

Clinical Consequences of Misinterpretations of Neuroradiologic CT Scans by On-Call Radiology Residents

Nirish R. Lal, Uwada M. Murray, O. Petter Eldevik, and Jeffrey S. Desmond

BACKGROUND AND PURPOSE: Studies have looked at the accuracy of radiologic interpretations by radiology residents as compared with staff radiologists with regard to emergency room plain films, emergency room body CT scans, and trauma head CT scans; however, to our knowledge, no study has evaluated on-call resident interpretations of all types of neuroradiologic CT scans. Both as a part of our departmental quality control program and to address concerns of clinical services about misinterpretation of neuroradiologic CT scans by on-call radiology residents, we evaluated the frequency of incorrect preliminary interpretations of neuroradiologic CT scans by on-call radiology residents and the effect of such misinterpretations on clinical management and patient outcome.

METHODS: As determined by the staff neuroradiologist the next day, all potentially clinically significant changes to preliminary reports of emergency neuroradiologic CT scans rendered by on-call radiology residents were recorded over a 9-month period. A panel of neuroradiologists reviewed and graded all the changed cases by consensus. An emergency department staff physician reviewed medical records of all submitted cases to determine clinical consequences of the misinterpretations.

RESULTS: Significant misinterpretations were made in 21 (0.9%) of 2388 cases during the study period. There was a significant change in patient management in 12 of the cases, with a potentially serious change in patient outcome in two cases (0.08%).

CONCLUSION: On-call radiology residents have a low rate of significant misinterpretations of neuroradiologic CT scans, and the potential to affect patient outcome is rare.

Neuroradiologic CT scans are essential in evaluating trauma and in examining acutely ill patients. Accurate interpretation of these studies is critical for proper patient management. Since many emergency studies are performed after regular working hours, on-call radiology residents play an integral role in the care of these patients.

Because clinical services in our medical center have questioned the accuracy of the on-call radiology residents' preliminary interpretations of neurologic CT scans, there has been an increased demand to have staff neuroradiologists in the hospital for longer portions of the day. Previous studies have evaluated the frequency of discrepancies in

interpretations made by residents and staff radiologists. Rhea et al found a 5% rate of significant or potentially significant errors in the interpretation of emergency room plain films (1). Wechsler et al found a discrepancy rate of 5.9% to 13.7% for significant errors between trainees (residents and fellows) and staff radiologists in interpreting emergency body CT scans (2). Roszler et al (3) and Wysoki et al (4) found error rates of 2.0% and 5.2%, respectively, between trainees and staff radiologists in interpreting trauma head CT scans. Other areas of neuroradiology have not been evaluated. As part of our departmental quality assurance program, we evaluated the frequency, consequences on clinical management, and effect on patient outcome of misinterpretations of neuroradiologic CT scans by on-call radiology residents in our hospital.

Methods

Data were collected prospectively during a 9-month period from June 23, 1997, to March 22, 1998. On-call radiology resident coverage at our hospital is from 5 PM to 8 AM on weekdays and all day on weekends and holidays. A prelimi-

Received March 9, 1999; accepted after revision August 11.

From the Departments of Radiology (N.R.L., U.M.M., O.P.E.) and Emergency Medicine (J.S.D.), University of Michigan Medical Center, Ann Arbor.

Presented in part at the annual meeting of the American Society of Neuroradiology, Philadelphia, May 1998.

Address reprint requests to Nirish R. Lal, MD, 2771 Page Ave, Ann Arbor, MI 48104.

nary report is rendered on all emergency cases. A staff neuroradiologist reviews all the neuroradiologic examinations the next morning, including weekends and holidays. At this time, they notify the referring physicians of any significant changes to the preliminary interpretations. During our study period, for any potentially clinically significant disagreement between the staff neuroradiologist's interpretation and the resident's preliminary interpretation, a study form was filled out by the staff neuroradiologist. All neuroradiologic CT examinations, including head, spine, neck, maxillofacial, temporal bone, and orbital CT scans, for which the on-call radiology resident rendered a preliminary interpretation were included in our analysis. Studies were performed predominantly on adult patients in the emergency department, but also included pediatric patients and emergency studies on inpatients.

