AJNR

Trainee Misinterpretations on Pediatric Neuroimaging Studies: Classification, Imaging Analysis, and Outcome Assessment

C.V.A. Guimaraes, J.L. Leach and B.V. Jones

ORIGINAL RESEARCH

C.V.A. Guimaraes
J.L. Leach
B.V. Jones

Trainee Misinterpretations on Pediatric Neuroimaging Studies: Classification, Imaging Analysis, and Outcome Assessment

BACKGROUND AND PURPOSE: The scope of trainee misinterpretations on pediatric neuroimaging studies has been incompletely assessed. Our aim was to evaluate the frequency of trainee misinterpretations on neuroimaging exams in children, describe a useful classification system, and assess related patient management or outcome changes.

MATERIALS AND METHODS: Pediatric neuroimaging examinations with trainee-dictated reports performed without initial attending radiologist assessment were evaluated for discrepant trainee interpretations by using a search of the RIS. The frequency of discrepant trainee interpretations was calculated and classified on the basis of the type of examination on which the error occurred, the specific type and severity of the discrepancy, and the effect on patient management and outcome. Differences relating to examination type and level of training were also assessed.

RESULTS: There were 143 discrepancies on 3496 trainee-read examinations for a discrepancy rate of 4.1%. Most occurred on CT examinations (131; 92%). Most discrepancies (75) were minor but were related to the clinical presentation. Six were major and potentially life-threatening. Thirty-seven were overcalls. Most had no effect on clinical management (97, 68%) or resulted simply in clinical reassessment or imaging follow-up (43, 30%). There was no permanent morbidity or mortality related to the misinterpretations. The most common misinterpretations were related to fractures (28) and ICH (23). CT examinations of the face, orbits, and neck had the highest discrepancy rate (9.4%). Third- and fourth-year residents had a larger discrepancy rate than fellows.

CONCLUSIONS: Trainee misinterpretations occur in 4.1% of pediatric neuroimaging examinations with only a small number being life-threatening (0.17%). Detailed analysis of the types of misinterpretations can be used to inform proactive trainee education.

ABBREVIATIONS: CAQ = Certificate of Added Qualification; CI = confidence interval; ICH = intracranial hemorrhage; MRA = MR angiography; MRI = MR imaging; RIS = radiology information system

On-call trainee (resident and fellow) initial interpretations of neuroimaging studies during off-hours are common in large pediatric medical centers. Although rendering preliminary reports is currently an integral part of imaging training, discrepancies can occur and potentially impact patient care. There have been a number of studies that have evaluated the discrepancy rates of radiology trainee interpretations of adult neuroimaging studies[1-7]; however, to our knowledge, no studies have evaluated this process related specifically to the unique neuroimaging issues encountered in children. Age-related changes in imaging appearance, different disease spectra, and limited exposure to pediatric neuroimaging by radiology trainees can produce additional and unique challenges to accurate interpretation. Our goals for this study were to assess the frequency of trainee misinterpretations on pediatric neuroimaging studies at a large pediatric hospital, describe a useful classification system for misinterpretations and their clinical importance, and assess related patient outcomes.

Materials and Methods

This study was approved by the institutional review board at the Cincinnati Children's Hospital Medical Center.

Trainee Misinterpretation Identification

At our institution, imaging examinations with trainee-dictated reports that were performed without the initial attending radiologist's assessment (off-hours examinations) are tracked within the RIS. Both a resident (years 1–4 during the study period) and a fellow performed in-house preliminary interpretations of off-hours examinations, with most being interpreted by the fellows. These examinations were subsequently interpreted by 1 of 7 staff radiologists working in the neuroradiology section, 4 with CAQs in neuroradiology, 2 with CAQs in pediatric radiology, and 1 with a CAQ in both pediatric radiology and neuroradiology.

For the purpose of this study, neuroimaging examinations included CT and MR imaging examinations of the head, face, neck, and spine as well as CT scans and MRAs of the head and neck. Our institution has a policy and procedure for the communication and documentation of changes to preliminary reports that are recognized as clinically significant by faculty. Clinically significant discrepancies are defined as any alteration in interpretation that results in clinically important changes in the primary diagnosis, differential diagnosis, or

Table 1: Categorization of discrepant trainee interpretation

Type	Number of Discrepant Trainee Reports (Rate)[a]
1: Major, life-threatening	6 (0.17%)
1A: Finding not originally identified	5
1B: Finding identified, incorrectly characterized	1
2: Minor, related to clinical presentation	75 (2.1%)
2A: Finding not originally identified	56
2B: Finding identified, incorrectly characterized	19
3: Minor, unrelated to clinical presentation	8 (0.23%)
3A: Finding not originally identified	8
3B: Finding identified, incorrectly characterized	0
4: Possible abnormality	17 (0.49%)
4A: Confirmed on follow-up imaging	2
4B: Not confirmed on follow-up imaging	8
4C: No follow-up imaging performed	7
5: Abnormality called when none present (overall)	37 (1.1%)
5A: Resulting in inappropriate therapy	2
5B: Not resulting in inappropriate therapy	35

[a] % of total exams interpreted ($N = 3496$).

Table 2: Effect of discrepant trainee interpretations on clinical management and outcome

Type	Number of Discrepant Reports (% of Discrepancies)
1: No effect on clinical management/outcome	97 (68)
2: No direct treatment change but imaging or clinical follow-up performed related to the discrepancy	43 (30)
3: Direct treatment change, no sequelae	3 (2.1)
4: Direct treatment change (morbidity)	0 (0)
5: Death potentially related	0 (0)

recommendations for follow-up imaging or clinical assessment. The policy dictates that the preliminary report is not changed and a standardized macro stating "Final Impression After Attending Radiologist Review: Note: There has Been a Change From the Preliminary Report:" is used to clearly identify that clinically significant discrepancies have been documented by the faculty signing off on that study. When a clinically significant discrepancy is found, the staff radiologist communicates the findings to the clinical service caring for the patient, per department policy.

