Original Paper

Accuracy of Reporting the Hyperdense Middle Cerebral Artery Sign as a Function of Clinical Experience

Patrick Aouad^{a, b} Andrew Hughes^{a, b} Terry Neeman^c
Christian J. Lueck^{a, b}

^aDepartment of Neurology, Canberra Hospital, ^bAustralian National University Medical School, and ^cStatistical Consulting Unit, Australian National University, Canberra, A.C.T., Australia

Key Words

Hyperdense middle cerebral artery · Ischaemic stroke · Reporting accuracy · Interobserver variability

Abstract

Background/Aim: The hyperdense middle cerebral artery sign (HMCAS) is a useful clinical sign in the management of acute stroke and may alter time-critical decisions within an emergency setting. Though gold standards have been published, these are rarely used in clinical practice and scans tend to be reported subjectively. It is therefore possible that the level of experience of the doctor reporting the scan may impact on the accuracy of the reporting and hence patient management. This study was designed to evaluate the accuracy in detecting HMCAS across doctors with varying levels of experience. **Methods:** Forty doctors were recruited into four categories of experience. Each subject received a brief computer-based tutorial on how to identify an HMCAS and was then asked to report on the presence or absence of an HMCAS in 19 pre-prepared CT scans using a standardised viewing template. **Results:** The mean ($\pm$ SE) percentage correct scores increased with experience from 76.8 ± 3.69 among interns and residents to 90.1 ± 2.23 (neurologists and radiologists; $p < 0.01$). Sensitivity and specificity as well as positive and negative predictive values all increased with experience. In addition, more experienced clinicians were better able to distinguish scans which met the radiological criteria for HMCAS from those which only just failed to do so. **Conclusions:** Experienced neurologists and radiologists consistently and accurately reported the presence or absence of HMCAS, whereas less experienced clinicians tended to over-report the presence of HMCAS. This may have implications for the acute management of thromboembolic stroke.

© 2015 S. Karger AG, Basel

Prof. Christian J. Lueck
Department of Neurology, Canberra Hospital
PO Box 11
Woden, ACT 2606 (Australia)
E-Mail christian.lueck@act.gov.au

Introduction

The hyperdense middle cerebral artery sign (HMCAS) was first described about 30 years ago as a high-density structure within the territory of the middle cerebral artery (MCA) on non-enhanced computed tomography (NECT) scans of patients with acute ischaemic stroke [1]. The HMCAS has been shown to be specific for acute ischaemic stroke affecting the MCA territory, and its presence is associated with a poorer outcome [2]. Its presence or absence has recently assumed increasing importance in the hyperacute management of patients presenting with ischaemic stroke, specifically in relation to whether or not consideration of intra-arterial thrombolysis or clot retrieval is appropriate [3, 4].

While radiological gold standards have been published [5, 6], the criteria are rarely applied in day-to-day clinical practice, with the decision generally being made simply on the visual appearance of the NECT. A study of the interobserver variability among neurologists found moderate to substantial agreement [7]. However, the individuals detecting and reporting on the presence or absence of the sign early on in patient management (when this may make a significant difference to acute management decisions) tend to be more junior staff rather than experienced senior clinicians or neuroradiologists (who generally report the scans hours to days later). This means that there is the potential for error in diagnosis and, therefore, management arising as a result of inexperience.

Since the HMCAS is a useful and important sign and may alter time-critical decisions within an emergency setting [3, 4], it is important to know whether less experienced clinicians can detect the presence or absence of the sign with a degree of accuracy comparable to that of experts. Accordingly, this study was set up to evaluate the accuracy of detecting the HMCAS across doctors with varying levels of experience practising within an Australian tertiary medical facility. The study aimed to assess whether non-experts could, after a brief tutorial, report the presence or absence of the sign as well as experts. This research further aimed to evaluate how consistently clinicians of varying degrees of experience reported the presence or absence of the sign and to correlate subjective reporting at all levels of experience with the radiological density of the MCAs as measured in Hounsfield units (HU).

Methods

Subject Selection

Forty subjects were recruited into four separate groups (10 in each) based on clinical experience. All subjects were qualified medical doctors working at the Canberra Hospital and all gave informed consent. The subjects were grouped as follows: experience group 1: interns and resident medical officers; experience group 2: basic physician trainee registrars and senior registrars in medical and surgical specialities, excluding neurology and radiology; experience group 3: emergency department staff specialists and senior registrars in neurology and radiology, and experience group 4: staff specialist neurologists and radiologists.