Head CT scans were performed with 5-mm contiguous axial images every 5 mm (5q5) through the posterior fossa and 10q10 through the remainder of the brain. Spine CT scans were performed primarily 3q3. Neck CT scans were 5q5. Maxillofacial and orbital CT scans were 3q3 axial images, with coronal images obtained if possible. Temporal bone CT scans were 1q1 axial and coronal. Examinations were performed without intravenous contrast material, except for most neck CT studies or when there was a strong indication for contrast. Bone windows were obtained when indicated and for all trauma cases.

At our institution, radiology call for all CT, sonographic, and MR examinations is split between junior and senior residents; the former have 20 to 31 months of training in general radiology, while the latter have 32 to 43 months of such training before starting call. Not including call, our residents work an average of 50 to 60 hours a week. At least 7 to 8 weeks of this training is within dedicated neuroradiology rotations prior to starting call. Additionally, there is a weekly 45-minute neuroradiology teaching conference throughout the residency.

Staff neuroradiologists were educated about the project and regularly reminded to record potentially clinically significant discrepancies with preliminary interpretations made by on-call residents on a form, which was readily available at the alternator. For all submitted forms, cases were reviewed by a panel of neuroradiologists, who graded the misinterpretations by consensus according to the following scale: grade 1, case thrown out because the panel determined that either the resident interpretation was correct, the resident interpretation was adequate, or the error made was of no clinical significance; grade 2, the finding or diagnosis could not be expected to be identified; grade 3, the finding or diagnosis would not ordinarily be expected to be identified; grade 4, the finding or diagnosis should have been identified. All submitted cases, including those that were thrown out, were independently evaluated by an emergency department staff physician to determine the effect of the misinterpretation on clinical management and patient outcome. To accomplish this, he reviewed all available clinical records and follow-up information. Additionally, physicians involved with the case were contacted when the other information was inconclusive for determining the effect of the misinterpretation. On the basis of the information he gathered and his clinical experience, the evaluating physician graded the effect on clinical management as no change, minor change, or significant change, and the effect on patient outcome was graded as no change, possible minor change, or potentially significant change.

Results

Preliminary reports were rendered in 2388 cases during the study period. Of these, 1734 (73%) were adult head CT scans, 302 (13%) were pediatric head CT scans, 162 (7%) were adult spine CT

scans, 29 (1%) were pediatric spine CT scans, and 161 (7%) were other CT scans. Thirty-five forms (1.5% of cases) indicating significant resident misinterpretations were completed. These were graded by the neuroradiology panel as follows: thrown out, 14; grade 1, one; grade 2, seven; grade 3, nine; and grade 4, four. Twenty-one cases (0.9%) were not thrown out, and consisted of 17 adult head CTs, one adult spine CT, one pediatric spine CT, one temporal bone CT, and one orbital CT.

Of the misinterpretations not thrown out, junior residents made nine of the errors and senior residents made 12. Fourteen of the 18 residents in the call pool made at least one misinterpretation, with three errors being the most by any one resident. From the actual time the scan was performed, the average delay in contacting the clinical service about the change in interpretation was approximately 10 hours 40 minutes (range, 4 hours 15 minutes to 15 hours 50 minutes).

A review of clinical records revealed that three patients had a minor change and 12 patients had a significant change in their clinical management because of the preliminary misinterpretation. The significant changes in patient management were as follows: needed additional imaging (n = 2) (Fig 1); consulted subspecialty service (n = 1); additional imaging and subspecialty consult (n = 4) (Fig 2); would have been admitted (n = 2) (Figs 3 and 4); delay in treatment (n = 1) (Fig 5); and change in treatment (n = 2).