Using the text search capabilities of our institutional RIS, the standardized macro allowed identification of all preliminary studies in which a clinically significant discrepancy (as defined above) was identified. We retrospectively searched the RIS for all neuroimaging studies interpreted off-hours by a trainee during a consecutive 18-month period (Jan 1, 2008, through June 30, 2009). This included emergency department, outpatient, and in-house patients. We then identified all the examinations in which the standardized macro for clinically significant changes in the preliminary report was used.

Discrepant Interpretation Classification

All imaging examinations with discrepant trainee interpretations were reviewed by 2 CAQ-certified neuroradiologists with fellowship training in both adult and pediatric neuroradiology with 14 and 15 years of postsubspecialty training experience (expert reviewers). Decisions as to the validity of the identified discrepancy were made by consensus. A full patient chart review was performed for each discrepant examination. Each discrepancy was categorized by the type of error and the severity (type 1 through type 5), as demonstrated in Table 1, by using the best clinical judgment of the expert reviewers and the clinical context of each examination. In our method, discrepancy classification as major (type 1) or minor was based on the potential immediate clinical impact of the misinterpretation.[8]

Type 1 (major, life-threatening) discrepancies were defined as those in which the findings were of major clinical importance and the knowledge of which could result in immediate therapeutic plan alterations. These could include (but were not restricted to) the following: subdural or epidural hemorrhage >2 mm in maximum thickness, intraparenchymal hemorrhage >5 mm in greatest dimension, any ICH demonstrating mass effect, diffuse edema with signs of increased

intracranial pressure, acute hydrocephalus with signs of increased intracranial pressure, missed mass lesion (with significant mass effect or secondary hydrocephalus), nontraumatic subarachnoid hemorrhage, and missed signs of large arterial distribution infarct.

Minor discrepancies (types 2 and 3) were those that were not deemed immediately life-threatening. Although classified as minor for this study, these cases could include serious conditions such as smaller intracranial masses, small foci of traumatic ICH, or a missed cortical malformation. Type 4 discrepancies were possible abnormalities of uncertain significance. Examples would include areas of altered attenuation on CT examinations that could be artifacts but were suspicious enough to warrant comment or further investigation, signal-intensity abnormalities on MR imaging that could be related to technical factors, or questionable fractures versus normal variants of bone (atypical sutures, vascular channels). In this study, we also describe those examinations in which an abnormality was called by the trainee and was thought to be normal by the attending radiologist (overcalls). These were further subclassified by whether they resulted in inappropriate therapy. Although none were identified in this study, an overcall that resulted in an inappropriate surgical or potentially dangerous medical intervention would have been classified as type 1—major, life threatening.

Outcome Assessment

The effects of the discrepant trainee interpretation on patient management and outcome are classified as type 1 through type 5 as demonstrated in Table 2. A direct treatment change is defined as a surgical or medical intervention that was performed on the basis of the new information provided by the final imaging report. An example of a type 1 clinical outcome would be a small subdural hemorrhage in a trauma patient with another correctly identified ICH, resulting in no management changes. Type 2 clinical outcomes would include cases in which a follow-up imaging study or clinical assessment was performed that otherwise might not have been. An example would be the recall of the patient to the emergency department for follow-up imaging and clinical assessment, with no change in initial treatment plan and no patient morbidity. Type 3 outcomes reflect a change in treatment plan without associated morbidity as a result of the delay, and type 4 outcomes have a change in treatment associated with additional morbidity. A type 5 outcome would occur if a patient death could have been related to the discrepancy in interpretation.

A detailed subclassification was used for assessing the types of discrepant trainee interpretations as demonstrated in Table 3. Statistical assessment between groups was performed by using a 2-sample 2-sided Z-test for proportions, continuity-corrected to

Table 3: Discrepancy subclassification

Discrepancy Subclass	No. (N = 143)
Fracture	28
ICH	23
Focal brain parenchymal attenuation/signal abnormality	16
Other	15
Ventricle size	11
Diffuse edema/attenuation/signal	9
Mass lesion	7
Vascular	7
Abscess/fluid/edema	6
Extra-axial collection	5
Osseous, nonfracture	5
Cerebellar tonsil position	3
Intracranial/soft-tissue air	3
Vertebral alignment	3
Intraspinal hemorrhage	2

Table 4: Discrepancy rate by exam type

Exam Type	Discrepant Exams (% of Discrepant Exams)	Total Exams Read (% Discrepant)[a]
CT Total	131 (91.6)	3102 (4.22 [3.49–4.94])
CT head	103 (72.0)	2748 (3.75 [3.02–4.48])
CT face/orbit/neck	20 (14.0)	213 (9.39 [5.24–13.54])
CT spine	8 (5.6)	141 (5.67 [1.5–9.84])
MRI total	12 (8.4)	394 (3.04 [1.22–4.86])
MRI brain	7 (4.9)	211 (3.32 [0.66–5.98])
MRI spine	4 (2.8)	113 (3.54 [0.0–7.39])
MRA	1 (0.7)	70 (1.43 [0.0–4.92])
Total exams	143	3496

[a] Discrepancy rate (% [95% CI]). The discrepancy rates for CT scans of the face, orbit, and neck were significantly larger compared with CT head, combined CT head and spine, and MRI. The discrepancy rate was not significantly different comparing MRI and CT, or MRI subcategories.

produce 95% CIs. Aberrant CI values (<0.0%) are listed as 0.0 in the Tables.

Results

Overall Discrepancy Rate

There were a total of 3496 trainee-interpreted examinations during the study period. Of these examinations, there were 158 trainee interpretations that were identified as potentially harboring clinically significant discrepancies by the attending radiologist for an overall discrepancy rate of 4.52%. Of the initial 158 discrepant interpretations, 15 did not harbor a true discrepancy when evaluated by the expert reviewers (9.5% of all discrepant trainee interpretations). The inappropriate identification of interpretations as discrepant was a result of 3 error types on the part of the faculty: 1) re-emphasis or restating the findings dictated by the trainee with no significant change in the findings or differential diagnosis (6 cases), 2) correction of grammatical or voice-recognition errors (5 cases), or 3) discrepancies in which the 2 expert reviewers concurred that the trainee interpretation was a valid one, based on review of the imaging study and the chart (4 cases). Therefore, the final assessed discrepancy rate was 4.1% (143 discrepant interpretations of 3496 examinations). The age range of the patients examined was 1 day to 21 years (mean, 7 years 6 months).