This study was approved by the ACT Health Human Research Ethics Committee (ETHLR.11.202) and the ANU Human Research Ethics Committee (211/567).

Scan Selection

Scans suitable for this study were de-identified and selected from a database of NECT brain scans using pre-specified criteria based on standard background viewing values for the scans, i.e. a level (centre) of 40 HU and a window of 90 HU. The criteria (table 1) were designed to select scans that would cover a broad range of radiological densities of MCAs. The scans had to have clearly discernible MCAs and could not be affected by radiological artefact (as

Table 1. Classification of NECT scans according to absolute and relative densities of MCAs

	Absolute density of MCA		
	high (46 HU ≤ MCA _{max} ≤ 55 HU)	moderate (41 HU ≤ MCA _{max} < 46 HU)	low (35 HU ≤ MCA _{max} < 41 HU)
Relative density			
High category (MCA _{rel} ≥ 1.2 HU)	A	B	C
Moderate category (1.1 HU ≤ MCA _{rel} < 1.2 HU)	B	C	D
Low category (MCA _{rel} < 1.1 HU)	C	D	E

determined independently by an expert radiologist). In total, 19 scans were selected. The density of the MCAs across these scans ranged from sub-normal, through normal, to hyperdense, and the scans covered a range of relative densities between left and right MCAs (table 1).

The densities were calculated from oval regions of interest (ROIs; 0.03–0.07 cm²) and were measured in HU. These values were divided into five scan groups (from group A to group E, where scan group A comprised the scans meeting the definition in the literature) according to how closely they met the definition of HMCAS, i.e. MCA_{attenuation} ≥ 46 HU and MCA_{attenuation ratio} > 1.2 using oval ROIs (online suppl. Appendix 1; for all online suppl. material, see www.karger.com/doi/10.1159/000370009) [5, 6].

Experimental Template Development

For each of the 19 NECT brain scans, the right and left MCAs were identified using the scroll function. Sufficient screenshots were taken of consecutive 1-mm slices to include both MCAs and at least 1 slice caudal and rostral to this region. The resolution of the images was preserved in each screenshot. The number of slices required in order to demonstrate both left and right MCAs adequately ranged in different scans from 4 to 10.

The 1-mm axial slices were then copied into a Microsoft PowerPoint document and layered in such a way as to mimic the exact format within the standard hospital picture archive and communicating system without any degradation in quality. Five introductory slides were added to provide the information required for the subjects to use the template, including an advice on how to detect an HMCAS (online suppl. Appendix 2). The hyperlinks were programmed into the PowerPoint presentation to allow the subjects to report the scans in an allocated order.

Data Collection

Each of the 40 subjects was given a reporting sheet along with an order in which to report the scans. The order of reporting the 19 scans was randomised to minimise bias due to learning or fatigue. The subjects were not allowed to return to a scan once it had been reported.

The subjects viewed the PowerPoint presentation on the standard display monitors available at the Canberra Hospital (Hewlett Packard HP Compaq LA2405wg). They were asked to indicate if they could see an HMCAS and, if so, on which side.

Data Analysis

The 3 scans in scan group A were selected because they met the definition of HMCAS in the literature [5, 6] and were therefore considered to be ‘true positives’. The other 16 scans were considered to be ‘true negatives’. All subjects’ answers were initially analysed to generate a mean percentage correct. One-way ANOVA was performed to assess the difference

Table 2. Individual results

Group 1		Group 2		Group 3		Group 4	
subject	% correct	subject	% correct	subject	% correct	subject	% correct
101	63.2	201	89.5	301	94.7	401	84.2
102	89.5	202	84.2	302	79.0	402	89.5
103	68.4	203	84.2	303	79.0	403	100.0
104	89.5	204	73.7	304	79.0	404	94.7
105	73.7	205	68.4	305	79.0	405	79.0
106	68.4	206	73.7	306	94.7	406	94.7
107	94.7	207	73.7	307	94.7	407	94.7
108	79.0	208	63.2	308	89.5	408	84.2
109	63.2	209	73.7	309	84.2	409	89.5
110	79.0	210	79.0	310	63.2	410	†

Group 1: mean 76.8 ± 3.7 , range 57.9–94.7. Group 2: mean 76.3 ± 2.5 , range 63.2–89.5. Group 3: mean 83.7 ± 3.2 , range 63.2–94.7. Group 4: mean 90.1 ± 2.4 , range 79.9–100.0.

† Omitted from all analyses.

in the results across the four experience groups. Post hoc analyses were then performed using the Dunnett test (two-sided) to make comparisons between the less experienced groups and the specialist group (group 4).

