Whole-Body MR Imaging for Staging of Malignant Tumors in Pediatric Patients: Results of the American College of Radiology Imaging Network 6660 Trial[1]

Marilyn J. Siegel, MD
Suddhasatta Acharyya, PhD
Frederic A. Hoffer, MD
J. Brad Wyly, MD
Alison M. Friedmann, MD, MSc
Bradley S. Snyder, MS
Paul S. Babyn, MD
Geetika Khanna, MD, MS
Barry A. Siegel, MD

Purpose: To compare whole-body magnetic resonance (MR) imaging with conventional imaging for detection of distant metastases in pediatric patients with common malignant tumors.

Materials and Methods: This institutional review board–approved, HIPAA-compliant, multicenter prospective cohort study included 188 patients (109 male, 79 female; mean age, 10.2 years; range, < 1 to 21 years) with newly diagnosed lymphoma, neuroblastoma, or soft-tissue sarcoma. Informed consent was obtained and all patients underwent noncontrast material–enhanced whole-body MR imaging and standard-practice conventional imaging. All images were reviewed centrally by 10 pairs of readers. An independent panel verified the presence or absence of distant metastases. Detection of metastasis with whole-body MR and conventional imaging was quantified by using the area under the receiver operating characteristic curve (AUC). The effects of tumor subtype, patient age, and distant skeletal and pulmonary disease on diagnostic accuracy were also analyzed.

Results: Of the 134 eligible patients, 66 (33 positive and 33 negative for metastasis) were selected for image review and analysis. Whole-body MR imaging did not meet the noninferiority criterion for accuracy when compared with conventional imaging for detection of metastasis (difference between average AUCs was −0.03 [95% confidence interval: −0.10, 0.04]); however, the average AUC for solid tumors was significantly higher than that for lymphomas (P = .006). More skeletal metastases were detected by using whole-body MR imaging than by using conventional imaging (P = .03), but fewer lung metastases were detected (P < .001). Patient age did not affect accuracy.

Conclusion: The noninferior accuracy for diagnosis of distant metastasis in patients with common pediatric tumors was not established for the use of whole-body MR imaging compared with conventional methods. However, improved accuracy was seen with whole-body MR imaging in patients with nonlymphomatous tumors.

Supplemental material: *http://radiology.rsna.org/lookup/suppl/doi:10.1148/radiol.12112531/-/DC1*

[1] From the Mallinckrodt Institute of Radiology and Siteman Cancer Center, Washington University School of Medicine, 510 S Kingshighway Blvd, St Louis, MO 63110 (M.J.S., G.K., B.A.S.); Center for Statistical Sciences, Brown University, Providence, RI (S.A., B.S.S.); Department of Radiology, University of Washington, Seattle, Wash (F.A.H.); Department of Radiology, Egleston Children's Hospital CHOA, Emory University School of Medicine, Atlanta, Ga (J.B.W.); Massachusetts General Hospital for Children, Boston, Mass (A.M.F.); and Department of Medical Imaging, University of Saskatchewan, Saskatoon, Sask, Canada (P.S.B.). Received November 24, 2011; revision requested January 4, 2012; revision received June 5; accepted July 20; final version accepted August 15. **Address correspondence to** M.J.S. (e-mail: *siegelm@mir.wustl.edu*).

In the evaluation of malignant diseases, the accurate determination of tumor burden at diagnosis is a critical factor in planning treatment and predicting outcome. The frequency of distant metastasis of malignant tumors in pediatric patients varies according to disease type, from 10%–12% for sarcomas to 70% for neuroblastomas with unfavorable prognostic factors (1–4). Computed tomography (CT), magnetic resonance (MR) imaging, various scintigraphic methods and, more recently, positron emission tomography (PET) with fluorine 18 fluorodeoxyglucose (FDG) are important techniques for the staging and management of pediatric solid tumors. CT, scintigraphy and FDG PET, and PET/CT all involve exposure to ionizing radiation, which has been associated with an increased lifetime risk of cancer (5). As survival rates for childhood malignant tumors continue to improve, with 80% of patients expected to reach adulthood, the long-term sequelae of ionizing radiation are increasingly important (6,7). Thus, alternative imaging methods that do not use ionizing radiation but allow similar diagnostic accuracy would be particularly attractive. Because of its inherently high contrast resolution, MR imaging is the most important alternative imaging method.

Results of several studies in adults have shown the usefulness of whole-body MR imaging in oncologic evaluation (8–15). Results of several studies in children with a variety of tumors showed that whole-body MR imaging was superior to bone scintigraphy for allowing detection of skeletal metastases (16–22). Whole-body MR imaging also has been shown to be as effective as scintigraphy with radioiodine-labeled meta-iodobenzylguanidine (MIBG) in the detection of extraskeletal metastases of neuroblastomas (17,22). Important limitations of these studies were small sample size, the absence of histologic confirmation, pretreatment of some patients, and nonuniformity of experimental or reference-standard imaging for individual patients or cohorts.

The potential of whole-body MR imaging led us to propose a study comparing diagnostic accuracy when whole-body MR imaging versus when conventional imaging modalities are used for detection of distant metastases in patients with common pediatric tumors. We believed that if whole-body MR imaging could allow accurate detection of distant metastases, this technique could be a substitute for ionizing radiation–based imaging methods for the staging of common malignant tumors in the pediatric population.

Advances in Knowledge

- The study results did not establish that diagnostic accuracy with whole-body MR imaging was noninferior to that with conventional imaging for detection of distant metastatic disease in patients with common pediatric tumors.
- Whole-body MR imaging allows for more accurate detection of metastatic disease in pediatric patients with solid tumors than in those with lymphomas (average area under the receiver operating characteristic curve, 0.92 vs 0.72, respectively; $P = .006$).
- Whole-body MR imaging allows for detection of more skeletal lesions, on average, than does conventional imaging ($P = .03$) but does not enhance detection of lung metastases (average sensitivity, 0.53 vs 0.83, respectively; $P < .001$).

Materials and Methods

The American College of Radiology Imaging Network (ACRIN), a cooperative group that is funded by the National Cancer Institute, conducted a prospective multi-institutional study of whole-body MR imaging for the evaluation of pediatric malignant tumors (ACRIN 6660). Twenty institutions participated in the study. Each participating site obtained institutional review board approval before patient recruitment and conducted the trial in compliance with the Health Insurance Portability and Accountability Act (or with applicable equivalent regulations for sites outside of the United States). Parental written informed consent was obtained for participants younger than 18 years old, and assent was obtained from older minors. Written informed consent was obtained from participants aged older than 18 years.