Discrepancy Rate by Examination Type

The types of examinations in which the discrepant interpretations were made are listed in Table 4. Most occurred in the interpretation of CT studies (131, 92%), most of which were of the head (103, 72%). MR imaging examinations were responsible for a much smaller percentage of the discrepancies (12, 8%). There were 131 discrepancies in 3102 CT examinations and 12 discrepancies in 394 MR imaging examinations for a technique-related discrepancy rate of 4.2% for CT and 3.0% for MR imaging. There was no statistical difference between the discrepancy rates of MR imaging and CT examinations in total. CT examinations of the face, orbit, and neck had a discrepancy rate of 9.4%, which was greater than that for head CT (3.75%), combined head and spine CT (3.84%), and MR imaging.

Clinical Classification of Discrepancies

The type and severity of discrepancies are summarized in Table 1. Most discrepant trainee interpretations were related to minor findings that were potentially related to the patient's clinical presentation. Major life-threatening discrepancies (type 1) were rare, occurring 6 times for a rate of 0.17%. Four of the 6 were cases of misinterpreted brain edema or hypoxic-ischemic change (Fig 1), and 2 were cases of potentially significant traumatic ICH (Fig 2). Of the 75 discrepant trainee interpretations classified as type 2 (minor, related to clinical presentation), the most common findings included minimal traumatic ICH (14), calvarial fractures (10), abnormal ventricle size (9), focal parenchymal changes (6), and mass lesions (3).

There were 37 type 5 discrepant interpretations (overcalls). Of these, 12 involved possible fractures, typically caused by the misinterpretation of asymmetric or accessory sutures (5 cases) and vascular grooves (3 cases) (Fig 3A). Seven involved the misinterpretations of ICH, caused by overcalling punctate regions of hyperattenuation in the brain parenchyma and subarachnoid space, likely caused by vessels (3 cases) and typical streak artifacts along the inner table (2 cases). Two of these type 5 discrepancies were classified as resulting in inappropriate therapy (performance of unnecessary follow-up imaging examinations). One of these cases was the overcall of a normal cisterna magna on CT as an arachnoid cyst. Another was the overcall of a potential aneurysm that represented normal pericallosal arteries on CT. Both resulted in inappropriate MR imaging examinations (these patients did not have symptoms that would otherwise warrant MR imaging). No overcalls resulted in inappropriate medical or surgical intervention.

Discrepancy Etiology

Discrepancies were subclassified by etiology, summarized in Table 3. Misinterpreted fractures were the most common discrepancy (20%) (Fig 3). There were 11 calvarial fractures, 4 temporal bone or skull base fractures, 11 facial/orbital fractures (2 nasal, 1 mandibular, 5 orbital wall, 1 zygomatic arch, 2 maxilla), and 2 vertebral fractures (1 was missed edema within an upper thoracic vertebral body on MR imaging, and another was missed subtle wedging of the C6 vertebral body on CT). Misinterpreted ICH was the next most common discrepancy (16%) (Fig 4). Locations of misinterpreted ICH (23 cases, all traumatic) included 11 subdural, 5 subarachnoid, 4 intrapa-
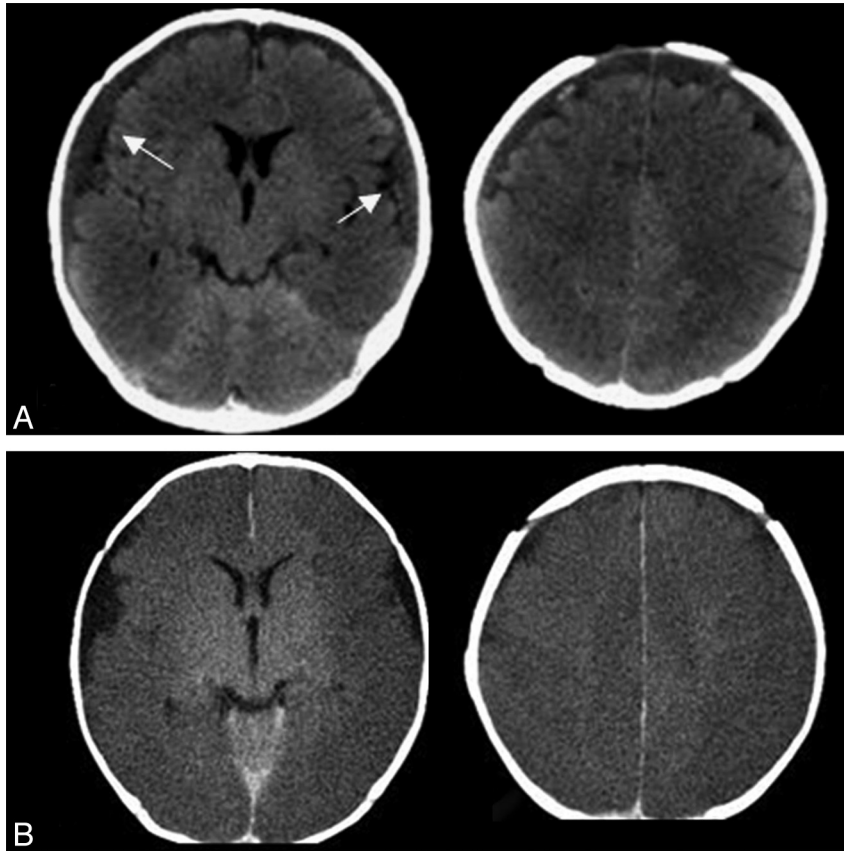
Fig 1. Type 1 discrepancy (major, life-threatening). A 15-week-old male infant found unresponsive. *A*, CT at the time of presentation (top) was initially interpreted by the trainee as a small falcine subdural hematoma and normal brain parenchyma. The staff final report documented bilateral subdural collections with regions of hemorrhages (*arrows*) and diffuse loss of gray-white matter differentiation consistent with edema/ischemia. *B*, Follow-up head CT (bottom) 7 hours later demonstrates evolution of diffuse cortical edema consistent with diffuse ischemic injury. At the time of clinician notification of the discrepancy, the patient was already being treated for presumed diffuse brain injury clinically and was on ventilator respiratory support and intracranial pressure management. Further investigation of the clinical scenario coupled with the imaging abnormalities were concerning for non-accidental trauma with a subsequent confession of inflicted injury by a caretaker.

