Second Opinion Interpretations by Specialty Radiologists at a Pediatric Hospital: Rate of Disagreement and Clinical Implications

Christopher Eakins[1]
Wendy D. Ellis[2]
Sumit Pruthi[2]
David P. Johnson[2]
Marta Hernanz-Schulman[2]
Chang Yu[3]
J. Herman Kan[4]

OBJECTIVE. The objective of our study was to identify whether a substantive difference exists between the imaging interpretations of radiologists at outside referring institutions and those of radiologists at a tertiary care children's hospital and whether such reinterpretation affects the clinical management of pediatric patients.

MATERIALS AND METHODS. This retrospective chart review examined the diagnostic imaging reports of all pediatric patients referred to a tertiary care freestanding children's hospital over a 17-month period (January 1, 2009–May 31, 2010); 773 examinations met the inclusion criteria. The original and second interpretations were compared. A fellowship-trained pediatric radiologist and neuroradiologist categorized each case using the content of the two radiology reports as agreement versus minor or major disagreement, and the results were analyzed for statistical significance. A cohort of cases in which a final diagnosis could be confirmed was also analyzed to evaluate the accuracy of both interpretations.

RESULTS. Disagreements were found in 323 of 773 reports (41.8%): 168 (21.7%) were major and 155 (20.0%), minor. Neurologic studies were most frequently requested for reinterpretation, 427 (55.2%), most commonly in the setting of trauma, 286 (67.0%). Among the 427 neuroimaging studies, major and minor disagreements occurred in 54 (12.6%) and 91 (21.3%) cases, respectively. Major disagreements most frequently observed were about the presence of fracture and hemorrhage. Among 305 body imaging cases, major and minor disagreements occurred in 99 (32.6%) and 57 (18.7%) cases, respectively. The most common setting for nontraumatic body imaging was concern for appendicitis (168/305 [55.1%]); this indication for imaging was responsible for 40.3% of major disagreements in nontraumatic abdominal imaging. Reinterpretation was rarely requested for radiographic studies (41/773 [5.3%]), which had major and minor disagreement rates of 36.6% and 17.1%, respectively. In the cohort of cases analyzed for final diagnosis, the second interpretation was more accurate than the original in 90.2% of cases with a p value of less than 0.0001.

CONCLUSION. Our findings suggest that discrepancy rates for second interpretations in studies of pediatric patients transferred to tertiary care pediatric institutions are substantial. Although the original and second interpretations in the majority of cases were in agreement, major discrepancies were prevalent—12.6% and 32.6% of neuroimaging and body studies, respectively—and the second interpretations were significantly correlated with the final diagnosis. These results indicate that interpretations by subspecialty radiologists at a point-of-care facility provide important clinical information about the pediatric patient and should be recognized by payers as integral to optimal care.

Keywords: body imaging, comparative study, diagnostic imaging, neuroimaging, pediatric imaging, second opinion, specialty radiologists

DOI:10.2214/AJR.11.7662

Received August 8, 2011; accepted after revision March 5, 2012.

Statistical support was funded in part by the Vanderbilt CTSA grant UL1 RR024975 from the National Center for Research Resources of the National Institutes of Health.

[1]Department of Internal Medicine, Gunderson Lutheran Medical Center, La Crosse, WI.

[2]Department of Radiology and Radiological Sciences, Monroe Carell Jr. Children's Hospital at Vanderbilt, Nashville, TN.

[3]Department of Biostatistics, Vanderbilt Clinical Research Center, Vanderbilt University, Nashville, TN.

[4]Department of Radiology, Texas Children's Hospital, 6701 Fannin St, Ste 470, Houston, TX 77030. Address correspondence to J. H. Kan (jhkan@texaschildrens.org).

