ORIGINAL ARTICLE

Inter-Observer Agreement on Diffusion-Weighted Magnetic Resonance Imaging Interpretation for Diagnosis of Acute Ischemic Stroke Among Emergency Physicians

Deniz ORAY,[1] Onder LIMON,[1] Cem ERTAN,[1] Asli AYDINOGLU UGURHAN,[1] Erkan SAHIN[2]

[1]Department of Emergency Medicine, Izmir University School of Medicine, Izmir
[2]Department of Radiodiagnostics, Izmir University School of Medicine, Izmir, both in Turkey

SUMMARY

Objectives

Diffusion-weighted magnetic resonance imaging (DW-MRI) is a highly sensitive tool for the detection of early ischemic stroke and is excellent at detecting small and early infarcts. Nevertheless, conflict may arise and judgments may differ among different interpreters. Inter-observer variability shows the systematic difference among different observers and is expressed as the kappa (κ) coefficient. In this study, we aimed to determinate the inter-observer variability among emergency physicians in the use of DW-MRI for the diagnosis of acute ischemic stroke.

Methods

Cranial DW-MRI images of 50 patients were interpreted in this retrospective observational cross-sectional study. Patients who were submitted to DW-MRI imaging for a suspected acute ischemic stroke were included in the study, unless the scans were ordered by any of the reviewers or they were absent in the system. The scans were blindly and randomly interpreted by four emergency physicians. Inter-observer agreement between reviewers was evaluated using Fleiss' κ statistics.

Results

The mean kappa value for high signal on diffusion-weighted images (DWI) and for reduction on apparent diffusion coefficient (ADC) were substantial (k=0.67) and moderate (k=0.60) respectively. The correlation for detection of the presence of ischemia and location was substantial (k: 0.67). There were 18 false-positive and 4 false-negative evaluations of DWI, 15 false positive and 8 false-negative evaluations of ADC.

Conclusions

Our data suggest that DW-MRI is reliable in screening for ischemic stroke when interpreted by emergency physicians in the emergency department. The levels of stroke identification and variability show that emergency physicians may have an acceptable level of agreement.

Key words: Emergency department; diffusion weighted magnetic resonance imaging; inter-observer agreement, ischemic stroke.

Introduction

Clinical diffusion neuroimaging, introduced in the early 1990s, was quickly adopted in the evaluation of suspected acute ischemic brain injury.[1] Diffusion-weighted magnetic resonance imaging (DW-MRI) is a highly sensitive tool for the detection of early changes in water diffusion that characterize many brain pathologies, including acute ischemic stroke and is excellent at detecting small and early infarcts. These changes represent variations in the random motion of water molecules in tissues. They are expressed, in diffusion-weighted images (DWI), as changes in MRI signal intensity or as variations in the apparent diffusion coefficient (ADC) of water.[2–5] Acute ischemic stroke is characterized by very high signal on DWI and marked reduction in ADC values. In stroke

Submitted: August 09, 2014 **Accepted:** October 10, 2014 **Published online:** February 17, 2015

Correspondence: Deniz ORAY, M.D. Izmir Universitesi Tip Fakultesi, Izmir Medical Park Hastanesi Acil Tip Anabilim Dali. Izmir, Turkey.
e-mail: deniz.oray@yahoo.com

patients, reduced diffusion can be observed within minutes to less than 1 hour after witnessed ictus, before any findings are apparent on conventional MRI.[5] The appearance of DWI/ADC depends on the timing. Acute (0–7 days) findings of ischemic stroke in DWI are; a decrease in ADC value with maximal signal reduction at 1 to 4 days (hypo-intensity on ADC images), marked hyperintensity on DWI (a combination of T2 and diffusion weighting), and less hyperintensity on exponential images.[5,6]

DW-MRI can show hyperacute ischemic stroke which cannot be seen on computed tomography (CT); moreover, it only takes few minutes to scan and should be considered when the emergency physician evaluates a patient with acute ischemic stroke.[7] Lövblad et al reported a sensitivity of 88% and a specificity of 95% for DW-MRI.[8]

As high as these values are, there are also reports of negative DW-MRI studies in cases of clinically proven ischemic stroke.[9,10] Sylaja et al reported a 25% false negative report rate of DW-MRI in stroke and stroke like deficits.[11] Additionally, as in all imaging modalities, DW-MRI is also interpreter dependent to some degree, especially when evaluated by non-radiologist interpreters. Inter-observer variability is the reflection of this dependency and is defined as the systematic differences among different observers and expressed as the kappa (κ) coefficient.[12]

The primary aim of this study was to determine the inter-observer variability among emergency physicians in the use of DW-MRI for the diagnoses of acute ischemic strokes.