renchymal, and 1 intraventricular. Disagreements in evaluation of ventricular size assessment were noted in 11 (8%) cases. Incorrectly identified diffuse brain edema accounted for 9 (6%) errors.

For CT examinations of the face, orbit, and neck (20 discrepant cases), the most common misinterpretations were related to trauma of the orbit and facial structures (10 cases: 6 fracture misinterpretations and 4 missed soft-tissue findings) and to inflammatory disease (7 cases) (Fig 5). Of these 20 cases, 5 were overcalls. There were a relatively small number of MR imaging examinations with initial trainee interpretations during the study period. There were 12 discrepancies, 7 of which involved MR imaging scans of the brain. Of these, 6 were related to misinterpretation of focal (5 cases) or diffuse (1 case) signal changes (Fig 6). Of these, 2 were overcalls. There were misinterpretations in 4 spine MR imaging cases (Fig 6): a missed pars defect of L5, missed subtle traumatic edema in upper thoracic vertebral bodies, missed sub-endplate signal-intensity changes (likely degenerative), and overcall of cord signal-intensity abnormality in the cervical spine (related to artifacts).

Discrepancies by Level of Training

Most (92%) off-hours examinations were interpreted by pediatric radiology fellows during the study period (Table 5). A much smaller number of examinations were initially interpreted by residents (8%, 281 cases). The overall discrepancy rate for fellows was 3.7% compared with 8.5% (24/281) for residents (residents in their third and fourth years of training had the highest discrepancy rate overall, 11.7% and 14% respectively). The overall discrepancy rate for residents in the first or second year of training was 1.6% (2/123) compared with 12.36% in those in the third and fourth years of residency training. There were statistically significant differences in discrepancy rates comparing residents and fellows, third- and fourth-year residents and fellows, and third- and fourth-year residents and first- and second-year residents (Table 5).

Outcomes and Clinical Management Changes

Changes in management or outcome based on the discrepant trainee interpretations are listed in Table 2. Most errors had no identifiable effect on patient management (68%) or only resulted in watchful clinical follow-up or follow-up imaging without new medical or surgical intervention (30%). In 3 patients (2%), a direct treatment change occurred that was thought to be related to the discrepancy (neurosurgical consultation and extradural hemorrhage evacuation [Fig 2], ventricular drain placement for enlarging ventricles, and neurosurgical admission for nonemergent subdural drain placement). Of the 4 patients with life-threatening (type 1) discrep-
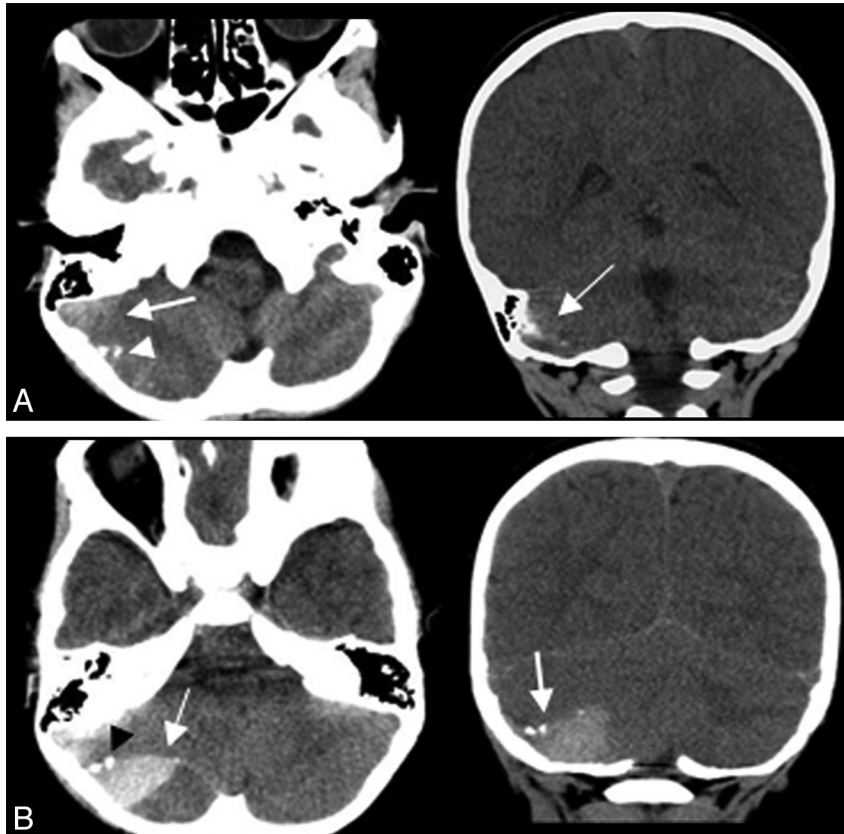
Fig 2. Type 1 discrepancy (major, life-threatening). A 3-year-old boy with a head injury and vomiting. *A*, Initial head CT axial images (left) and coronal reformat (right) were interpreted by the trainee as possible ICH versus streak artifacts (*arrows*). The attending radiologist thought the finding was artifacts, marking this as a discrepancy. The patient was monitored clinically for 24 hours in the hospital with no concerning symptoms and was discharged without a repeat head CT. The patient presented 3 days later with worsening headache and vomiting. *B*, Repeat head CT demonstrates a large mixed-attenuation posterior fossa epidural hematoma with mass effect in the location of the previously questioned artifacts (*arrows*). Small foci of bone attenuation are identified, displaced from the inner table (*black arrowhead*), retrospectively identified on the previous scan (*A*, *white arrowhead*). The patient was immediately taken to surgery for evacuation and made a complete recovery with no permanent deficits.