AJR 2012; 199:916–920

0361–803X/12/1994–916

Pediatric patients transferred to tertiary care centers for definitive evaluation and treatment of problems beyond the scope of the referring institution often arrive with radiologic imaging having already been performed at that institution. As part of the care to be delivered by the pediatric specialists, a second opinion interpretation of the imaging studies by tertiary care subspecialty radiologists at the referral center is often requested, and the formal reports are incorporated into the patient's permanent medical record at the referral institution where the patient's management and treatment are determined. The reasons for these requests include the fact that patients are transferred because of complex medical issues beyond the scope of the referring institution and the belief that subspecialty radiologists at a tertiary care center have incremental and essential expertise

in more complex pediatric disease entities and their pertinent findings in diagnostic imaging examinations.

In the adult population, the reported discrepancy rate between radiologic interpretations performed at primary facilities and those performed at tertiary care referral centers ranges from 12% to 41% [1–5]; in comparisons limited to general radiologists versus subspecialty neuroradiologists, the reported discrepancy rate ranges from 1.3% to 34% [6–8]. However, to our knowledge, the rate and type of discrepancies related to subspecialty reinterpretations of radiologic studies are not known in pediatric patients.

The purpose of this study was to evaluate the discrepancy rate for second interpretations of diagnostic imaging studies of pediatric patients transferred to our pediatric hospital.

Materials and Methods

This HIPAA-compliant retrospective study was approved by our institutional review board with a waiver of informed consent. From the electronic medical records, a list was generated of all pediatric patients (age range, 0 days–18 years) who were referred between January 1, 2009 and May 31, 2010 to our tertiary care children's hospital in a major metropolitan city and whose medical record included an official second opinion interpretation of the outside studies. In all cases, the reports and images from outside referring institutions had been imported into the electronic medical record from CDs or other digital media or by digitization of films. There were 834 requests for official reinterpretation of outside studies during that time period. When more than one examination of any one patient was reinterpreted, for the purposes of this analysis, each study was considered as an independent data point, totaling 1156. Of these 1156 studies, 773 studies (66.9%) had outside interpretations available at the time of review and thus met our inclusion criteria. The study population comprised 429 boys (55.5%) and 344 girls (44.5%) with a mean age of 9.55 years (range, 1 day–18 years). There was no significant difference in age between boys and girls.

The requested second interpretations were performed during the course of normal clinical work by board-certified subspecialty radiologists. For this investigation, the review of all outside and subspecialty reports in neuroradiology was conducted by a pediatric neuroradiologist with more than 4 years of postfellowship experience, and the review of all outside and subspecialty reports in body imaging was conducted by a pediatric radiologist with more than 20 years of postfellowship experience.

The examinations included digital images of radiographs; CT examinations of the head, neck, chest, abdomen, pelvis, and extremities; and MRI examinations of the head, spine, and extremities. CT and MRI studies of the head, neck, face, and cervical spine were considered neuroradiologic examinations. CT studies of the chest, abdomen, pelvis, and extremities were considered body examinations. Radiography was considered independently for the purpose of this review.

The official outside radiology report and the official second interpretation radiology report were compared and categorized in terms of concordance or discordance using a three-category grading system. Interpretations were classified as agreement, major disagreement, or minor disagreement. Substantive differences were considered as major disagreements, which were defined as being capable of altering patient care or prognosis as determined by the reviewing pediatric radiologist or neuroradiologist; minor disagreements were defined as not typically leading to an alteration in patient management or prognosis, again as determined by the reviewing pediatric radiologist or neuroradiologist. If both major and minor disagreements were recorded in the same examination, that examination was reported as a major disagreement.

After review of the initial data, a representative cohort of 96 cases in which the diagnosis could be independently proven was reviewed for confirmation of the final diagnosis to evaluate the accuracy, and thus clinical benefit, of the second interpretations. This group was initially obtained using 50 randomly selected cases from the major disagreements in the body and neurologic interpretations, for an initial total of 100 cases. Four of these cases were excluded because the patient was discharged or lost to follow-up and no clinical or pathologic proof of diagnosis was present in the patient's record. The final diagnosis in the group of 96 patients was verified by one or more of the following methods: pathologic studies, follow-up imaging studies, or clinical follow-up.