Implication for Patient Care

- Whole-body MR imaging as the initial study of choice for staging of pediatric malignant tumors is not justified, especially for tumors that have spread to the lungs and solid organs.

Study Patients

Study participant recruitment began in November 2004 and was completed in June 2007.

Inclusion criteria.—Inclusion criteria were male and female patients aged 21 years or younger; a clinically suspected or pathologically proved diagnosis of neuroblastoma, Hodgkin

Published online before print
10.1148/radiol.12112531 **Content code:** PD

Radiology 2013; 266:599–609

Abbreviations:
ACRIN = American College of Radiology Imaging Network
AUC = area under the receiver operating characteristic curve
CI = confidence interval
FDG = fluorine 18 fluorodeoxyglucose
MIBG = meta-iodobenzylguanidine
STIR = short tau inversion-recovery

Author contributions:
Guarantor of integrity of entire study, M.J.S.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; literature research, M.J.S., F.A.H., J.B.W.; clinical studies, M.J.S., F.A.H., A.M.F., P.S.B., G.K., B.A.S.; experimental studies, M.J.S., F.A.H.; statistical analysis, S.A., B.S.S.; and manuscript editing, M.J.S., S.A., F.A.H., J.B.W., A.M.F., B.S.S., G.K., B.A.S.

Funding:
This research was supported by the National Institutes of Health (grants U01 CA079778 and U01 CA080098).

Conflicts of interest are listed at the end of this article.

or non-Hodgkin lymphoma, rhabdomyosarcoma, nonrhabdomyosarcoma soft-tissue sarcoma, one of the Ewing sarcoma family of tumors, or a newly diagnosed mass strongly suspected to be one of the aforementioned tumors; and willingness to undergo imaging studies specified in the protocol.

Participants with prior CT, conventional MR imaging, MIBG or gallium scintigraphy, or FDG PET studies that were performed before registration in the trial, either at the participating site or at an outside institution, were eligible for inclusion if these studies met the technical standards specified in the protocol. Participants were eligible if they had undergone protocol imaging within 2 months of a biopsy or surgical procedure performed at the participating site or at an outside institution if the diagnosis was verified by a pathologist.

Exclusion criteria.—Nontechnical exclusion criteria were contraindications for MR imaging (including cardiac pacemakers or intracranial vascular clips), nonprotocol tumors determined at pathologic examination; previous malignant tumors, pregnancy or nursing, and claustrophobia. Patients for whom there were protocol violations including technically inadequate MR imaging, greater than 21-day time frame for all imaging studies, greater than 2-month time frame between whole-body MR imaging and biopsy or surgical procedure, radiation or chemotherapy before imaging, and incomplete clinical data, were also excluded.

Study Protocol

All patients underwent whole-body MR imaging with fast turbo short tau inversion-recovery (STIR) and the required conventional staging examinations at initial presentation. A panel of pediatric oncologists determined the staging requirements on the basis of standard clinical practice at the time and existing clinical trials through the Children's Oncology Group for which patients may have been eligible. CT of the neck, chest, abdomen, and pelvis and either gallium scintigraphy or FDG PET were required for patients with lymphomas. Chest CT, CT or MR imaging of the primary site, and bone scintigraphy were required for patients with sarcomas. CT or MR imaging of the primary site, abdominal and pelvic CT or MR imaging, and bone scintigraphy or MIBG scintigraphy were required for patients with neuroblastomas. FDG PET was optional for patients with sarcomas, non-Hodgkin lymphomas, and neuroblastomas. Bone marrow aspirates and biopsies were also required for patients with sarcomas, neuroblastomas, non-Hodgkin lymphomas, and advanced-stage or symptomatic Hodgkin lymphomas.

All studies were required to be performed within 21 calendar days of each other, before treatment, and within 2 months of any diagnostic or surgical procedure. Details of the imaging protocols are given in Appendices E1 and E2 (online).

Central Interpretation of Studies (Reader Studies)

A reader study was used to achieve sufficient statistical power to enable establishment of noninferior diagnostic accuracy with whole-body MR imaging compared with conventional imaging for staging. For the whole-body MR imaging, a reference set showing examples of interpretation data was created to standardize interpretation quality among the study radiologists.

For the central image interpretation, there were 10 pairs of readers (Appendix E3 [online]), each consisting of an experienced pediatric radiologist and an experienced nuclear medicine physician (to mimic clinical practice). The median experience of the pediatric radiologists, seven of whom were from participating sites, was 11.5 years (range, 5–31 years). The median experience of the nuclear medicine physicians, five of whom were from participating sites, was 15.5 years (range, 4–38 years).

The central image interpretation for each reader pair occurred during two sessions. In one session, the pediatric radiologist interpreted the whole-body MR imaging studies, and in the other session the pair interpreted the conventional imaging tests (CT, conventional MR imaging, scintigraphic studies, and optional FDG PET). To minimize recall bias, a period of at least 1 month was required between interpretation of the whole-body MR imaging studies and those of conventional CT and/or MR imaging studies for the same case. The reader pairs were the same throughout the entire study. The readers were aware of the location and type of the primary tumor for each patient but were blinded to the reference standard information throughout all reading sessions.

Pediatric radiologists interpreted the whole-body MR imaging and conventional imaging studies in separate sessions; their readings were used to determine the accuracy with whole-body MR imaging. For the interpretations of the conventional imaging studies, the pediatric radiologist first independently interpreted the CT and conventional MR imaging studies, while the nuclear medicine physician independently interpreted the scintigraphic and PET studies of the same set of patients. After the independent interpretations, a paired reading of the conventional imaging tests (ie, excluding whole-body MR imaging) was done, and this joint interpretation was used to determine accuracy with conventional imaging.

All imaging tests were assessed for distant tumor extent, including metastases to the skeleton, nonskeletal sites, and nonregional lymph nodes. Distant disease was considered to be disease beyond what was thought to be the primary site(s) or regional nodal sites of disease. However, in lymphoma, only extranodal and extrasplenic sites of disease were considered to be distant disease.