The sensitivity, specificity, positive predictive values (PPV), and negative predictive values (NPV) were then calculated for each of the four experience groups. Finally, the false-positive rate was calculated as the percentage of scans in each ‘radiologically negative’ category that were erroneously scored as positive. The mean false-positive rates were compared by experience group and scan group using a generalised fixed-effects model with physician and scan as random factors.

Results

Table 2 lists the percentage correct scores of each subject. A single subject from group 4 reported issues with utilising the PowerPoint template, and these data were discarded. The overall mean percentage correct scores increased and the range of results decreased with increasing level of experience.

ANOVA revealed a significant effect of the experience group on the mean percentage correct score [$F(3, 35) = 4.54$, $p < 0.01$]. Post hoc Dunnett t tests revealed that there was a significant difference between experience groups 1 and 4 ($p < 0.05$) and between groups 2 and 4 ($p < 0.05$), but not between groups 3 and 4 (fig. 1).

The sensitivity, specificity, PPV, and NPV are shown for each of the four experience groups in table 3. Group 4 performed best on all measures, but there was a gradual improvement across groups with increasing level of experience. Having said this, the PPV results were relatively low in all groups: even in the expert group (group 4), there was only a 65% chance that a scan reported as positive met the requirements of a true-positive scan as stated in the literature. This was due to the low prevalence of positive scans in the sample, consistent with what one would expect to see in clinical practice. However, the NPV results were very high in all four groups, meaning that if a doctor with any level of experience reported a scan as negative for an HMCAS, it was highly likely to be a true negative.

Fig. 1. Mean percentage correct score as a function of experience group. * $p < 0.05$

Fig. 2. False-positive rate as a function of scan and experience groups. Scan groups range from greatest similarity to positive scans (group B) to least similarity (group E). Experience groups range from least experienced (group 1) to most experienced (group 4).

Table 3. Sensitivity, specificity, PPV, and NPV

Group	Sensitivity, %	Specificity, %	PPV, %	NPV, %
1	63 (11–98)	79 (52–94)	36 (6–80)	92 (62–99)
2	67 (13–98)	78 (5–93)	36 (7–79)	93 (63–100)
3	70 (14–99)	86 (59–97)	49 (9–89)	94 (66–100)
4	81 (20–100)	92 (65–99)	65 (15–97)	96 (70–100)

As described in the Methods section, the 16 radiologically negative scans were classified according to how closely the parameters met the gold standard referred to in the literature [5, 6]. The false-positive rates for each of the four true-negative scan groups (B–E) are shown in figure 2 as a function of the experience group. Analysis using a fixed-effects model showed that there was a significant effect of the experience group on the results ($p < 0.05$), but there was no significant effect of the scan group, nor was there an interaction.

Discussion

The results demonstrated that, while there was a certain amount of interindividual variability, the mean number of correctly reported scans increased significantly with experience. Sensitivity, specificity, PPV, and NPV all increased with experience as well. In addition, more experienced clinicians were better able to distinguish those scans which met the radiological criteria for HMCAS [5, 6] from those which only just failed to do so. All these findings suggest that experience is important and that there is an increased chance of erroneous reporting by less experienced clinicians.

Though there was a clear increase with experience in both, specificity (79–92%) appeared to be higher than sensitivity (63–81%) for all experience groups. This suggests that there was a tendency to over-report the presence of an HMCAS, particularly in less experienced clinicians. This is consistent with the finding that the NPV was high across all groups (92–96%), whereas the PPV was much lower (35–65%). In practice, these findings suggest that if a doctor in any experience group felt an HMCAS was absent, this was highly likely to be correct. On the other hand, the declaration that a scan was positive for HMCAS was more likely to be correct with increasing levels of experience. Figure 2 demonstrates the effect of experience very well: doctors with more experience were significantly less likely to be ‘fooled’ by scans which almost met the radiological criteria for HMCAS [4] but did not actually do so.

We should make a comment about our choice of gold standard. Koo et al. [5] published a radiological gold standard of $MCA_{attenuation} \geq 43$ HU and $MCA_{attenuation\ ratio} > 1.2$ using oval ROIs in 2000, while Abul-Kasim et al. [6] published values of $MCA_{attenuation} \geq 46$ HU and $MCA_{attenuation\ ratio} > 1.2$ using oval ROIs in 2009. As the 95% confidence intervals for the HMCAS density in the paper by Koo et al. [5] ranged from 46.7 to 61.2 HU, we felt that their data were consistent with the slightly more stringent criteria published by Abul-Kasim et al. [6] and elected to use the latter.