Classification of Disagreements

Neuroradiology—Interpretive differences considered to be major disagreements included fractures, subdural hemorrhage, subarachnoid hemorrhage, epidural hemorrhage, parenchymal blood, neoplastic mass, abscess, foreign body, acute hydrocephalus, vasogenic cerebral edema, and cerebral calcifications.

Interpretive differences considered to be minor disagreements included scoliosis, additional nondisplaced traumatic fractures not necessitating management change, sinusitis, congenital cortical abnormality, arachnoid granulation, additional finding of ligamentous injury in a patient with multiple cervical spine injuries, chronic encephalomalacia, cavum septum pellucidum and septum cavum vergae, benign cystic lesion without associated mass effect, congenital anatomic malformations not necessitating management change, stable ventriculomegaly, dentigerous cyst, pseudosubluxation, optic dysplasia, and mucous retention cyst.

Body imaging—Interpretive differences considered to be major disagreements included appendicitis, omental torsion, cholecystitis, colitis, hepatitis, solid organ hematoma, solid organ laceration, fracture, small-bowel obstruction, inflammatory bowel disease, luminal perforation and free air, endometriosis, pneumothorax, pneumonia, diaphragmatic hernia, pyelonephritis, primary and metastatic neoplasia, abscess, lymphadenitis, lymphangioma, biloma, recommendation for further unnecessary studies, ascites, pelvic hematoma, thymoma, and Baker cyst.

Interpretive differences considered to be minor disagreement included incidental benign renal cyst, incidental benign hepatic cyst, splenic cyst, ovarian cyst, atelectasis, incidental pulmonary nodule, pulmonary contusion, adrenal hematoma, retractile testes, hepatic steatosis, additional undocumented traumatic injury not necessitating management change, incidental or isolated appendicolith, vascular anatomic malformation, and scoliosis.

Radiography—Interpretive differences considered to represent major disagreements included the following: nonaccidental trauma, diaphragmatic hernia, pneumonia, bronchiolitis, pulmonary edema, interstitial pulmonary disease, pneumothorax, and fracture.

Interpretive differences considered to represent minor disagreements included atelectasis, osteoarthropathy, and soft-tissue edema.

Statistical Analysis

For binary endpoints, the number of events is reported. The proportion and its 95% asymptotic or exact CI, depending on the number of events, are calculated. Statistical analyses were conducted using the statistical package SAS (version 9, SAS Institute) for Microsoft Windows.

Results

The 773 subspecialty interpretations comprised 427 neuroradiologic studies (55.2%), 305 cross-sectional body imaging studies (39.5%), and 41 radiographic studies (5.3%). Although the majority of reports showed congruence between the interpretations generated by radiologists at the referring institutions and interpretations generated by subspecialty radiologists at a tertiary care hospital (450/773 [58.2%]), there were disagreements in a sizeable minority. In 323 of 773 examinations (41.8%), review of the official interpretations identified a major disagreement in 168 of 773 cases (21.7%) or a minor disagreement in 155 of 773 cases (20.1%).

TABLE 1: Major and Minor Disagreement Rates Among Neurologic Examinations by Study Type and Anatomic Region

	Agree			Disagree								
				Total		Minor			Major			
Neurologic Study	No.	%	95% CI	No.	%	No.	%	95% CI	No.	%	95% CI	Total
CT												
Cervical spine	59	74.7	63.6–84.3	20	25.3	17	21.5	12.5–30.6	3	3.8	0.0–8.00	79
Neck	8	66.7	40.0–93.3	4	33.3	4	33.3	6.6–60.0	0	0.0	0.0–100.0	12
Face	23	53.5	38.6–68.4	20	46.5	13	30.2	16.5–44.0	7	16.3	5.24–27.3	43
Head	177	65.6	59.9–71.2	93	34.4	49	18.1	13.6–22.7	44	16.3	11.9–20.7	270
MRI and MRA												
Brain	15	65.2	45.8–84.7	8	34.8	8	34.8	15.3–54.3	0	0.0	0.0–100.0	23
Total	282	66.0		145	34.0	91	21.3		54	12.6		427

Note—MRA = MR angiography.