Materials and Methods

This retrospective observational cross-sectional study was performed in xxx University xxx Hospital. Images selected from all cranial DW-MRI examinations that were referred to the Radiology Department of xxx University School of Medicine with suspected acute ischemic stroke from 01.06.2013-01.01.2014. Local ethics committee approval was obtained prior to data collection. Data were obtained from Medin© v3.1.24.115 software. The inclusion criterion was suspected acute ischemic stroke in patients who were submitted to DW-MRI imaging. The exclusion criterion was absence of DW-MRI imaging or DW-MRI scans ordered by any of the reviewers. The scans were interpreted by four emergency physicians separately in a blind and random fashion under emergency departments (ED) conditions, where actual cases were being evaluated. Reviewers were unaware of the official radiology report and patients' clinical status at the time of imaging. All reviewing physicians were emergency physicians and have the designation of assistant professor and more than 5 years' experience in their specialty. The reviewers were trained by the radiology department for interpretation of DWI-MRI for two hours. Also, the emergency medicine residency program in xxx have at least a one month mandatory radiology rotation. Most of the hospitals do not have a radiologist and neurologist for 24 hours in xxx, therefore many emergency physicians have to interpret their patients' scans and activate the stroke protocol themselves according to the local hospital protocols.

The following parameters were taken into consideration:

a- High signal on DWI

b- Marked reduction on ADC values

c- Location of these findings

The reviewers noted on the mentioned parameters as present or absent. Inter-observer agreement between reviewers was evaluated using Fleiss' κ statistics. The kappa coefficient measures pairwise agreement among a set of interpreters making category judgments, correcting for expected chance agreement.[13] The agreement on interpreter opinions are supposed to be acceptable beginning at a correlation co-efficient of 0.41–0.60 (Table 1).[14,15] Later, reviewers' interpretations were compared to official radiology reports; the radiologists had more than 10 year experience in cranial radiology.

MRI Protocol

All MRI scans were performed on a Siemens 1.5 Tesla Avanto© MRI scanner using 4-mm slice thickness; B-0, B-1000 mm^2/sec images and ADC.

Statistical Analysis

Kappa statistics on https://www.statstodo.com were used to evaluate the inter-observer agreement. StatsToDo© website provides free statistics calculators for clinical research and quality control.

Results

There were 97 patients who underwent DWI-MRI between 01.06.2013-01.01.2014. Images of 47 patients were ordered by one of the reviewers and these were excluded from the study, leaving 50 patients for further evaluation. The mean kappa value for high signal on DWI and for reduction on ADC were κ=0.67 (0.56-0.78) and κ=0.60 (0.49-0.71) respectively. The inter-observer agreement between reviewers was substantial for high signal on DWI and moderate for reduction on ADC respectively. Accuracy of diagnosis for reviewers one to four, according to the official radiologist's reports, were 88%, 90%, 80% and 86% for ischemic stroke (mean accuracy rate: 86%) respectively. Correlation between the reviewers was substantial (κ: 0.67) in detecting the presence of ischemia

Table 1. The interpretation of kappa value ranges[14]

Kappa	Interpretation (Level of agreement)
<0	Poor
0.01–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–1.00	Almost perfect

Table 2. Numbers of false-negative and false-positive comments for high signal on DWI and ADC respectively

	FN[1] DWI/ADC	FP[2] DWI[3]/ADC[4]
Observer 1	1/2	2/2
Observer 2	1/2	3/2
Observer 3	1/1	9/9
Observer 4	1/3	4/2

[1]FN: False negative; [2]FP: False positive; [3]DWI: Diffusion weighed images; [4]ADC: apparent diffusion coefficient.

and location. On the other hand, 14% of ischemic strokes related changes on DWI and ADC were missed by interpreters.

In total, observers recorded 18 false-positive high signals on DWI, 15 false positive marked reductions on ADC, 4 false-negative high signals on DWI and 8 false-negative marked reductions on ADC (Table 2). Six of the false-positive stroke lesions (33.3%) were located at the posterior cerebral region (5 at the brainstem and one at the cerebellum) and the rest were equally distributed. Of the 12 false-negative stroke lesions; 8 were located in the brainstem (66.7%), 2 were located in the cerebellum (16.3%), one was periventricular (8.3%) and one was at the occipital lobe (8.3%).

In one case, a large hematoma (10 cm) was defined by all observers, and it was interpreted to be a non-ischemic lesion despite the diffusion deficiency findings; in another, a cortical mass <1 cm was falsely interpreted as an ischemic lesion; all four observers misinterpreted a mass with surrounding edema as a positive finding for ischemia and an artifact was misinterpreted as lacunar infarct by all the reviewers. Since all observers expressed the same opinions about the same images, these misinterpretations did not affect the inter-observer agreement, which is the main objective of our study.