The diagnosis of skeletal (either bone or bone marrow) metastases with conventional T2-weighted and STIR sequence MR imaging was based on demonstration of one or more areas of diffuse or focal high signal intensity compared with adjacent muscle and bone marrow. The diagnosis of bone or bone marrow metastases on scintigraphy (bone, MIBG, or gallium) or FDG PET was based on demonstration of one or more focal areas of increased activity

on the images. Skeletal metastases included those to the skull, cervical spine, thoracic spine, lumbar spine, sternum, scapulae, clavicles, ribs, pelvis, humeri, radii, ulnae, hands, femora, tibiae, fibulae, and feet.

Nonskeletal sites of distant metastatic disease included the common sites of metastases (lung, liver, spleen) and other parenchymal sites (brain, head and/or neck, soft tissues of upper and lower extremities, chest wall, pleura, abdominal wall, pancreas, adrenal glands, kidneys, bowel, and/or mesentery system). In patients with lymphoma, splenic involvement was considered local disease. Lung metastases were defined as foci of high signal intensity on MR images and as soft-tissue attenuating lesions at CT. Other nonskeletal metastases were defined as foci of high signal intensity on MR images and low or high attenuation at CT. Diagnosis of nonskeletal metastases at scintigraphy or PET involved the identification of greater than normal uptake of radiopharmaceuticals.

In children younger than 10 years, any visualized intrathoracic or intraabdominal lymph node was considered abnormal, regardless of size. In older children, a lymph node was considered abnormal on CT, conventional MR, or whole-body MR images if the largest transverse diameter exceeded 1 cm. The diagnosis of lymph node involvement of tumor at scintigraphy or PET was based on the identification of uptake that was greater than normal activity. Nodal sites evaluated included cervical, hilar, supraclavicular, axillary, mediastinal, abdominal, pelvic, the upper and lower extremities, and other sites.

The readers recorded their findings on standardized case report forms and included the location of metastases, lesion size, and degree of suspicion of distant metastatic disease. These findings were recorded separately for each imaging examination. For conventional imaging, each reader pair provided joint assessment of the degree of suspicion of distant metastatic disease on a five-point ordinal categorical scale (1 = definitely not present; 2 = probably not present; 3 = indeterminate; 4 = probably present; and 5 = definitely present).

Reference Standard Information

The final diagnosis of metastatic disease was established by an independent diagnosis verification panel, which included four pediatric oncologists and one diagnostic radiologist (Appendix E3 [online]). The chairperson of the panel was a pediatric oncologist (A.M.F., with 15 years of experience) and the panel included three other pediatric oncologists with expertise in each of the major tumor groups included in the study and with at least 10 years of clinical experience. The panel reviewed all available clinical history and initial and 3-month and/or 6-month follow-up images (CT, conventional MR imaging, bone, MIBG or gallium scintigraphy, or FDG PET) and other diagnostic tests (bone marrow aspiration or biopsy, surgical biopsy of skeletal or visceral lesions) and surgical reports. The whole-body MR images were not used to determine proof of diagnosis. The panel used their best clinical judgment and all available information to arrive at a consensus about the reference standard for each case.

Lesions suspected of being metastases on the basis of the imaging examinations were diagnosed as metastases by the diagnosis verification panel either when the imaging findings were sufficiently convincing or when tissue was sampled and proved to be metastatic disease. The results of follow-up imaging studies were also reviewed in patients with equivocal images and with lesions diagnosed as metastases when the lesions became larger during the follow-up period or smaller after treatment.

A post hoc review of patients who had metastasis was performed (M.J.S., F.A.H., and J.B.W., with 31, 21, and 25 years of experience, respectively) to further categorize sites of disease into skeletal, lung, liver, distant node, and other sites. This information was only used for secondary analyses.

Statistical Analysis

The primary goal of the study was to establish noninferior diagnostic accuracy with whole-body MR imaging compared with conventional imaging for the detection of distant metastases in the staging of common pediatric tumors. The primary endpoint was diagnostic accuracy with whole-body MR imaging and conventional imaging, as quantified by the area under the receiver operating characteristic curve (AUC) obtained in the reader study, averaged among readers (whole-body MR imaging) or reader pairs (conventional imaging).

For the primary analysis, the comparison of average AUC between the experimental (whole-body MR imaging) and control (conventional imaging) modalities was done by using the mixed effects model proposed by Obuchowski and Rockette (25,26). To establish noninferior diagnostic accuracy of whole-body MR imaging, the lower bound of the 95% confidence interval (CI) for the difference in average AUC (AUC for whole-body MR imaging minus AUC for conventional imaging) was required to to exceed −0.03. The study was designed to ensure sufficient enrollment so that the lower bound of the 95% CI was greater than −0.03 when 10 reader pairs read the images of 64 participants (32 positive and 32 negative for metastasis). This provided more than 82% power for testing the primary noninferiority hypothesis when the true AUCs for whole-body MR imaging and conventional imaging were 0.85 and 0.80, respectively. The noninferiority margin of 0.03 implied that a difference in AUC of 3.75% (or less) for whole-body MR imaging compared with conventional imaging would not be considered to be clinically important. The rate of enrolled patients who were positive for metastasis turned out to be higher than that initially assumed, and as a result, the diagnosis verification panel determined that more than 32 patients were positive for metastasis. Primary endpoint analysis was ultimately done for a set of 66 patients (33 positive and 33 negative for metastasis) selected from among the 134 patients who met inclusion and exclusion criteria and had available reference standard information. To allow for initiation of the reader study before

the diagnosis panel completed patient assessment, patients were selected in two stages (Appendix E2 [online]).

Patients in whom metastasis was most commonly missed at whole-body MR imaging were defined as those for whom a majority (≥ five readers) of whole-body MR imaging readers recorded an incorrect diagnosis. These patients were further examined at image re-review by two radiologists (M.J.S. and F.A.H.). On the 5-point ordinal scale for suspicion of distant metastatic disease, patients who were understaged (ie, received a false negative diagnosis) received diagnoses of 3 = indeterminate, 2 = probably not present, or 1 = definitely not present. Patients who were overstaged (ie, received a false positive diagnosis) received diagnoses of 4 = probably present or 5 = definitely present. In particular, location and size of missed disease were reported for patients who were understaged.

Secondary analyses are described in Appendix E2 (online). Statistical computations were performed by using SAS software v9.2 (SAS Institute, Cary, NC), S-plus software v8.0 (Insightful Corporation, Seattle, Wash), ROCKIT software (University of Chicago, Chicago, Ill) (23,24), and OBUMRM software (Cleveland Clinic, Cleveland, Ohio) (25,26). For all reported analyses, *P* values less than .05 were considered to indicate a statistically significant difference.