Two patients had a potentially serious change in their outcome, while the other 33 patients had no change in outcome. The first patient with a potentially serious change in outcome presented with right hemiparesis and aphasia. A head CT (Fig 6) showed unilateral basal ganglia high attenuation, initially interpreted as most likely representing hemorrhage. Given the possibility of hemorrhage, the patient was not considered a candidate for thrombolytic therapy. The CT was later interpreted to represent a somewhat atypical appearance of basal ganglia calcification. This misinterpretation resulted in a change in clinical management, as the patient was not offered intravenous thrombolytic therapy for his presumed brain infarct. Although it is uncertain if thrombolysis would have beneficially or adversely affected this particular patient, there was a potentially serious change in the patient's outcome. The second patient had a history of renal failure and presented with a change in mental status. The preliminary interpretation of the head CT (Fig 7) was negative. The head CT findings were later interpreted as being consistent with a cerebellar stroke. This misinterpretation resulted in a change in the patient's clinical management, as the patient underwent emergency dialysis after the initial interpretation to exclude uremia as the cause of symptoms, given what was believed to be a negative head CT. Additionally, the management of the patient's stroke was delayed. Although the patient would have been unlikely to be a candidate for

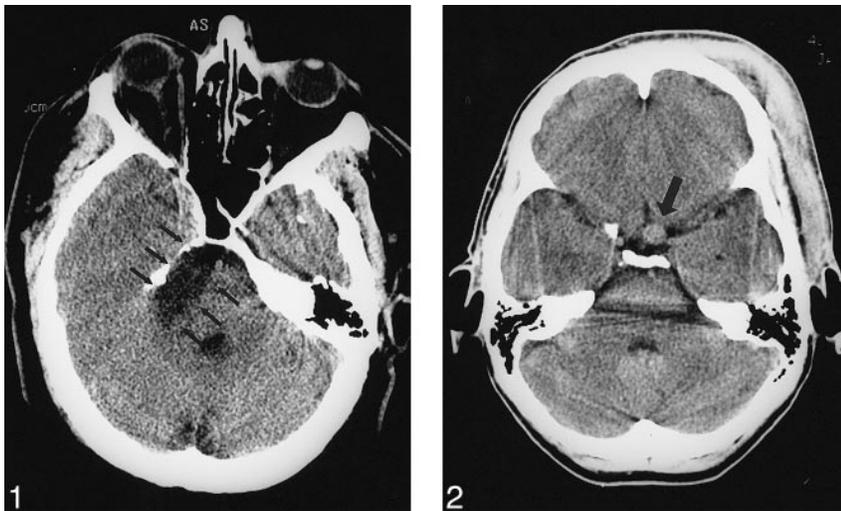
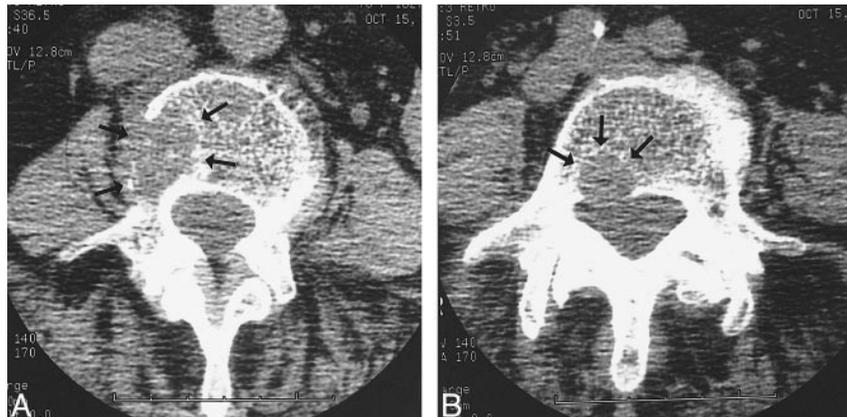


FIG 1. 51-year-old man with headache, neck pain, and loss of consciousness after a fall. An emergency noncontrast head CT was performed to determine the presence of hemorrhage or fracture. The initial interpretation by the on-call resident was no significant abnormality. The next morning, the staff neuroradiologist identified a 2 × 3-cm right cerebellopontine angle lesion (arrows) causing mass effect on the pons and fourth ventricle, most likely representing an epidermoid. This error in interpretation was given a grade 4 (should have been identified) by the panel. The patient was scheduled for a follow-up MR examination; however, the 10½-hour delay in diagnosis did not change his outcome.