The results suggest that a negative report from a doctor of any level of experience in the emergency setting can generally be taken to mean that an HMCAS was truly absent, so that the patient would be managed appropriately [3, 4]. However, a positive report would not necessarily mean there was an HMCAS present. This has significant implications for day-to-day practice as there is the potential for increased error in the patient management arising from the fact that the doctors acting on scans in the emergency setting often have less experience than those who issue the formal reports in hindsight.

The use of imaging software to determine whether scans meet radiological criteria has been shown elsewhere to improve accuracy and consistency [8]. Unfortunately, the use of imaging software takes time and requires familiarity with how to use the software, making it relatively impractical in a busy emergency department. Nevertheless, this study highlights the importance of ensuring that reporting is not left to relatively inexperienced junior doctors making a decision based simply on the visual appearance of the scan.

There are, in theory, three ways to overcome this problem, and these are not mutually exclusive. First, less experienced clinicians can be provided with better training, though it must be borne in mind that even the level of training provided in this study (online suppl. Appendix 2) was not adequate. Second, doctors reporting scans can be required to use imaging software to improve the accuracy. Third, emergency scans could be reported immediately by senior clinicians – this is becoming increasingly possible as the remote access to radiological images becomes easier.

This experiment has a number of limitations. First, the number of subjects was small, and the allocation of subjects to each group was based on the current job description rather than a detailed examination of previous experience. Nevertheless, the fact that the results demonstrated a steady trend towards greater accuracy with increasing level of experience suggests

that this empirical allocation was reasonable. This is also consistent with the observation made elsewhere that increasing clinical experience improves the general diagnostic accuracy when interpreting CT data in the context of acute ischaemic stroke [8].

Second, the number of scans assessed by each subject was only 19, meaning that the number of scans in each of the five categories was very small. However, it was felt that this was a reasonable total number because busy clinicians were unlikely to take part in a preliminary study if they were asked to give up any more time than the current study demanded.

Third, no clinical data were provided with each of the scans. It has previously been shown that when clinical information is provided, experts improve their level of agreement and accuracy [7]. If clinicians had been provided with clinical data to accompany the scans they were reporting, the level of accuracy and consistency amongst and within the groups would almost certainly have been better. However, this study was specifically designed to investigate the ability to report on the presence or absence of the HMCAS on NECT per se.

In conclusion, the reported absence of an HMCAS by a clinician of any level of experience is highly likely to imply a true negative. The accurate detection of the presence of an HMCAS, on the other hand, increases significantly with experience. This suggests that the reporting of NECT scans in the assessment of acute stroke is best undertaken by more experienced clinicians, although better education of junior staff in the future could change this.

Disclosure Statement

The authors have no conflicts of interest to disclose.

References

- 1 Gács G, Fox A, Barnett H, Vinuela F: CT visualization of intracranial arterial thromboembolism. *Stroke* 1983; 14:756–762.
- 2 Aouad P, Hughes A, Valecha N, et al: Prevalence, comorbid associations and prognostic value of the hyperdense middle cerebral artery sign: a prospective study. *ISRN Stroke* 2013; article ID 954825: 6 pages.
- 3 Agarwal P, Kumar S, Hariharan S, et al: Hyperdense middle cerebral artery sign: can it be used to select intra-arterial versus intravenous thrombolysis in acute ischemic stroke? *Cerebrovasc Dis* 2004;17:182–190.
- 4 Mattle HP, Arnold M, Georgiadis D, et al: Comparison of intraarterial and intravenous thrombolysis for ischemic stroke with hyperdense middle cerebral artery sign. *Stroke* 2008;39:379–383.
- 5 Koo CK, Teasdale E, Muir KW: What constitutes a true hyperdense middle cerebral artery sign? *Cerebrovasc Dis* 2000;10:419–423.
- 6 Abul-Kasim K, Selariu E, Brizzi M, Petersson J: Hyperdense middle cerebral artery sign in multidetector computed tomography: definition, occurrence, and reliability analysis. *Neurol India* 2009;57:143–150.
- 7 von Kummer R, Holle R, Gizyska U, et al: Interobserver agreement in assessing early CT signs of middle cerebral artery infarction. *AJNR Am J Neuroradiol* 1996;17:1743–1748.
- 8 Hand PJ, Kwan J, Lindley RI, Dennis MS, Wardlaw JM: Distinguishing between stroke and mimic at the bedside: the brain attack study. *Stroke* 2006;37:769–775.