Figure 1

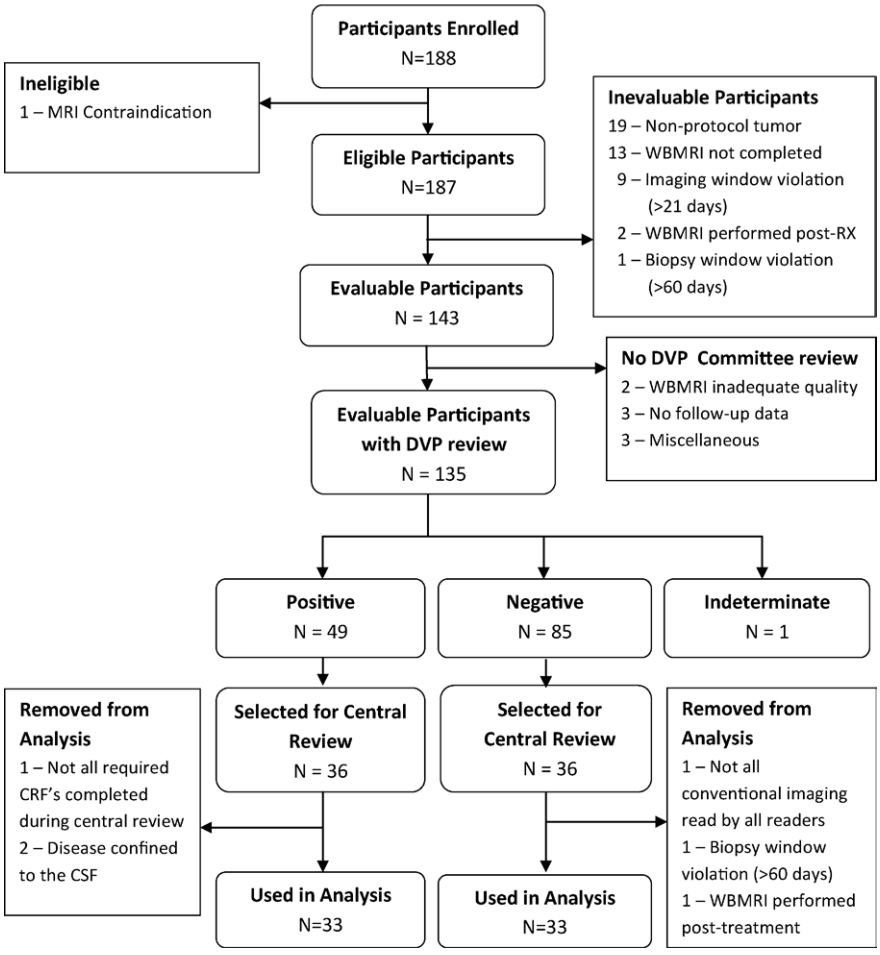


Figure 1: Study flowchart. *CRF* = case report form, *CSF* = cerebrospinal fluid, *DVP* = diagnosis verification panel, *WBMRI* = whole-body MR imaging.

Results

Study Cohort

A total of 188 patients from 20 sites (Appendix E3 [online]) were enrolled and 54 patients were excluded (Fig 1). Among the 134 remaining patients for whom reference standard information was obtained, there were 49 patients who were positive and 85 patients who were negative for metastasis, from which 33 patients with positive and 33 patients with negative diagnoses for metastasis were included in the central reader study.

The patients in the eligible (n = 187) and analysis (n = 66) sets had comparable sociodemographics, medical history, and biopsy verification (Table 1). The mean ± standard deviation difference in imaging time between whole-body MR imaging and conventional studies was 9.5 days ± 9.5 for the eligible patients and 7.2 days ± 4.7 for the analyzed patients. The mean imaging time between whole-body MR imaging and bone marrow aspiration, biopsy, or surgical biopsy was 8.1 days ± 15.9 for the eligible patients and 6.6 days ± 9.7 for the analyzed patients.

Final Pathologic Examination

Other than a higher percentage of neuroblastomas in the analysis set compared with that in the eligible set (19 of 66 [29%] vs 34 of 187 [18%]), the final pathologic examination results did not differ substantially between the two groups (Table 1). Conventional imaging according to tumor subset for patients in the analysis set is shown in Table 2.

Primary Analysis: Location and Size of Distant Disease

On the basis of the final diagnosis, the distribution of disease in the evaluable and analysis sets was similar. In the analysis set, five of 33 (15%) had only skeletal metastases, 14 of 33 (42%) had only nonskeletal disease, and 14 of 33 (42%) had both skeletal and

Table 1

Characteristics of Eligible and Analysis Sets in the Study

Participant Characteristic	Eligible Set (n = 187)	Analysis Set (n = 66)	Eligible, Not Analyzed Set (n = 121)	P Value*
Age (y)				.0355
Mean ± standard deviation	10.2 ± 6.1	9.0 ± 6.4	10.9 ± 5.8	
Range	<1–21	<1–21	<1–20	
Sex				.5118
Male	108 (58)	36 (54)	72 (60)	
Female	79 (42)	30 (45)	49 (40)	
No. of biopsy procedures				.9281
Mean ± standard deviation	2.7 ± 1.6	2.7 ± 1.1	2.7 ± 1.8	
Range	0–10	0–7	0–10	
Length of biopsy window (d)†				.2916
Mean ± standard deviation	8.1 ± 15.9	6.6 ± 9.7	9.2 ± 19.2	
Range	0–168	0–58	0–168	
Length of imaging window (d)‡				.0601
Mean ± standard deviation	9.5 ± 9.5	7.2 ± 4.7	9.4 ± 10.4	
Range	0–64	0–18	0–64	
Final Tumor Type				.2184
No pathology report	4 (2)	0 (0)	4 (3)	
Nonprotocol tumor	19 (10)	0 (0)	19 (16)	
Rhabdomyosarcoma	21 (11)	9 (14)	12 (10)	
Ewing sarcoma	26 (14)	7 (11)	19 (16)	
Neuroblastoma	34 (18)	19 (29)	15 (12)	
Non-Hodgkin lymphoma	25 (13)	7 (11)	18 (15)	
Hodgkin lymphoma	53 (28)	22 (33)	31 (26)	
Other	5 (3)	2 (3)	3 (2)	

Note.—Unless otherwise indicated, data are number of patients with percentages in parentheses.