Fig. 1—CT scan through posterior fossa and temporal lobes of 3-year-old boy. Radiologist at outside referring institution initially diagnosed hydrocephalus, whereas specialty radiologist at tertiary care hospital recognized posterior fossa tumor as cause of hydrocephalus.

A second interpretation was requested most frequently for neurologic studies. A majority of these studies was performed for evaluation of traumatic injury (286/427, 70.0%) (Table 1). Subspecialty interpretations of neuroimaging studies disagreed with the original interpretations in 145 of 427 cases (34.0%) with 12.6% major and 21.3% minor discrepancies. Among the trauma patients, there were disagreements between the original and second interpretations in 79 of 286 cases (27.6%), 41 (14.3%) of which were considered major and 38 (13.3%), minor. The most common finding missed in this group was fracture (21/286, 7.3%). Among the nontrauma patients, there were disagreements between the original and second interpretations in 65 of 141 cases (46.1%): The disagreements were considered major in 12 cases (8.5%) and minor in 53 cases (37.6%). Figure 1 is an example of a major disagreement in this category.

Of the 305 body examinations reviewed, disagreements between radiologists at the referring institution and subspecialty radiologists occurred in 156 cases (51.1%): 99 (32.6%) were major disagreements and 57 (18.7%), minor disagreements (Table 2). Unlike neurologic studies, trauma indications represented less than half of the body imaging examinations (136/305, 44.6%), and 169 body imaging examinations were performed for nontraumatic indications. The most frequent indication for nontrauma body examinations was concern for appendicitis (64/169, 37.9%). This was also the single most frequently observed indication in a discrepant report (31/64, 48.4%) (Table 3). Cases in which the requisition stated concern for appendicitis comprised 40.3% (25/62) of reports with major disagreements in all nontraumatic body studies and 22.5% (31/138) of all disagreements observed in abdominopelvic CT interpretations specifically. Figure 2 is an example of a major disagreement in this category. When exclusively considering studies performed for nontraumatic indications, other inflammatory processes represented a major disagreement in 4.1% (n = 11) and were represented by diagnoses such as colitis (n = 7). Among body examinations performed for traumatic indications (n = 136), there were 18 major disagreements. Of these 18 cases, the most common original interpretation of a study was "normal" (n = 7, 38.9%). Other diagnoses with high disagreement rates included lymphadenopathy and small-bowel obstructions (Table 3).

Interpretations by subspecialty radiologists were requested less frequently for radiographic examinations than for other types of studies, with only 41 radiographic reports meeting our inclusion criteria during the observational period. There was a history of trauma in 18 of the 41 radiographic studies (43.9%). Among the 41 total radiographic studies, 15 (36.6%) and seven (17.1%) were found to have major and minor disagreements, respectively (Table 4). In the trauma cases, fracture was the most commonly missed finding (5/18, 27.8%). The diagnoses most commonly missed on the remaining 23 conventional radiographic studies of patients without a clinical history of trauma (56.1%) were lung abnormalities; six of these imaging studies were reinterpreted as either no lung disease or a lung disease other than the one diagnosed (i.e., reactive airway disease, bronchiolitis, or pneumonia).

Review of the final diagnosis in the cohort of 96 patients revealed one body imaging case and three neuroradiology cases in which neither the original nor second interpretation was consonant with the final diagnosis (Table 5). If these four cases are excluded, the second interpretations were more accurate in predicting final diagnosis than the original interpretations in 84.4% of neurologic cases (38/45), 95.7% of body cases (45/47), and 90.2% of combined body and neurologic cases (83/92). All p values were less than 0.0001.