FIG 2. 43-year-old woman with loss of consciousness after a motor vehicle accident. An emergency noncontrast head CT was performed to evaluate for hemorrhage. The initial interpretation by the on-call resident identified a scalp hematoma. The next morning, the staff neuroradiologist also identified an 8-mm soft-tissue mass left of the midline in the suprasellar cistern (arrow). This error in interpretation was given a grade 3 (could usually be identified) by the panel. The patient was scheduled for a follow-up MR angiogram, which showed the mass to be an aneurysm of the left supraclinoid carotid artery, and she later underwent elective surgery. The 15-hour delay in diagnosis did not change this patient's outcome.

FIG 3. 78-year-old woman with back pain 5 years after a mastectomy for breast cancer.

A and B, An emergency lumbosacral spine CT was performed to evaluate for L3 compression fracture. The initial interpretation by the on-call resident identified multi-level degenerative changes with a questionable lesion at L5 versus degenerative change. The next morning, the staff neuroradiologist noted destructive lesions on the right side of L4 (A, arrows) and the posterior aspect of L5 (B, arrows), consistent with metastases. This error in interpretation was given a grade 4 (should have been identified) by the panel. The patient had been sent home instead of being admitted for metastatic workup, so she was called back the next day; however, the 6-hour delay in diagnosis did not change her outcome.



thrombolysis, given the changes already seen on the CT scan, it is uncertain if any treatment, such as heparin or hyperventilation or mannitol, would have benefited this patient if it had been performed earlier, but nonetheless there was a potentially serious change in the patient's outcome.

Discussion

Our study differs in several ways from previous studies performed to evaluate radiology residents' interpretations of neuroradiologic CT scans (3, 4). First, we looked at residents' interpretations of all emergency neuroradiologic CT scans, while prior

studies evaluated only trauma head CT scans (3, 4). Second, a greater number of examinations were reviewed by our residents than in previous studies. Third, a panel of neuroradiologists reviewed all cases to ensure that the individual staff neuroradiologist was "correct" when determining a significant misinterpretation by a resident. Since the definition of a potentially clinically significant error is subjective, the standard of reference was based on group consensus. Additionally, Hillman et al (5) showed that diagnostic accuracy of over 75% of the radiologists in their study improved after group consultation. Finally, a clinical staff physician outside the department of radiology reviewed all sub-