ancies related to misinterpreted brain edema, no direct treatment change occurred related to the final report. Three cases were secondary to missed findings of diffuse brain ischemic injury, and 1 was related to posterior reversible encephalopathy syndrome. In this case, the patient was already admitted and was being treated for hypertension and clinically presumed hypertensive encephalopathy, though follow-up imaging was recommended (type 2 outcome). For the 3 patients with missed diffuse brain ischemic injury, all were admitted by the time of the final report. Diffuse brain edema and ischemic injury had been suspected clinically, resulting in treatment with ventilatory respiratory support and intracranial pressure management.

Discussion
The trainee discrepancy rate for pediatric neuroimaging studies was low at 4.1%. This is similar to that in previous studies that evaluated trainee discrepancy rates for neuroimaging studies performed primarily in adults. The overall trainee discrepancy rates described in studies primarily evaluating adult patients have ranged from 2% to 8% (most commonly 3%–5%).[1-7] Most of the discrepancies in our study, as in previous reports, were classified as minor.

Concerning the type of findings most commonly missed on pediatric neuroimaging studies, fractures were the most common, occurring in 28 cases and making up 20% of all trainee

discrepancies. Missed fractures were also common in some of the studies of adult patients.[1,5] ICH was the second most common source of discrepancies, occurring in 23 cases and making up 16% of all trainee misinterpretations. Other common misinterpreted findings included focal brain parenchymal abnormalities and misinterpretations of ventricular size. In our experience, multiplanar reformats are an important tool in the interpretation of head CT for trauma, especially in the depiction of subtle extra-axial hemorrhage (Fig 4*C*, -*D*). This technique has been recently documented in the literature[9] and has been adopted as standard protocol at our institution in this patient population.

Overcalls were not uncommon in our series, occurring in 37 cases. The most frequent cause of trainee overcalls was the misidentification of fractures, typically related to asymmetric or accessory sutures. We believe the presence of unfused or partially fused sutures in children may play a role in the high incidence of overcalled fractures. As previously reported in the literature,[10,12] multiplanar 2D and especially 3D reconstructions are helpful for better delineation of an area of questionable skull abnormality, and the use of this technique is emphasized to our trainees in our educational curriculum.

There were only 6 major (life-threatening) discrepancies, for a rate of 0.17%. This low rate of major discrepancies is similar to that in previous studies evaluating trainee misinterpretations.[1-7] The most common cause of major discrepancies
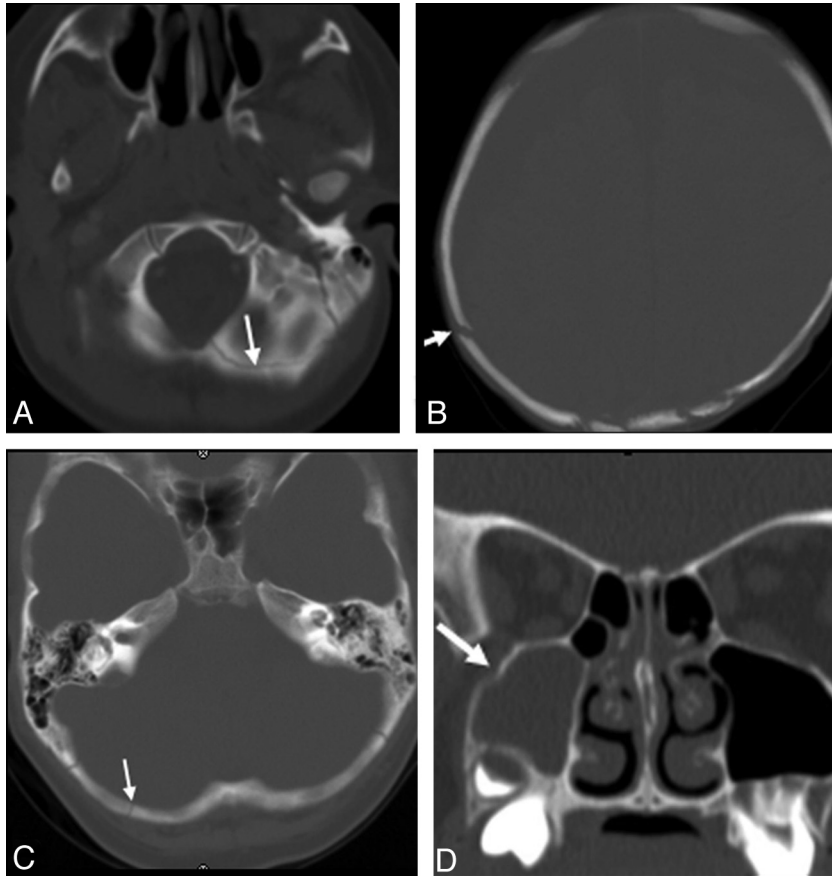
Fig 3. Examples of misinterpreted fractures (*arrows*). *A*, Overcall of the normal posterior intraoccipital synchondrosis as a fracture. *B*, Right parietal calvarial fracture. *C*, Nondisplaced right occipital fracture. *D*, Minimally displaced right maxillary fracture.

was the misinterpretation of brain edema (4 cases). Misinterpretation in these cases was due to failure in recognizing diffuse (3 cases) or focal (1 case) loss of gray-white matter differentiation. The 3 cases in which diffuse loss of gray-white matter differentiation was missed were the result of diffuse brain edema from hypoxic-ischemic insult. This is an important finding in our study that differs from that in previous studies based primarily on the adult population. Although misinterpretation of brain attenuation has been described as a cause of major discrepancies in some adult studies,[8] those cases were typically the result of focal stroke or vasogenic edema, rather than diffuse brain insult as in our study. Findings of diffuse brain edema can be subtle,[13,14] and misidentification may result in inadequate or delayed treatment. Although in our study the correct identification of these findings by the staff radiologist did not lead to a specific treatment change on the part of the clinical team because the diagnosis was also suspected clinically, in a different clinical scenario, the misinterpretation could have been much more clinically significant. On the basis of the findings in this study, the imaging appearances of diffuse brain edema/injury on CT should be specifically emphasized during resident and fellow training.