Discussion

The quality of health care has become a target of increasing public scrutiny and governmental concern while radiologic evaluation has assumed an increasingly important role in the diagnosis and management of patients of all ages [5]. Patients who are referred to a tertiary care facility typically represent more com-

TABLE 2: Major and Minor Disagreement Rates Among Body Examinations by Study Type and Anatomic Region

	Agree			Disagree								
				Total		Minor			Major			
Body Study	No.	%	95% CI	No.	%	No.	%	95% CI	No.	%	95% CI	Total
CT												
Abdomen and pelvis	114	45.2	39.1–51.4	138	54.8	49	19.4	14.8–24.9	89	35.3	29.4–41.2	252
Chest	33	66.0	52.9–79.1	17	34.0	8	16.0	5.8–26.2	9	18.0	8.9–28.7	50
Angiography	1	100.0	2.5–100.0	0	0.0	0	0.0	0.0–100	0	0.0	0–100	1
MRI												
Body	1	50.0	1.3–98.7	1	50.0	0	0.0	0.0–100	1	50.0	1.3–98.7	2
Total	149	48.9		156	51.1	57	18.7		99	32.6		305

TABLE 3: Rates of Disagreement Among Abdominopelvic CT Examinations Performed for a Nontraumatic Indication

Initial Interpretation by Community Radiologist at Outside Institution	Second Interpretation by Specialty Radiologist at Tertiary Care Hospital			
	Agree		Disagree	
	No.	%	No.	%
Appendicitis	33	51.6	31	48.4
Colitis	5	41.7	7	58.3
Gastroenteritis	1	20.0	4	80.0
Hydronephrosis	1	33.3	2	66.7
Intussusception	4	57.1	3	42.9
Lymphadenopathy	3	42.9	4	57.1
Normal	12	41.4	17	58.6
Small-bowel obstruction	3	60.0	2	40.0
Soft-tissue mass	2	40.0	3	60.0
Total	64	46.7	73	53.3

plex cases than those who do not need such a transfer, and appropriate diagnostic information is a requisite for the delivery of safe and quality care by the clinicians receiving the patients and entrusted with their care. Our study shows a substantial discrepancy rate between the interpretations of radiologists at the referring institution and those of subspecialty radiologists at a pediatric tertiary care institution. Additionally, our results show that there is a statistically significant correlation between the second opinion interpretations and the final diagnoses, thus assuming a crucial role in patient management decisions. The added value of the point-of-care second interpretations can be viewed from the medical perspective of guiding and expediting appropriate treatment as well as from the financial perspective of avoiding unnecessary studies, at times incurring additional patient radiation exposure, when initial examinations need to be repeated or when additional examinations are suggested by the radiologist at an outside institution. However, interpretation of outside studies represents a substantive additional workload for specialty radiologists at a referral center, one of which is currently underrecognized and largely unfunded [9].

Discrepancy rates currently reported in the literature cover a relatively wide spectrum, between 1.3% and 41% [1–8]; this large range of rates likely reflects the variability in the groups examined by the various investigators: from comparisons of radiologists and subspecialists within the same group to comparisons of interpretations among biopsy-proven subspecialty patients. Our reported discrepancy rates reflect the rates of discrepancy at a children's hospital and compare interpretations made at the referring institutions with those of experienced specialty radiologists at a tertiary care children's hospital in an unselected pediatric population. The study group patients range in age from neonates to teenagers and have conditions so complex that transfer to our institution was required. Our major disagreement rates of 14.3% and 32.6% for neurologic and body examinations, respectively, encompass conditions such as fractures, appendicitis, and epidural hemorrhage, all of which imply substantial alterations in management and prognosis. The frequency and type of major discrepancies as well as the improved accuracy of the second interpretations in relation to the final clinical diagnosis underscores the added value of interpretations by specialty radiologists to optimize the care that the patient's transfer to a tertiary care pediatric hospital was intended to accomplish.

Fig. 2—CT scan through pelvis of 4-year-old girl. Radiologist at outside referring institution initially interpreted study as "normal," whereas specialty radiologist at tertiary care hospital recognized distended appendix (*arrow*) with increased enhancement, periappendiceal inflammation, and small fecalith.