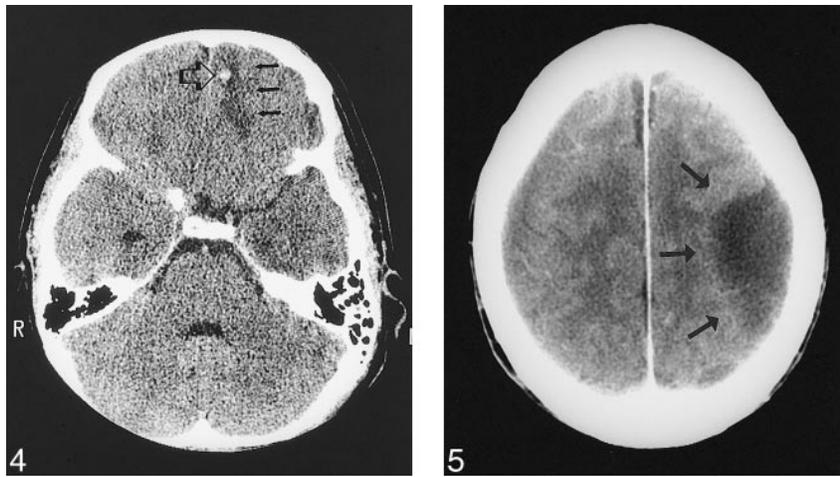


FIG 4. 15-year-old girl with left hemotympanum. An emergency temporal bone CT was performed to evaluate for temporal bone fracture. The initial interpretation by the on-call resident identified a left temporal bone fracture. The next morning, the staff neuroradiologist also noted a hemorrhagic contusion of the left frontal lobe (*open arrow*), with surrounding edema (*closed arrows*). This error in interpretation was given a grade 2 (would not ordinarily be expected to be identified) by the panel. The patient was called back for reevaluation and a follow-up head CT. The 11-hour delay in diagnosis did not change her outcome.

FIG 5. 70-year-old man with metastatic prostate cancer, presented with right-sided weakness and mental status changes. An emergency head CT was performed to evaluate for metastasis or infarct. The initial interpretation by the on-call resident was intra-axial edema (*arrows*) from a metastasis or infarct. The next morning, the staff neuroradiologist interpreted this lesion to represent an extra-axial fluid collection. This error in interpretation was given a grade 2 (would not ordinarily be expected to be identified, difficult to distinguish) by the panel. The initial error in interpretation resulted in a 13-hour delay in treatment, as the patient was taken to the operating room for drainage of the subdural hematoma the next day; however, his outcome was not changed.

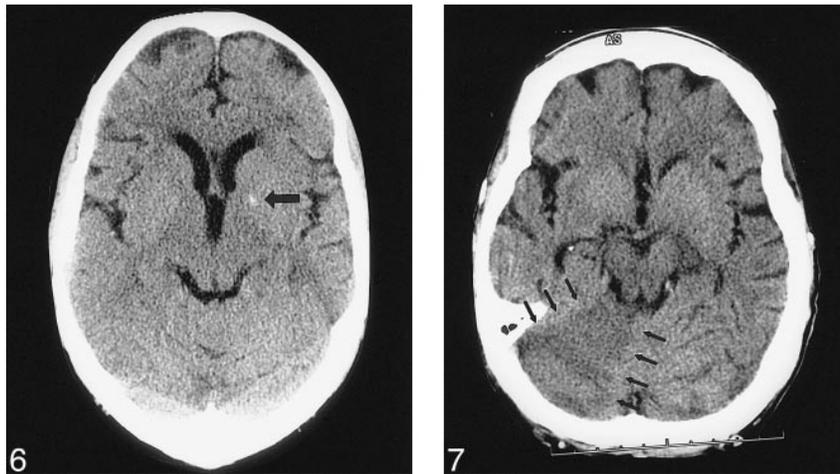


FIG 6. 66-year-old man with right hemiparesis and aphasia. An emergency noncontrast head CT was performed to determine whether the patient was a potential candidate for thrombolysis. The initial interpretation by the on-call resident identified a high-attenuation focus in the left basal ganglia (*arrow*), most likely representing hemorrhage. The next morning, the staff neuroradiologist interpreted this lesion as calcification, not hemorrhage. This error in interpretation was given a grade 3 (could usually be identified) by the panel. The initial misinterpretation resulted in thrombolytic therapy being withheld, denying the patient the potential benefit of that treatment and thereby representing a potentially serious change in his outcome.

FIG 7. 61-year-old woman with end-stage renal disease presented with a change in mental status. An emergency noncontrast head CT was performed to evaluate for septic emboli or hemorrhage. The initial interpretation by the on-call resident was negative. The next morning, the staff neuroradiologist noted a right cerebellar infarct (*arrows*). This error in interpretation was given a grade 3 (could usually be identified) by the panel. The initial misinterpretation resulted in a change in management, as the patient received dialysis to exclude uremia as the possible cause of her symptoms, and a delay in the treatment of her stroke, resulting in a potentially serious change in the patient's outcome.

mitted cases to assess potential effects of the preliminary misinterpretation on clinical management and patient outcome.

The rate of significant misinterpretations by on-call radiology residents was low (21/2388, 0.9%).