As has been previously suggested,[4] the final interpretation by the staff radiologist was an imperfect criterion standard. A recent study noted an 8.3% discrepancy rate between different board-certified subspecialty-trained neuroradiologists in interpreting outside imaging studies in a tertiary care setting, though only 1.2% were clinically important.[15] In our study, 15

(9.5%) of the reports initially described as discrepancies in trainee interpretation by the original interpreting radiology staff did not harbor a true discrepancy when subsequently evaluated by the 2 expert reviewers. In most of these cases, the faculty essentially restated the trainee interpretation with no change in the original findings or differential diagnosis. In 4 cases (2.8% of reviewed examinations), the expert reviewers disagreed with the initial attending radiologist's interpretation. In 1 important case (Fig 2), the initial trainee interpretation suggested the correct finding, only to be overruled by the attending radiologist's interpretation. Follow-up CT performed to assess new symptoms showed an enlarging epidural hematoma necessitating surgical evacuation. Because the trainee was not definitive in identifying hemorrhage, as well as the importance of the case, this was classified as a type 1 trainee misinterpretation for the purposes of this study.

The most common neuroimaging examination type in which discrepancies were noted was head CT (72%). This is also the most common neuroimaging study performed on off-hour shifts. Educational programs aimed at training rotating residents and fellows for off-hour shifts should focus on CT and, in particular, head CT interpretation. Of CT examinations, those of the face and neck had the highest discrepancy rates and should also be emphasized early in trainee education, particularly with regard to fracture patterns and normal variants that may cause diagnostic confusion.

Although the relatively small number of examinations read by resident trainees (versus fellows) in this study and our study
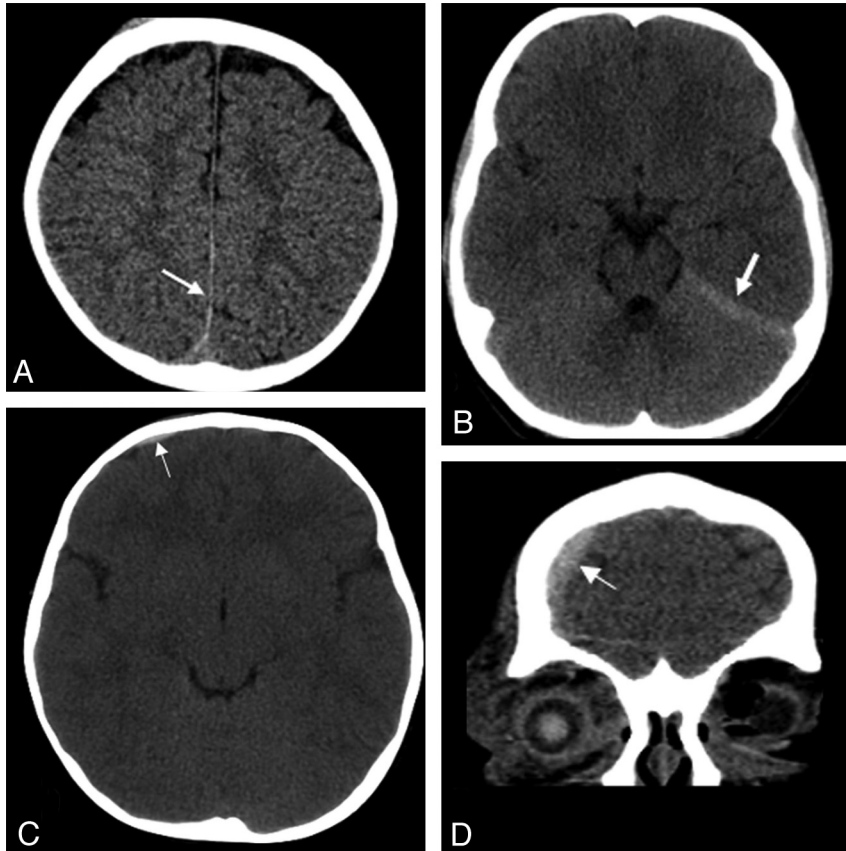
Fig 4. Examples of misinterpreted ICH (*arrows*). *A*, Small subdural hemorrhage along the posterior interhemispheric fissure. *B*, Small subdural hemorrhage along the left tentorial leaflet seen as asymmetric hyperattenuation compared with the contralateral side. *C* and *D*, Small right frontal extra-axial hemorrhage (*C*), more obvious on subsequently performed coronal reformats (*D*).
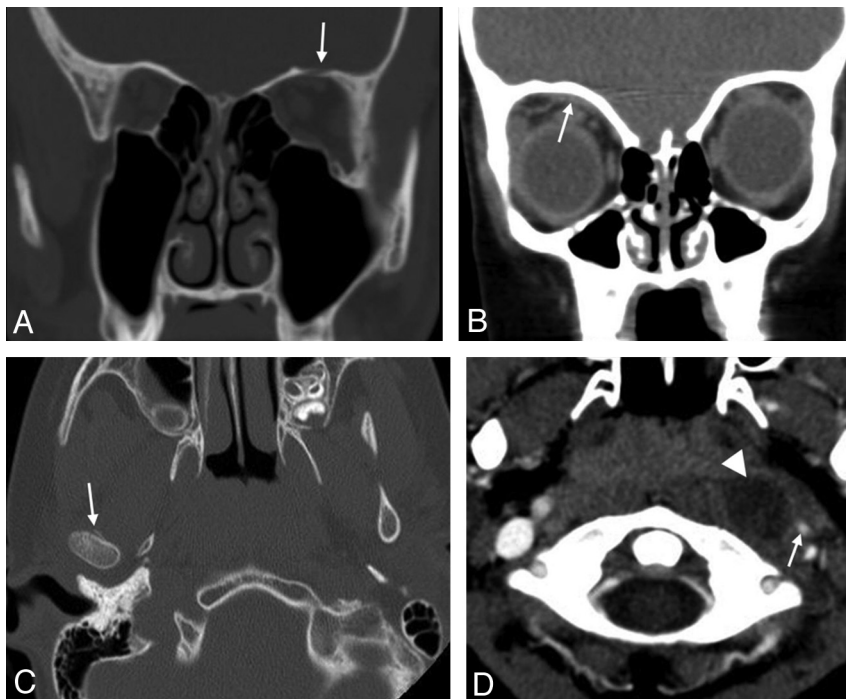