* P values are for comparison of analysis set to those patients eligible but not analyzed.

† One patient with a recorded interval from biopsy to whole-body MR imaging outside of protocol limits was excluded.

‡ One patient with a CT scan reportedly obtained before consent (possibly incorrectly recorded) was excluded.

nonskeletal metastases. Among the patients in the analysis set, seven of 33 (21%) had liver involvement, 16 of 33 (48%) had lung involvement, 11 of 33 (33%) had distant node involvement, and 15 of 33 (45%) had other involvement (Table 3).

Primary Analysis: Tumor Subset versus Location of Distant Disease

For analysis of the results according to tumor type, the data were divided into two groups, those with Hodgkin or non-Hodgkin lymphoma (n = 29) and those with all other tumor types (n = 37). For the lymphoma group, the diagnosis verification panel determined that distant metastatic disease was present in 11 of 29 (38%) and not present in 18 of 29 (62%) patients. For other tumors, distant metastatic disease was determined to be present in 22 of 37 (59%) and not present in 15 of 37 (41%) patients.

Among the 33 analyzed patients who were positive for metastasis, skeletal metastases were present in six of 10 (60%) patients with neuroblastomas, eight of 12 (67%) with sarcomas, and five of 11 (45%) with lymphomas (Table 3). Nonskeletal metastases were present in the lung, liver, and distant nodes, respectively, in 0 of 10 (0%), five of 10 (50%), and four of 10 (40%) neuroblastomas, and in 10 of 12 (83%), 0 of 12 (0%) and seven of 12 (58%) sarcomas. Lung and liver metastases were present in six of 11 (55%) and two of 11 (18%) lymphomas.

Primary Analysis: Accuracy in Detection of Distant Disease

The difference between the average AUC for the experimental and conventional examinations was −0.03 (95% CI: –0.10, 0.04; Table 4). Because the lower bound of the 95% CI was not greater than −0.03, noninferiority of whole-body MR imaging compared with conventional imaging could not be established. This implies that the accuracy of whole-body MR imaging could not be shown to be close enough to that of conventional imaging to preclude a clinically meaningful difference in performance. In addition, a sensitivity analysis was performed in which the four patients whose images were read by the central readers but were subsequently removed from the analysis set (Fig 1) were included in the primary comparison. The results were similar (difference in average AUC, −0.04; 95% CI: −0.11, 0.03) and the overall conclusion remained the same.

Examples of whole-body MR imaging detection of distant metastatic disease in patients with neuroblastoma are shown in Figure 2.

Primary Analysis: Errors in Diagnosis of Distant Disease

Among the 33 patients who were negative for metastasis, there were no instances in which a majority of whole-body MR imaging readers recorded a false-positive diagnosis. Disease was understaged (false-negative diagnoses) at whole-body MR imaging by a majority of readers in eight (24%) of the 33 patients with metastasis in the analysis set. These patients included two with neuroblastoma, one with sarcoma, and five with lymphoma. On a per-patient basis, distant disease was missed at whole-body MR imaging in the lung in four of eight patients, in the liver in two of eight patients, and in distant node (axillary) and tonsil in one of eight patients each. In all of the patients in whom metastasis was missed except for the patients with metastasis in the tonsil axillary node, lesions were approximately 1 cm or less in diameter. There were no false-negative diagnoses based on missed skeletal lesions at whole-body MR imaging.

Among the remaining correctly diagnosed 25 patients who were positive for metastasis, there were two patients

Table 2

Conventional Imaging Included in the Central Reader Study for Patients in the Analysis Set

Tumor Type	CT	MR Imaging	Bone Scintigraphy	MIBG Scintigraphy	Gallium Scintigraphy	PET	Total No. of Patients
Rhabdomyosarcoma	9 (100)	7 (78)	9 (100)	0 (0)	0 (0)	1 (11)	9
Ewing sarcoma	7 (100)	3 (43)	7 (100)	0 (0)	0 (0)	2 (29)	7
Neuroblastoma	18 (95)	8 (42)	19 (100)	12 (63)	0 (0)	0 (0)	19
Lymphoma	29 (100)	3 (10)	0 (0)	0 (0)	17 (59)	15 (52)	29
Other, specify	2 (100)	1 (50)	2 (100)	0 (0)	0 (0)	1 (50)	2
Total	65 (98)	22 (33)	37 (56)	12 (18)	17 (26)	19 (29)	66

Note.—Data are number of patients with percentages in parentheses.

Table 3

Tumor Subset versus Sites of Distant Disease among Positive Patients in the Analysis Set

Final Tumor Type	Skeletal	Lung	Liver	Distant Nodes	Other	Total
Sarcoma	8 (67)	10 (83)	0 (0)	7 (58)	6 (50)	12 (36)
Neuroblastoma	6 (60)	0 (0)	5 (50)	4 (40)	5 (50)	10 (30)
Lymphoma	5 (45)	6 (55)	2 (18)	NA	4 (36)	11 (33)
Total	19 (58)	16 (48)	7 (21)	11 (33)	15 (45)	33 (100)

Note.—Data are number of patients with percentages in parentheses. NA = not applicable.

(8%) in whom most of the readers identified a lesion of less than 1 cm in the lung, and no patients in whom a majority identified a lesion less than 1 cm in the liver.

Secondary Analysis: Tumor Type Subset Analysis

For whole-body MR imaging, the average AUC for solid tumors was significantly higher than that for lymphomas (Table 4; $P = .006$). In particular, both sensitivity (0.86 vs 0.56; $P < .001$) and specificity (0.91 vs 0.84; $P = .04$) were significantly higher for solid tumors. For conventional imaging, although the average AUC for solid tumors was higher than that for lymphomas, the difference was not statistically significant ($P = .19$). Of note, the sensitivity was also significantly higher for conventional imaging in the group with solid tumors compared with that in the group with lymphomas (0.89 vs 0.68; $P < .001$), but there was no significant difference in the specificity (0.84 vs 0.87; $P = .55$).