Our results also showed that a great majority of the false negative high signals in DWI and reduction in ADC were located at the posterior cerebral area or the brain stem.

Discussion

Ischemic stroke is the fourth leading cause of death in United States.[16] Diagnoses of stroke in the ED depend on clinical evaluation and imaging studies. Although most EDs use (CT) as the imaging modality of choice due to factors such as availability and low costs, it's a fact that 30% to 60% of ischemic lesions are not visible on CT in the acute stage of stroke.[3] DW-MRI is successfully being used in the evaluation of suspected acute ischemic brain injury due to the fact that reduced diffusion of brain tissue can be observed within the first hour following the ischemic event.[1] Gonzalez et al reported that DW-MRI is highly accurate for diagnosing stroke within 6 hours of symptom onset and is superior to CT and conventional MRI.[17] Recently, DW-MRI has been preferred over CT in many EDs due to the same reasons, and the literature suggests the benefits of this choice.[16]

Despite the well reported benefits of the DW-MRI, the interpretation of the imaging studies is still dependent on the reviewing physician, who is often not a radiologist, but an emergency physician practicing within the routine circumstances of an ED.[18] Although the diagnostic criteria for ischemia on DWI studies are well known, the effects of personal experience, differences in MRI devices, patients and even the screens which reviewers evaluate the images cannot be ignored.[19] Thus, the inter-observer agreement is a very important issue on this life threatening diagnosis. However, we could not find any studies on inter-observer variability for DW-MRI interpretation of ischemic brain injury at the ED in the literature. We therefore designed this study to determinate the inter-observer variability among emergency physicians in the use of DW-MRI for the diagnosis of acute ischemic strokes.

Inter-observer agreement gives a score of the homogeneity or consensus in the ratings given by the reviewers. The level of agreement is shown by means of kappa values, and classified as shown in Table 1.[12,15] In our study, the mean kappa values for high signal on DWI and reduction on ADC were found to be substantial and moderate respectively, meaning that four reviewers had an acceptable to high levels of understanding and agreement on ischemic lesions on DW-MRI.

The evaluation of radiologic work-up by non-radiologist physicians, although inevitably necessary, is continuously an issue of debate. The presence of false negative CT and DW-MRI studies is manifested in the literature by many researchers. In a study, Kothari et al proposed that misdiagnosis rates for ischemic stroke by emergency physicians ranged from about 5% to 33%.[19] In their study, Ferro et al reported a 9%

rate of misdiagnosis of stroke for non-neurologist physicians.[18] Savits et al also reported that DW-MRI and CT scans which were initially interpreted as negative in the ED, but were reported to be positive for ischemic findings by radiologists.[20] In our study, 14% of ischemic stroke related findings in DW-MRI were missed by emergency physicians.

Several case series have shown that false-negative DW-MRI results are not rare in brain stem strokes.[8,21–23] Sylaja et al also reported that 30% of their patients with negative DWI scans had either an imaging or clinical diagnosis of stroke located at the brain stem.[11] Our results also showed that a great majority of the false negative high signals in DWI and reductions in ADC were located at the posterior cerebral area or the brain stem.

There were four cases in which all observers misinterpreted some lesions or artifacts as ischemic findings. Although these false positive interpretations did not affect the overall inter-observer agreement, they also pointed the fact that some technical or non-ischemic pathology may mimic ischemia. Sylaja et al suggested that technical concerns including the magnetic susceptibility artifacts might cause misinterpretations.[11] Löveblad et al also reported the misinterpretation of a cerebral abscess and a cranial tumor as ischemic lesions in their study.[8] We observed in our study that false positive diffusion deficiency judgments were commonly expressed for posterior cerebral area. This may be due to the small volume of the area as well as a higher rate of artifacts.

Limitations

Four reviewers have different medical education backgrounds from different medical schools; therefore the qualitative levels of radiological evaluation based on this background may differ in perspective. The evaluation of the DW-MRI images were performed under ED circumstances to mimic original judgment conditions. This also might have negatively affected the reviewers' quality of decisions due to the level of illumination of the environment, distractive issues and the image quality of the monitors.

Conclusion

DW-MRI evaluation is an invaluable tool for ischemic stroke diagnosis in the ED. There is no question that diffusion-weighted imaging is an important diagnostic tool in stroke management. Emergency physicians have to be capable of identifying stroke findings in DW-MRI images for timely diagnosis and the initiation of the appropriate treatment. The levels of stroke identification and variability show that emergency physicians may have an acceptable agreement on DW-MRI interpretation for stroke. According to our data, DW-MRI seems to be a reliable method in screening for ischemic stroke when interpreted by emergency physicians in the ED. Our results also showed that false-negative DW-MRI studies are not uncommon, especially if findings are suggestive of a stroke in the posterior circulation and brain stem.

