar infarct. Figure 7 shows the T_1- and T_2-weighted images obtained from case no. 1 (Fig. 6). Although all the observers marked this lacunar infarct, three observers changed their confidence levels to lower values after using the computer output. Hence, this error is not considered as a "detection error" but rather as an "interpretation error," because three observers accurately detected the lacunar infarct. We also investigated 15 cases without lacunar infarcts. Figure 8 shows the average percentage change in the radiologist confidence levels without and with the use of computer outputs. While many cases were affected beneficially by using computer outputs, three cases were affected detrimentally. Figure 9 shows the T_1- and T_2-weighted images obtained from case no. 9 (Fig. 8). Our CAD scheme yields a relatively large number of FPs. These FPs can be classified into four types: a part of the cerebral sulcus, a part of the cerebral ventricle, enlarged Virchow-Robin spaces, and others [7]. The majority of FPs were different from those detected by radiologists and therefore radiologists can disregard these obvious FPs identified by the computer. However, it is noteworthy that some FPs detected by the computer were difficult for assisting radiologists in distinguishing between lacunar infarcts. If the radiologists were to be strongly influenced by these FPs, unnecessary medical treatments might be prescribed for the patient.

Conclusions

Our results provide evidence of a potential benefit of CAD in the detection of lacunar infarcts on T_1- and T_2-weighted images. However, it is also obvious that there are potential pitfalls associated with the use of computer outputs, in that false negatives and FPs identified by computers may increase interpretation errors by radiologists. Therefore, in order to decrease interpretation errors in the future, we need to develop a CAD scheme for distinguishing between enlarged Virchow-Robin spaces and lacunar infarcts, and to combine two independent CAD schemes for the detection and classification of lacunar infarcts. Moreover, a limitation of this observer study was in stipulating a criterion whereby

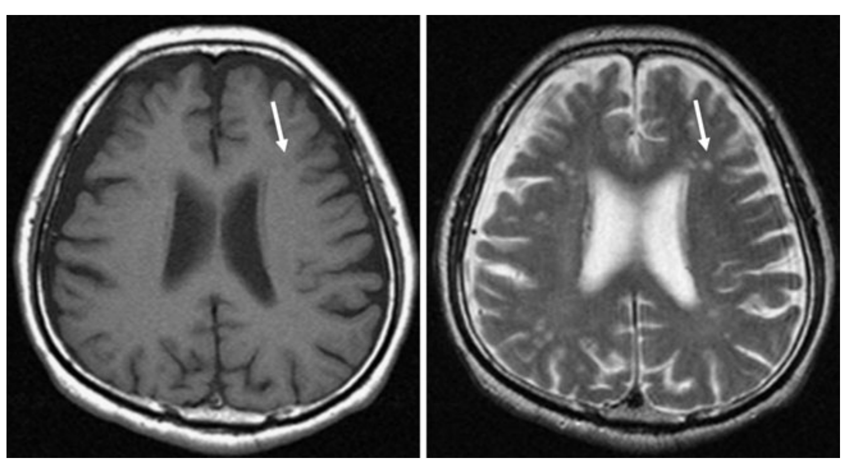

Fig. 9 T_1- and T_2-weighted images from case no. 9 in Fig. 8. Our CAD scheme considered this lesion as a lacunar infarct

patients with only one lacunar infarct were included. Therefore, further investigation of the use of our CAD scheme in patients with several lacunar infarcts is also required.

Acknowledgments The authors would like to thank Satoshi Honda, MD, Nobumaro Ihara, MD, Kota Sakurai, MD, Toshihisa Kojima, MD, Kouzou Kuwabara, MD, Ishikawa, MD, Yusuke Tuge, MD, Oikawa, MD, and Yasuo Sasaki, MD, for their participation in the observer study; Junji Shiraishi, PhD, and Satoshi Kasai, PhD, for their helpful comment on the ROC analysis; Charles E. Metz for using the ROC analysis software. This work was partly supported by a grant for the "Knowledge Cluster Initiative: Gifu Ogaki Robotics Advanced Medical Cluster" from the Ministry of Education, Culture, Sports, Science, and Technology, Japan.