This is slightly lower than that reported by Roszler et al (3) (2%) and Wysoki et al (4) (5.2%, with 1.7% for major discrepancies). In contrast to Roszler's study, in which the most junior residents made a majority of the errors (3), we found no trend for

an increased number of errors among the junior residents; in fact, more of the errors were made by senior residents. These two discrepancies with Roszler's results may reflect differences in resident training, as our residents do not read neuroradiologic CT scans on call until they have had at least 20 months of radiologic training, including at least 7 to 8 weeks of dedicated neuroradiologic training. Additionally, the lack of difference between junior and senior residents may also relate to the hypothesis by Herman and Hessel (6), based on their study of chest X-ray interpretation by physicians at four different levels of training, that once radiologic training has reached a certain fundamental level, individual reader's characteristics overshadow experience in determining accuracy.

Of the 21 misinterpretations judged to be potentially clinically significant, 13 were graded as having findings that could usually be or that should have been identified. The other eight cases had findings that could not be expected to be or that would not ordinarily be expected to be identified. One may then argue that it was reasonable for the residents to have made an error in these eight cases.

Although the neuroradiology panel agreed that the preliminary interpretations in 21 cases contained potentially clinically significant errors, clinical management and patient outcome were often not affected. The delay in notifying the clinical service of the change in interpretation led to a significant change in clinical management in 12 cases. Even with an alteration in patient management in these 12 cases, patient outcome was potentially affected in only two (0.08%) of the 2388 examinations. This is similar to results obtained by Wysoki et al (4), who found no cases that affected patient outcome in their study, even in the presence of delays of up to 15 hours between resident and staff interpretations.

There are several potential limitations of this study. First, noncompliance with completion of the forms to document a misinterpretation would have resulted in an underestimation of the true rate of misinterpretations. To minimize this occurrence, the forms were easy to fill out and readily available at the alternator. Additionally, the neuroradiology staff were educated about the study and continually reminded to fill out the forms. Second, the requirement that forms be completed only for potentially significant discrepancies allowed subjective judgments by the neuroradiology staff as to whether to submit a form. Since there appeared to be a tendency to submit forms for minor changes in interpretation, it is unlikely that any significant changes were missed. In addition, the preliminary report log book was checked by the authors to try to pick up any altered reports. Third, the intent of the grading scale was to rate the residents on the basis of how a neuroradiologist, not a resident, would interpret the cases. The actual standard used by the neuroradiology panel undoubtedly varied slightly among cases and, on average, was probably somewhere

between these two reference levels. Fourth, only a single physician reviewed the clinical information to determine effects on patient management and outcome, which increased the chance for bias. For this reason, a physician outside the department of radiology performed this function to decrease any potential bias to try to show no effect of the misinterpretations. Additionally, because detailed follow-up was not available for a few patients, it was helpful to have a physician who was more familiar with the clinical management of these patients consider the effect of the misinterpretations.

The panel threw out 14 of 35 cases because they either agreed with the resident, judged the resident's interpretation to be adequate, or determined that the error was insignificant. One may wonder why so many cases were thrown out. A possible explanation is that the neuroradiology staff were trying to point out minor and subtle findings to educate residents the next morning. These "educational" changes by the staff were often then thrown out by the panel, with the initial neuroradiologists often changing their mind and agreeing that the case should be thrown out. Although it is laudable for staff to identify and point out subtle potential findings to educate residents the next morning, one must be careful. Communicating minor changes of the initial interpretation to referring physicians can lead to frustration on the part of doctors and patients. Physicians may then need to change what they have told a patient, who may then have to return for additional tests or clinic visits. This is particularly a problem with the emergency department, because patients have usually left the department by the time the changed interpretation is communicated. Interestingly, more cases were thrown out during the first half of our study (eight of the first 16 cases, or 50%) than in the second half (six of the next 19 cases, or 32%), probably because the peer review of submitted cases during the first half of the study period modified the behavior of the neuroradiology staff, resulting in a decreased tendency for staff to "overcall" subtle questionable findings.