Secondary Analysis: Age Subset Analysis

To determine the effect of patient age on diagnostic accuracy with whole-body MR imaging, the analysis set was stratified into two age groups, those with an age at enrollment of 2 years or younger ($n = 16$), and those older than 2 years ($n = 50$). The small number of patients younger than 2 years precluded comparison of AUCs. The average sensitivity with whole-body MR imaging for children 2 years or younger was 0.77 (95% CI: 0.64, 0.86) and the average specificity for the same group was 0.92 (95% CI: 0.83, 0.96). The average sensitivity with whole-body MR imaging for children aged older than 2 years was 0.76 (95% CI: 0.70, 0.82) and the average specificity was 0.85 (95% CI: 0.77, 0.91). There was no statistically significant difference in the average sensitivity ($P = .95$) or specificity ($P = .11$) with whole-body MR imaging between these two groups.

Table 4

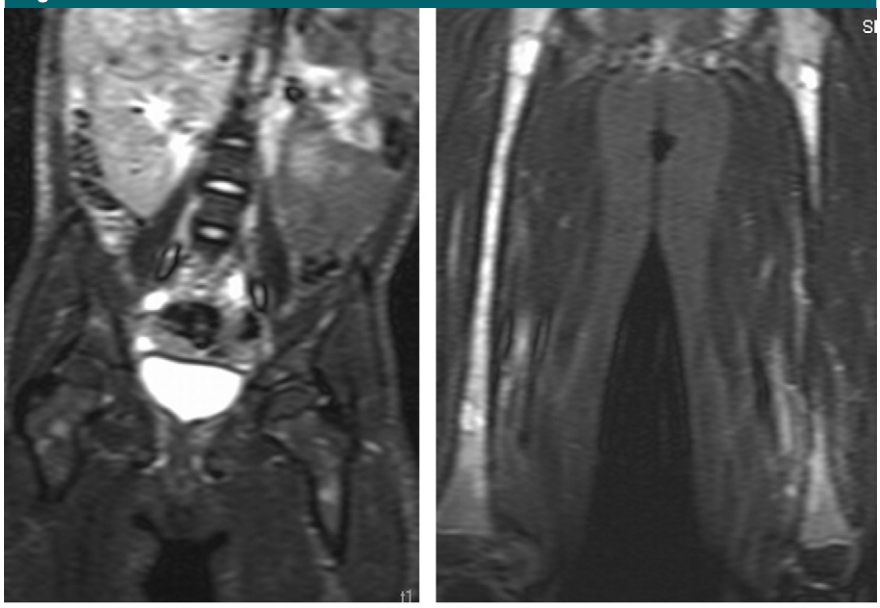
Subset Analysis: Lymphoma versus Nonlymphoma Patients

Modality and Tumor Type	No. of Tumors	Average AUC	95% CI
Whole-body MR imaging			
Lymphoma tumors	29	0.72	0.57, 0.87
Nonlymphoma tumors	37	0.92	0.85, 0.99
Overall	66	0.85	0.77, 0.92
Conventional imaging			
Lymphoma tumors	29	0.83	0.68, 0.97
Nonlymphoma tumors	37	0.91	0.83, 0.99
Overall	66	0.87	0.79, 0.96

Secondary Analysis: Skeletal Metastatic Disease Subset Analysis

The total number of patients who had disease in skeletal regions as determined by most of the readers of whole-body MR images and by most of thereaders of conventional images, regardless of histopathologic results (which were not available for every lesion), is shown in Table 5. A per-case analysis among

Figure 2



a. b.

Figure 2: Coronal whole-body MR images show metastatic neuroblastomas. **(a)** Image in a 4-year-old girl with right adrenal primary tumor shows multiple foci of increased signal intensity in iliac bones and proximal femora. Skeletal and iodine 123 MIBG scintigraphy (not shown) results were negative in this patient. **(b)** Image in a 3-year-old boy with left adrenal primary tumor shows diffusely increased signal intensity in both femora. Skeletal and iodine 123 MIBG scintigraphy results (not shown) were positive in this patient.

the 19 of 33 patients who were positive for distant skeletal metastases (Table 3) showed that, on average, more anatomic regions of skeletal disease were detected with whole-body MR imaging than with conventional imaging (5.9 vs 3.7; P = .03). The conventional imaging that was used in this comparison included scintigraphy in 17 patients and PET in two patients.

Secondary Analysis: Pulmonary Metastatic Disease Subset Analysis

The total number of nonskeletal regions with disease determined by most of the readers of whole-body MR and conventional images, regardless of histopathologic results (which were not available for every lesion), is shown in Table 5. In the 16 of 33 patients who were positive for distant pulmonary metastases (Table 3), whole-body MR imaging was found to be less accurate than conventional imaging for detection of pulmonary lesions. Most of the readers (≥ five readers) identified pulmonary metastases on conventional images of 13 of 16 (81%) patients versus eight of 16 (50%) on whole-body MR images. The average sensitivity of the 10 readers was 0.83 (0.75, 0.89) with conventional imaging and 0.53 (0.39, 0.67) with whole-body MR imaging ($P < .001$).

Discussion

The results of several studies have shown that whole-body MR imaging allows for more accurate detection of skeletal and extraskeletal metastases in children than does conventional oncologic imaging (16–22,27,28). These studies were limited by inclusion of patients from a single site, small sample sizes, inclusion of patients with prior treatment, and nonuniformity of the experimental or reference standard imaging. Our study differed in that it was a large, prospective multi-institutional study that used uniform techniques to compare the performance of whole-body MR imaging and conventional imaging for staging of common tumors in pediatric patients. In our study, noninferior diagnostic accuracy with whole-body MR imaging compared with conventional imaging for detection of distant metastatic disease was not established. However, whole-body MR imaging was significantly more accurate for the detection of distant metastatic deposits in solid tumors than in the detection of lymphomas.

When only skeletal metastases were evaluated, whole-body MR imaging allowed detection of more skeletal metastases than did conventional imaging. Among the 19 patients with distant skeletal disease, whole-body MR imaging allowed detection of substantially more regions of disease per patient than did conventional imaging (5.9 vs 3.7, P = .03). On a per-patient basis, there were no instances in which most readers recorded an incorrect diagnosis because of missed skeletal lesions. These results are in agreement with prior studies that have shown that whole-body MR imaging with STIR was superior to bone scintigraphy for detection of skeletal metastases (16–22).