References

1. Rosamond W, Flegal K, Furie K, et al: Heart disease and stroke statistics—2008 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Circulation 117:e25–e146, 2008
2. Shintani S, Shiigai T, Arinami T: Silent lacunar infarction on magnetic resonance imaging (MRI): risk factors. J Neurol Sci 160:82–86, 1998
3. Kobayashi S, Okada K, Koide H, et al: Subcortical silent brain infarction as a risk factor for clinical stroke. Stroke 28:1932–1939, 1997
4. Vermeer SE, Hollander M, Dijk EJ, et al: Silent brain infarcts and white matter lesions increase stroke risk in the general population: the Rotterdam scan study. Stroke 34:1126–1129, 2003
5. Bokura H, Kobayashi S, Yamaguchi S: Discrimination of silent lacunar infarction from enlarged Virchow-Robin spaces on brain magnetic resonance imaging: clinicopathological study. J Neurol 245:116–122, 1998
6. Yokoyama R, Zhang X, Uchiyama Y, et al: Development of an automated method for detection of chronic lacunar infarct regions on brain MR images. IEICE Trans. Inf. & Syst. E90-D(6):943–954, 2007
7. Uchiyama Y, Yokoyama R, Ando H, et al: Computer-aided diagnosis scheme for detection of lacunar infarcts on MR image. Acad Radiol 14(12):1554–1561, 2007
8. Uchiyama Y, Yokoyama R, Asano T, et al: Improvement of automated detection method of lacunar infarcts in brain MR images. Conf. Proc. of IEEE Engineering in Medicine and Biology 2007:1599–1602, 2007
9. Free TM, Ulissey MJ: Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast cancer. Radiology 220:781–786, 2001
10. Karssemeijer N, Otten JD, Verbeek AL, et al: Computer-aided detection versus independent double reading of masses on mammograms. Radiology 227:192–200, 2003
11. Gur D, Sumkin JH, Rockette HE, et al: Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system. J Natl Cancer Inst 96(3):185–190, 2004
12. Birdwell R, Bandodkar P, Ikeda D: Computer aided detection (CAD) with screening mammography in a university hospital setting. Radiology 236:451–457, 2005
13. Morton MJ, Whaley DH, Brandt KR, et al: Screening mammograms: interpretation with computer-aided detection prospective evaluation. Radiology 239:375–383, 2006
14. Dean JC, Ilvento CC: Improved cancer detection using computer-aided detection with diagnostic and screening mammography: prospective study of 104 cancers. AJR 187:20–28, 2006
15. Kobayashi T, Xu XW, MacMahon H, et al: Effect of a computer-aided diagnosis scheme on radiologists' performance in detection of lung nodules on radiographs. Radiology 199:843–848, 1996
16. Shiraishi J, Abe H, Engelmann R, et al: Computer-aided diagnosis to distinguish benign from malignant solitary pulmonary nodules on radiographs: ROC analysis of radiologists' performance—initial experience. Radiology 227:469–474, 2003
17. Li F, Arimura H, Suzuki K, et al: Computer-aided detection of peripheral lung cancers missed at CT: ROC analyses without and with localization. Radiology 237:684–690, 2005
18. Baker ME, Bogoni L, Obuchowski NA, et al: Computer-aided detection of colorectal polyps: can it improve sensitivity of less-experienced readers? Preliminary findings. Radiology 245:140–149, 2007
19. Mang T, Reloschek P, Plank C, et al: Effect of computer-aided detection as a second reader in multidetector-row CT colonography. Eur Radiol 17:2598–2607, 2007
20. Doi K: Computer-aided diagnosis in medical imaging: historical review, current status and future potential. Comput Med Imaging Graph 31:198–211, 2007
21. Hayashi N, Masutani Y, Matsumoto T, et al: Feasibility of curvature-based enhanced display system for detecting cerebral aneurysms in MR angiography. Magnetic Resonance in Medical Science 2:29–36, 2003
22. Arimura H, Li Q, Korogi Y, et al: Automated computerized scheme for detection of unruptured intracranial aneurysms in three-dimensional MRA. Acad Radiol 11(10):1093–1104, 2004
23. Uchiyama Y, Ando H, Yokoyama R, et al: Computer-aided diagnosis scheme for detection of unruptured intracranial aneurysms in MR angiography. Conf. Proc. of IEEE Engineering in Medicine and Biology 3:3031–3034, 2005
24. Kobayashi S, Kondo K, Hata Y: Computer-aided diagnosis of intracranial aneurysms in MRA images with case-based reasoning. IEICE Trans. Inf. & Syst. E89-D(1):340–350, 2006
25. Arimura H, Li Q, Korogi Y, et al: Computerized detection of intracranial aneurysms for three-dimensional MR angiography: feature extraction of small protrusions based on a shape-based difference image technique. Medical Physics 33(2):394–401, 2006
26. Hirai T, Korogi Y, Arimura H, et al: Intracranial aneurysms at MR angiography: effect of computer-aided diagnosis on radiologists' detection performance. Radiology 237:605–610, 2006
27. Nakayama R, Uchiyama Y, Yamamoto K, et al: Computer-aided diagnosis scheme using a novel filter bank for detection of clustered microcalcifications in mammograms. IEEE Trans. on Biomedical Engineering 53(2):273–283, 2006
28. Metz CE: ROC methodology in radiologic imaging. Invest Radiol 21:720–733, 1986
29. Metz CE, Herman BA, Shen JH: Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. Stat Med 17:1033–1053, 1998
30. Pan X, Metz CE: The "proper" binormal model: parametric receiver operating characteristic curve estimation with degenerate data. Acad Radiol 4:380–389, 1997
31. Hills SL, Obuchowski NA, Schartz KM, et al: A comparison of the Dorfman-Berbaum-Metz and Obuchowski-Rockette methods for receiver operating characteristic (ROC) data. Stat Med 24:1579–1607, 2005
32. Hills SL, Berbaum KS: Monte Carlo validation of the Dorfman-Berbaum-Metz method using normalized pseudo values and less data-based model simplification. Acad Radiol 12:534–1541, 2005