Although the staff radiologists are used as the reference standard, they are not always correct. A study by Rhea et al (1) of residents' interpretations of emergency plain films showed that in cases for which they could get clinical follow up, the residents ended up being correct in nearly 20% of the cases changed by the staff. We had a panel of neuroradiologists review all the cases to help ensure that changes to the preliminary interpretations were correct and to standardize grading of the seriousness of misinterpretations. Several studies have shown that radiologists have an inherent error rate in interpreting chest X-rays, double-contrast barium enemas, and skeletal and gastrointestinal studies, which will decrease when films are reviewed by additional radiologists (5, 7, 8). Although neuroradiologic CT scans have not specifically been evaluated, one would hypothesize that there must

be an inherent error rate in interpreting these studies as well. Therefore, when neuroradiology staff review cases the next day, one would expect them to make observations that were missed on the preliminary interpretation. To find the true misinterpretation rate by residents, one would then have to determine the error rate by staff neuroradiologists and subtract the difference. A possible future study to address this issue would be to evaluate whether there are any findings made by the on-call resident that are missed by the staff radiologist the next day if he does not have access to the preliminary report.

Although our resident error rate was low, what is an acceptable rate is something that is difficult to determine. Our finding of only a 0.08% rate of potentially serious changes in patient outcome has not caused us to change our current practices for neuroradiology staff coverage. Although the data from this study have not formally been presented to the clinical services in our hospital, they have been informally referred to at a weekly clinical trauma conference, and an emergency department physician was involved in the study. We therefore presume that the primary services responsible for ordering neuroradiologic CT scans at night are aware of the findings of this study. Both the department chairman and the neuroradiologist who attends the weekly trauma conference have subjectively noted a decrease in the number of complaints about resident on-call interpretations since the completion of the study. We have used the information gathered from this study to educate our residents. The missed cases have been shown to residents in both a grand rounds presentation as well as in educational conferences. In addition, we have expanded our quality assessment evaluation of on-call resident interpretations to include body CT scans,

sonograms, and plain films. The missed cases are used to educate our residents at a monthly noon conference.

Conclusion

On-call radiology residents have a low rate of significant misinterpretations of neuroradiologic CT scans. It is rare (0.08% of cases) that a potentially serious effect on patient outcome results from having radiology residents render a preliminary interpretation on call. On the basis of the results of our study, we have not altered our current practice in the evaluation of neuroradiologic CT scans other than to educate our residents about their misinterpretations.

References

1. Rhea JT, Potsaid MS, De Luca SA. **Errors of interpretation as elicited by a quality audit of an emergency radiology facility.** *Radiology* 1979;132:277-280
2. Wechsler RJ, Spettell CM, Kurtz AB, et al. **Effects of training and experience in interpretation of emergency body CT scans.** *Radiology* 1996;199:717-720
3. Roszler MH, McCarroll KA, Rashid T, Donovan KR, Kling GA. **Resident interpretation of emergency computed tomographic scans.** *Invest Radiol* 1991;26:374-376
4. Wysoki MG, Nassar CJ, Koenigsberg RA, Faro SH, Faerber EN. **Head trauma: CT interpretation by radiology residents versus staff radiologists.** *Radiology* 1998;208:125-128
5. Hillman BJ, Swensson RG, Hessel SJ, Gerson DE, Herman PG. **The value of consultation among radiologists.** *AJR Am J Roentgenol* 1976;127:807-809
6. Herman PG, Hessel SJ. **Accuracy and its relationship to experience in the interpretation of chest radiographs.** *Invest Radiol* 1975;10:62-67
7. Hessel SJ, Herman PG, Swensson RG. **Improving performance by multiple interpretations of chest radiographs: effectiveness and cost.** *Radiology* 1978;127:589-594
8. Markus JB, Somers S, O' Malley BP, Stevenson GW. **Double-contrast barium enema studies: effect of multiple reading on perception error.** *Radiology* 1990;175:155-156