Fig 5. Examples of misinterpreted examinations of the face, neck, and orbits. *A*, Nondisplaced orbital roof fracture (*arrow*). *B*, Small extraconal hematoma (*arrow*) in a patient with an orbital roof fracture. *C*, Subtle fracture of the right mandibular condyle. *D*, Missed caliber change of the left internal carotid artery, secondary to retropharyngeal inflammatory disease and abscess (correctly identified by the trainee, *arrowhead*).

design limit our ability to reach definitive conclusions regarding the effect of resident training experience and error rate, there are some significant findings: Third- and fourth-year residents committed more errors (12.4%) than fellows (3.7%)
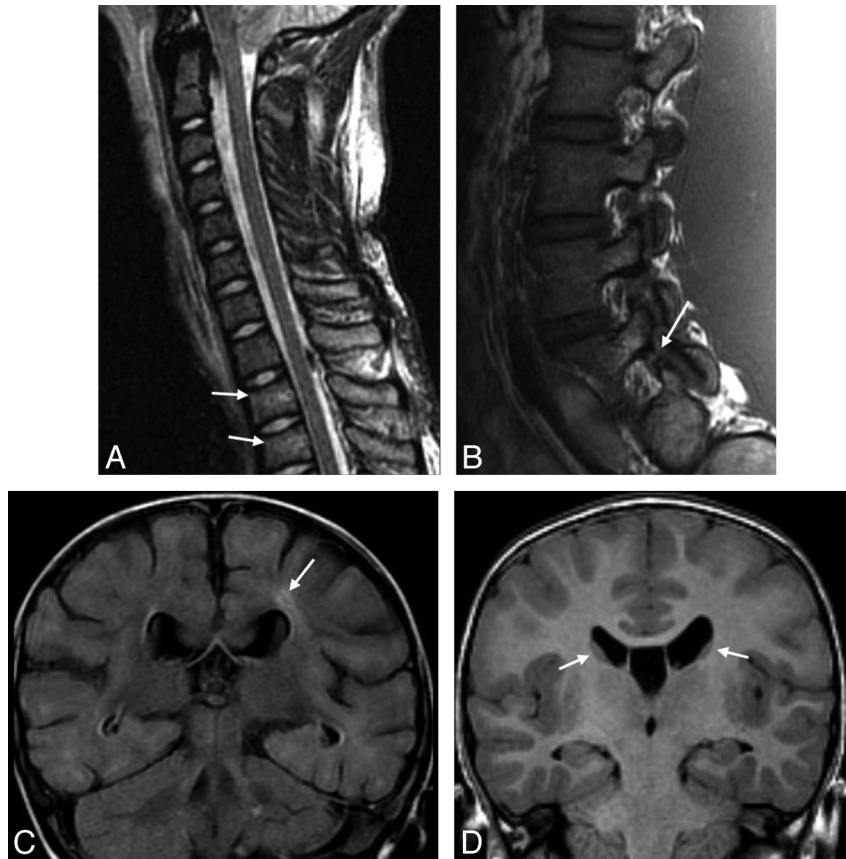
Fig 6. Examples of misinterpreted MR imaging examinations. *A,* Missed subtle upper thoracic vertebral body edema and slight height loss consistent with mild compression fractures (*arrows*). *B,* Missed L5 pars defect (*arrow*). *C,* Missed focal periventricular signal intensity in a 12-month-old child (fluid-attenuated inversion recovery sequence, *arrow*). *D,* Normal caudate nuclei misinterpreted as periventricular nodular heterotopia (*arrows*).

Table 5: Discrepancy rate by level of training

Year of Training	Discrepancies/Exams Read	Discrepancy Rate (95% CI)[a]
1	0/23	0%
2	2/80	2.5% (0.0–6.54)
3	15/128	11.7% (5.76–17.68)
4	7/50	14.0% (5.38–24.62)
Fellow	119/3215	3.7% (3.03–4.37)

[a] The combined discrepancy rate for residents was 8.54% (95% CI, 5.09%–11.99%). The combined discrepancy rate for first- and second-year residents was 1.94% (95% CI, 0.0%–5.08%). The combined discrepancy rate for third- and fourth-year residents was 12.36%, (95% CI, 7.24%–17.48%). There were statistically significant discrepancy rates comparing residents and fellows, third- and fourth-year residents and fellows, and third- and fourth-year residents and first- and second-year residents.

or first- and second-year residents (1.9%) during the study period. This finding is not intuitive, because error rates would be expected to decrease with more experience. Similar trainee-level patterns of discrepancies have been described by other researchers with body CT (major discrepancies)[16] and head CT[5] interpretations. Because of the small number of cases read by residents in this study, bias from sampling error cannot be discounted as a cause.