TABLE 4: Major and Minor Disagreement Rates Among Radiographic Studies by Anatomic Region

	Agree			Disagree								
				Total		Minor			Major			
Radiographic Study	No.	%	95% CI	No.	%	No.	%	95% CI	No.	%	95% CI	Total
Chest	5	26.3	6.5–46.1	14	73.7	5	26.3	6.5–46.1	9	47.4	24.9–69.8	19
Extremity	11	78.6	57.1–100	3	21.4	1	7.1	0.2–20.6	2	14.3	1.8–42.8	14
Bone series	1	20.0	0.5–71.6	4	80.0	0	0.0	0–100.0	4	80.0	28.4–99.5	5
Abdomen	1	50.0	28.4–99.5	1	50.0	1	50.0	1.26–98.7	0	0.0	1.26–98.7	2
Spine	1	100.0	2.5–100	0	0.0	0	0.0	0–100.0	0	0.0	0–100.0	1
Total	19	46.3		22	53.7	7	17.1		15	36.6		41

TABLE 5: Cohort of Cases Correlating Outside and Second Interpretations With Final Diagnosis

Major Disagreements	Initial Read Was Accurate	Second Read Was Accurate	Neither Read Was Accurate	Total No. of Cases
Body imaging	2	45	1	48
Neurologic imaging	7	38	3	48
Total no. of cases	9	83	4	96

There are several limitations of our retrospective study. First, this study evaluates only those radiology reports in which an official second interpretation was requested. Radiologic studies in which a second opinion was not rendered because it was not requested by the treating physician are obviously not included. Therefore, the discrepancy rates that we have found may be higher than if second interpretations of all transferred patients' studies had been performed. Second, there is an inherent interobserver variability, which may represent a component of the observed discrepancies. Third, there is subjectivity in the decision whether a discrepancy constitutes a major or minor disagreement, particularly in the absence of detailed knowledge of the patient and treatment decision tree at the time of the initial clinical encounter.

In summary, our study shows a high rate of major disagreements between interpretations of pediatric imaging studies by generalist community radiologists and those of specialty radiologists at a tertiary care pediatric hospital; these disagreements carry substantial implications for subsequent management and substantial implications for allocation of medical resources. These findings lend support to the premise that third-party payers should recognize the need for interpretation expertise by specialty or tertiary care radiologists at the point-of-care pediatric institution to which the child has been transferred and where these findings will contribute and extensively influence the management and treatment decisions for the patient.

References

1. Tilleman EH, Phoa SS, Van Delden OM, et al. Reinterpretation of radiological imaging in patients referred to a tertiary referral centre with a suspected pancreatic or hepatobiliary malignancy: impact on treatment strategy. *Eur Radiol* 2003; 13:1095–1099
2. Gollub MJ, Panicek DM, Bach AM, Penalver A, Castellino RA. Clinical importance of reinterpretation of body CT scans obtained elsewhere in patients referred for care at a tertiary cancer center. *Radiology* 1999; 210:109–112
3. Loevner LA, Sonners AI, Schulman BJ, et al. Reinterpretation of cross-sectional images in patients with head and neck cancer in the setting of a multidisciplinary cancer center. *AJNR* 2002; 23:1622–1626
4. Kalbhen CL, Yetter EM, Love L, Moncada R, Lawson TL, Albain KA. Outside film reviews of thoracic imaging studies in oncology patients. (abstract). *AJR* 1998; 170(suppl):77
5. Sung JC, Sodickson A, Ledbetter S. Outside CT imaging among emergency department transfer patients. *J Am Coll Radiol* 2009; 6:626–632
6. Jordan MJ, Lightfoot JB, Jordan JE. Quality outcomes of reinterpretation of brain CT imaging studies by subspecialty experts in neuroradiology. *J Natl Med Assoc* 2006; 98:1326–1328
7. Zan E, Yousem DM, Carone M, Lewin JS. Second opinion consultation in neuroradiology. *Neuroradiology* 2010; 255:135–141
8. Briggs GM, Flynn PA, Worthington M, Rennie I, McKinstry CS. The role of specialist neuroradiology second opinion reporting: is there added value? *Clin Radiol* 2008; 63:791–795
9. Bramson RT, Taylor GA. SOS: can we save pediatric radiology? *Radiology* 2005; 235:719–721