The role of whole-body MR imaging in the detection of extraskeletal disease is more controversial. In our study, there were eight patients with positive findings at conventional imaging whom most of the readers found to be negative for metastasis on the basis of whole-body MR imaging. Six of eight missed lesions were approximately 1 cm or less in diameter, and all were located in the lung or liver. Among the patients with correctly staged tumors who were positive for metastasis, there were only two patients (8%) in whom lesions were less than 1 cm. In particular, our results show lower sensitivity for whole-body MR imaging compared with that for conventional imaging for detection of pulmonary metastases (0.83 vs 0.53; $P < .001$).

The understaging of pulmonary and hepatic disease on the basis of whole-body MR imaging is not surprising and has been shown by other investigators. Whole-body MR imaging can help detection of relatively large pulmonary

nodules, but smaller lesions (< 10 mm in diameter) may go undetected (15,29). Similarly, prior reports in patients with lymphoma have shown that sensitivity with whole-body MR imaging is 92%–100% for detecting nodes between 10 and 12 mm in diameter, versus a sensitivity of 67% for nodes between 6 and 12 mm in diameter and 11% for nodes between 1 and 6 mm in diameter (28,30).

The advantage of this study was that all participating institutions used a state-of-the-art uniform STIR technique. Recruitment patterns were similar and there was no statistically significant difference in the pathologic results between the eligible patient set and the analysis set. These facts suggest that the results are likely to be reproducible and generalizable to other centers.

One weakness of this study was that we did not evaluate other sequences, such as diffusion-weighted or fast-acquisition parallel-imaging technology such as sensitivity encoding or simultaneous acquisition of harmonies, or 3-T imaging systems, which have the potential to improve accuracy for lesion detection (31,32). However, these were not available when we started this study. An additional weakness is that FDG PET and PET/CT imaging were not considered in the comparison of whole-body MR imaging with other staging procedures. FDG PET was not widely used in children when this study was proposed and, therefore, was not included in the protocol. Future studies are needed to determine the relative roles of MR imaging and PET/CT in pediatric oncologic patients.

Another limitation of our study was that histopathologic proof, other than aspiration and biopsy of marrow in the iliac crest, was not available for most distant metastatic lesions seen on whole-body MR images. However, this weakness in design is one confronted in most studies that evaluate detection of metastases at imaging. Although biopsy is the most reliable method of establishing the diagnosis of metastatic disease, confirmation of every metastatic lesion will not change therapeutic strategies and is not possible for ethical reasons.

Table 5

Total Metastases in 33 Patients by a Majority of Readers

Location of Tumor Skeletal Regions	Whole-body MR Imaging	Conventional Imaging
Skeletal regions		
Skull	2 (6)	6 (18)
Cervical spine	1 (3)	1 (3)
Humerus (right or left)	11 (33)	6 (18)
Radius and/or ulna (right or left)	1 (3)	0 (0)
Hand (right or left)	0 (0)	0 (0)
Ribs (right or left)	3 (9)	7 (21)
Scapula and/or clavicle	3 (9)	1 (3)
Sternum	1 (3)	4 (12)
T-spine	11 (33)	9 (27)
L-spine	10 (30)	10 (30)
Pelvis	13 (39)	10 (30)
Femur (right or left)	16 (48)	8 (24)
Tibia and/or fibula (right or left)	14 (42)	5 (15)
Foot (right or left)	0 (0)	0 (0)
Nonskeletal regions		
Brain	0 (0)	0 (0)
Head and/or neck	0 (0)	0 (0)
Soft tissue, upper extremity (right or left)	0 (0)	0 (0)
Lungs (right or left)	9 (27)	14 (42)
Pleura	2 (6)	3 (9)
Chest wall	0 (0)	0 (0)
Liver	4 (12)	4 (12)
Spleen	2 (6)	3 (9)
Pancreas	0 (0)	1 (3)
Adrenal gland (right or left)	0 (0)	0 (0)
Kidney (right or left)	0 (0)	2 (6)
Bowel, mesentery, and/or peritoneum	0 (0)	0 (0)
Abdominal wall	0 (0)	0 (0)
Pelvis	1 (3)	0 (0)
Soft tissue, lower extremity (right or left)	1 (3)	0 (0)

Note.—Data are number of patients with percentage in parentheses.

The next best method for the establishment of the nature of the lesions detected in an imaging study is to compare the findings with those of another study or to re-evaluate the patient during the course of treatment, which is the method we used in this study.

Finally, the study design was a potential limitation. Rather than conducting the usual superiority study (with a null hypothesis of equality of the AUCs), we made a choice to test a noninferiority hypothesis. In part, this was based on the practical consideration that the required sample size would be unachievable. In addition, although there was an a priori belief among the study team that whole-body MR imaging could be as good as conventional imaging (from a clinical perspective); there was no strong rationale for assuming that it could actually be shown to be substantially better, especially in a controlled, multireader study. Furthermore, from a medical perspective, given the fact that whole-body MR imaging involves no ionizing radiation (an important consideration in a pediatric population), we felt that establishing noninferior diagnostic accuracy in detection of distant disease would be sufficient to warrant its clinical use.

In conclusion, because this trial did not establish that the diagnostic

accuracy of whole-body MR imaging was noninferior to that of conventional imaging, whole-body MR imaging cannot be justified as the study of choice for detection of distant disease in pediatric patients with common malignant tumors, especially those tumors that typically spread to lung and solid organs.

Acknowledgments: The authors gratefully acknowledge the participating institutions (and the principal investigator at each site), the radiologists who participated in the central reader study, and the members of the diagnosis verification panel (Appendix E3 [online]). We also thank the many radiologists, oncologists, surgeons, radiologic technologists, and research coordinators at the participating sites, and the ACRIN staff who supported the ACRIN 6660 trial at both ACRIN headquarters, in Philadelphia, Penn and the Biostatistic and Data Management Center at Brown University in Providence, RI. Without the diligent efforts of these many individuals, this study would not have been possible.

Disclosures of Conflicts of Interest: M.J.S. No relevant conflicts of interest to disclose. **S.A.** No relevant conflicts of interest to disclose. **F.A.H.** Financial activities related to the present article: author received consulting fee or honorarium, support for travel to meetings or other purposes, and fees for participation in review activities from the ACRIN; institution received grant and writing assistance, medicines, equipment, or administrative support from the ACRIN. Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **J.B.W.** No relevant conflicts of interest to disclose. **A.M.F.** Financial activities related to the present article: received reimbursement for travel and time for committee work. Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **B.S.S.** No relevant conflicts of interest to disclose. **G.K.** Financial activities related to the present article: institution received support from ACRIN. Financial activities not related to the present article: none to disclose. Other relationships: none to disclose. **B.A.S.** Financial activities related to the present article: received payment from ACRIN for work as a reader and as Deputy cochair and received travel expense reimbursement from ACRIN. Financial activities not related to the present article: none to disclose. Other relationships: none to disclose.