Conflict of Interest

The authors declare that there is no potential conflicts of interest.

References

1. Maas LC, Mukherjee P. Diffusion MRI: Overview and clinical applications in neuroradiology. Applied Radiology 2005;34:44–60.
2. Lutsep HL, Albers GW, DeCrespigny A, Kamat GN, Marks MP, Moseley ME. Clinical utility of diffusion-weighted magnetic resonance imaging in the assessment of ischemic stroke. Ann Neurol 1997;41:574–80. CrossRef
3. van Everdingen KJ, van der Grond J, Kappelle LJ, Ramos LM, Mali WP. Diffusion-weighted magnetic resonance imaging in acute stroke. Stroke 1998;29:1783–90. CrossRef
4. Srinivasan A, Goyal M, Al Azri F, Lum C. State-of-the-art imaging of acute stroke. Radiographics 2006;26 Suppl 1:75–95.
5. Allen LM, Hasso AN, Handwerker J, Farid H. Sequence-specific MR imaging findings that are useful in dating ischemic stroke. Radiographics 2012;32:1285–99. CrossRef
6. Chung SP, Ha YR, Kim SW, Yoo IS. Diffusion-weighted MRI as a screening tool of stroke in the ED. Am J Emerg Med 2002;20:327–31. CrossRef
7. Keyik B, Yanık B, Bakdık S, et al. Difüzyon ağırlıklı manyetik rezonans görüntülemenin akut iskemik inmede tanı değeri. (The diagnostic value of diffusion weighed magnetic resonance imaging in ischemic stroke) Yeni Tıp Dergisi 2009;26:42–5.
8. Lövblad KO, Laubach HJ, Baird AE, Curtin F, Schlaug G, Edelman RR, et al. Clinical experience with diffusion-weighted MR in patients with acute stroke. AJNR Am J Neuroradiol 1998;19:1061–6.
9. Wang PY, Barker PB, Wityk RJ, Uluğ AM, van Zijl PC, Beauchamp NJ Jr. Diffusion-negative stroke: a report of two cases. AJNR Am J Neuroradiol 1999;20:1876–80.
10. Lefkowitz D, LaBenz M, Nudo SR, Steg RE, Bertoni JM. Hyperacute ischemic stroke missed by diffusion-weighted imaging. AJNR Am J Neuroradiol 1999;20:1871–5.
11. Sylaja PN, Coutts SB, Krol A, Hill MD, Demchuk AM; VISION Study Group. When to expect negative diffusion-weighted images in stroke and transient ischemic attack. Stroke 2008;39:1898–900. CrossRef
12. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. Br J Math Stat Psychol 2008;61(Pt 1):29–48. CrossRef
13. Carletta J. Squibs and Discussions Assessing Agreement on Classification Tasks: The Kappa Statistic. Computational Linguistics 1996;22:248–54.
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74. CrossRef
15. Sim J, Wright CC. The kappa statistic in reliability studies:

use, interpretation, and sample size requirements. Phys Ther 2005;85:257–68.

16. Jauch EC, Saver JL, Adams HP Jr, Bruno A, Connors JJ, Demaerschalk BM, et al. Guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2013;44:870–947. CrossRef
17. González RG, Schaefer PW, Buonanno FS, Schwamm LH, Budzik RF, Rordorf G, et al. Diffusion-weighted MR imaging: diagnostic accuracy in patients imaged within 6 hours of stroke symptom onset. Radiology 1999;210:155–62. CrossRef
18. Ferro JM, Pinto AN, Falcão I, Rodrigues G, Ferreira J, Falcão F, et al. Diagnosis of stroke by the nonneurologist. A validation study. Stroke 1998;29:1106–9. CrossRef
19. Kothari RU, Brott T, Broderick JP, Hamilton CA. Emergency physicians. Accuracy in the diagnosis of stroke. Stroke 1995;26:2238–41. CrossRef
20. Savitz SI, Caplan LR, Edlow JA. Pitfalls in the diagnosis of cerebellar infarction. Acad Emerg Med 2007;14:63–8. CrossRef
21. Oppenheim C, Stanescu R, Dormont D, Crozier S, Marro B, Samson Y, et al. False-negative diffusion-weighted MR findings in acute ischemic stroke. AJNR Am J Neuroradiol 2000;21:1434–40.
22. Ay H, Buonanno FS, Rordorf G, Schaefer PW, Schwamm LH, Wu O, et al. Normal diffusion-weighted MRI during stroke-like deficits. Neurology 1999;52:1784–92. CrossRef
23. Küker W, Weise J, Krapf H, Schmidt F, Friese S, Bähr M. MRI characteristics of acute and subacute brainstem and thalamic infarctions: value of T2- and diffusion-weighted sequences. J Neurol 2002;249:33–42. CrossRef