Independent reading cannot be assured in all cases interpreted by trainees in our study. Fellows and residents in our institution share the same reading room and are on call at the same time, and the fellows can be easily consulted on cases before dictation. It may be that the first- and second-year residents seek out second opinions from the fellows more commonly than the third- and fourth-year residents who are more

confident in their ability to read independently. Experience in pediatric-specific imaging is still relatively limited for most residents during training, and pediatric rotations are commonly interrupted by many months of nonpediatric imaging. It seems reasonable to have residents of all training levels undergo basic pediatric neuroradiology–specific training, focusing on commonly misinterpreted findings, before each rotation throughout their training. Building an educational curriculum around common errors has been shown to improve trainee performance,[9,17] and continued diligence in resident training with feedback on commonly missed findings is important to help limit these errors.

Concerning patient management and outcome, 98% of the discrepancies noted in this study had either no effect on clinical management or outcome or had no direct treatment change but did lead to imaging or clinical follow-up related to the discrepancy. There was no permanent harm related to the discrepancies and no directly related mortality. Some of the misinterpretations in a different clinical scenario could have had a more direct and significant impact on outcome (diffuse edema, missed epidural hematoma). Although they were small in number in our study, the performance of unnecessary follow-up imaging examinations based on a misinterpretation adds to the cost of medical care and, in the case of CT, adds an additional risk of unnecessary exposure to ionizing radiation.

We recognize some limitations to our study, in addition to those described above. First, the identification of a report as discrepant is dependent on the faculty using the official macro

for changing preliminary reports. Faculty use of this process is randomly audited as part of our quality assurance process and has demonstrated high compliance during the past 2 years. Second, the assessment of outcome was subjective, based on chart review and the best clinical judgment of the expert reviewers. Although this type of assessment can potentially lead to bias, every effort was made, by detailed chart review, to accurately identify any changes in clinical management based on the misinterpretation. Although it has some limitations, this method is practical and is similar to that in previous studies of trainee misinterpretations.[1,4]

Conclusions

The discrepancy rate of trainee interpretation of pediatric neuroimaging studies is low at 4.1% and similar to the rates described in adult imaging series.[1-7] Most of the discrepancies occurred in the interpretation of pediatric head CT. The rate of major life-threatening discrepancies was very low at 0.17% and was related to misinterpretation of diffuse cerebral edema or ICH. The overwhelming majority of discrepancies had either no or minimal effect on clinical management. The most common findings associated with minor discrepancies included misinterpretation of fractures, ICH, ventricular size, and focal brain parenchymal abnormalities. CT studies of the face, orbits, and neck had the highest rate of discrepancies. Education programs aimed at improving trainee performance in the off-hour interpretation of pediatric neuroimaging studies should be aimed at those specific findings that are either commonly missed or are sometimes missed and potentially life-threatening.

References

1. Miyakoshi A, Nguyen QT, Cohen WA, et al. **Accuracy of preliminary interpretation of neurologic CT examinations by on-call radiology residents and assessment of patient outcomes at a level 1 trauma center.** *J Am Coll Radiol* 2009;6:864–70

2. Walls J, Hunter N, Brasher PM, et al. **The DePICTORS study: discrepancies in preliminary interpretation of CT scans between on-call residents and staff.** *Emerg Radiol* 2009;16:303–08

3. Filippi CG, Schneider B, Burbank HN, et al. **Discrepancy rates of radiology resident interpretations of on-call neuroradiology MR imaging studies.** *Radiology* 2008;249:972–79

4. Strub WM, Leach JL, Tomsick T, et al. **Overnight preliminary head CT interpretations provided by residents: locations of misidentified intracranial hemorrhage.** *AJNR Am J Neuroradiol* 2007;28:1679–82

5. Wysoki MG, Nassar CJ, Koenigsberg RA, et al. **Head trauma: CT scan interpretation by radiology residents versus staff radiologists.** *Radiology* 1998;208:125–28

6. Strub WM, Leach JL, Ying J, et al. **First year radiology residents not taking call: will there be a difference?** *Emerg Radiol* 2007;13:231–35

7. Erly WK, Berger WG, Krupinski E, et al. **Radiology resident evaluation of head CT scan orders in the emergency department.** *AJR Am J Roentgenol* 2002;23:103–07

8. Strub WM, Vagal A, Tomsick T, et al. **Overnight resident preliminary interpretations on CT examinations: should the process continue?** *Emerg Radiol* 2006;13:19–23

9. Gutmark R, Halsted MJ, Perry L, et al. **Use of computer databases to reduce radiograph reading errors.** *J Am Coll Radiol* 2007;4:65–68

10. Sanchez T, Stewart D, Walvick M, et al. **Skull fractures vs. accessory sutures: how can we tell the difference?** *Emerg Radiol* 2010;17:413–18

11. Wei SC, Ulmer S, Lev MH, et al. **Value of coronal reformations in the CT evaluation of acute head trauma.** *AJNR Am J Neuroradiol* 2010;31:334–339

12. Ringl H, Schernthaner RE, Schueller G, et al. **The skull unfolded: a cranial CT visualization algorithm for fast and easy detection of skull fractures.** *Radiology* 2010;255:553–62

13. Robertson RL, Ball WS, Barnes PD. **Skull and brain.** In: Kirks DR, Griscom NT, eds. *Practical pediatric imaging: diagnostic radiology of infants and children.* 3rd ed. Philadelphia: Lippincott; 1998:65–200

14. Moore MJ, Vagal AS, Strub WM, et al. **Reducing the gray zone: imaging spectrum of hypoperfusion and hypoxic brain injury in adults.** *Emerg Radiol* 2009;17:123–30

15. Zan E, Yousem DM, Carone M, et al. **Second-opinion consultations in neuroradiology.** *Radiology* 2010;255:135–41

16. Wechsler RJ, Spettell CM, Kurtz AB, et al. **Effects of training and experience in interpretation of emergency body CT scans.** *Radiology* 1996;199:717–20

17. Halsted MJ, Kumar H, Paquin JJ, et al. **Diagnostic errors by radiology residents in interpreting pediatric radiographs in an emergency setting.** *Pediatr Radiol* 2004;34:331–36