References

1. Breneman JC, Lyden E, Pappo AS, et al. Prognostic factors and clinical outcomes in children and adolescents with metastatic rhabdomyosarcoma—a report from the Intergroup Rhabdomyosarcoma Study IV. J Clin Oncol 2003;21(1):78–84.
2. Pizzo PA, Poplack DG. Neuroblastoma. In: Principles and practice of pediatric oncology. Philadelphia, Pa: Lippincott Williams & Wilkins, 2011; 886–922.
3. Nachman JB, Sposto R, Herzog P, et al; Children's Cancer Group. Randomized comparison of low-dose involved-field radiotherapy and no radiotherapy for children with Hodgkin's disease who achieve a complete response to chemotherapy. J Clin Oncol 2002;20(18):3765–3771.
4. Pizzo PA, Poplack DG. Ewing sarcoma. In: Principles and practice of pediatric oncology. Philadelphia, Pa: Lippincott Williams & Wilkins, 2011; 987–1014.
5. Brenner DJ, Hall EJ. Computed tomography—an increasing source of radiation exposure. N Engl J Med 2007;357(22):2277–2284.
6. Ahmed BA, Connolly BL, Shroff P, et al. Cumulative effective doses from radiologic procedures for pediatric oncology patients. Pediatrics 2010;126(4):e851–e858.
7. Horner MJ, Ries LAG, Krapcho M, et al, eds. SEER Cancer Statistics Review, 1975–2006. Bethesda, Md: National Cancer Institute, 2008. http://seer.cancer.gov/csr/1975_2006. Accessed August 10, 2009.
8. Lauenstein TC, Freudenberg LS, Goehde SC, et al. Whole-body MRI using a rolling table platform for the detection of bone metastases. Eur Radiol 2002;12(8):2091–2099.
9. Eustace S, Tello R, DeCarvalho V, Carey J, Melhem E, Yucel EK. Whole body turbo STIR MRI in unknown primary tumor detection. J Magn Reson Imaging 1998;8(3):751–753.
10. Walker R, Kessar P, Blanchard R, et al. Turbo STIR magnetic resonance imaging as a whole-body screening tool for metastases in patients with breast carcinoma: preliminary clinical experience. J Magn Reson Imaging 2000;11(4):343–350.
11. Lauenstein TC, Goehde SC, Herborn CU, et al. Three-dimensional volumetric interpolated breath-hold MR imaging for whole-body tumor staging in less than 15 minutes: a feasibility study. AJR Am J Roentgenol 2002;179(2):445–449.
12. Engelhard K, Hollenbach HP, Wohlfart K, von Imhoff E, Fellner FA. Comparison of whole-body MRI with automatic moving table technique and bone scintigraphy for screening for bone metastases in patients with breast cancer. Eur Radiol 2004;14(1):99–105.
13. Hargaden G, O'Connell M, Kavanagh E, Powell T, Ward R, Eustace S. Current concepts in whole-body imaging using turbo short tau inversion recovery MR imaging. AJR Am J Roentgenol 2003;180(1):247–252.
14. Iizuka-Mikami M, Nagai K, Yoshida K, et al. Detection of bone marrow and extramedullary involvement in patients with non-Hodgkin's lymphoma by whole-body MRI: comparison with bone and ^{67}Ga scintigraphies. Eur Radiol 2004;14(6):1074–1081.
15. Lauenstein TC, Goehde SC, Herborn CU, et al. Whole-body MR imaging: evaluation of patients for metastases. Radiology 2004;233(1):139–148.
16. Daldrup-Link HE, Franzius C, Link TM, et al. Whole-body MR imaging for detection of bone metastases in children and young adults: comparison with skeletal scintigraphy and FDG PET. AJR Am J Roentgenol 2001;177(1):229–236.
17. Goo HW, Choi SH, Ghim T, Moon HN, Seo JJ. Whole-body MRI of paediatric malignant tumours: comparison with conventional oncological imaging methods. Pediatr Radiol 2005;35(8):766–773.
18. Goo HW, Yang DH, Ra YS, et al. Whole-body MRI of Langerhans cell histiocytosis: comparison with radiography and bone scintigraphy. Pediatr Radiol 2006;36(10):1019–1031.
19. Kumar J, Seith A, Kumar A, et al. Whole-body MR imaging with the use of parallel imaging for detection of skeletal metastases in pediatric patients with small-cell neoplasms: comparison with skeletal scintigraphy and FDG PET/CT. Pediatr Radiol 2008;38(9):953–962.
20. Laffan EE, O'Connor R, Ryan SP, Donoghue VB. Whole-body magnetic resonance imaging: a useful additional sequence in paediatric imaging. Pediatr Radiol 2004;34(6):472–480.
21. Mazumdar A, Siegel MJ, Narra V, Luchtman-Jones L. Whole-body fast inversion recovery MR imaging of small cell neoplasms in pediatric patients: a pilot study. AJR Am J Roentgenol 2002;179(5):1261–1266.
22. Mentzel HJ, Kentouche K, Sauner D, et al. Comparison of whole-body STIR-MRI and ^{99m}Tc-methylene-diphosphonate scintigraphy in children with suspected multifocal bone lesions. Eur Radiol 2004;14(12):2297–2302.
23. Metz CE, Wang P-L, Kronman HB. A new approach for testing the significance of differences between ROC curves measured from correlated data. In: Deconinck F, ed. Information processing in medical imaging.

The Hague, the Netherlands: Martinus Nijhoff, 1984; 432–445.

24. Metz CE, Herman BA, Roe CA. Statistical comparison of two ROC-curve estimates obtained from partially-paired datasets. Med Decis Making 1998;18(1):110–121.
25. Obuchowski NA, Rockette HE. Hypothesis testing of the diagnostic accuracy of multiple diagnostic tests: an ANOVA approach with dependent observations. Commun Stat Simul Comput 1995;24(2):285–308.
26. Obuchowski NA. Multireader, multimodality receiver operating characteristic curve studies: hypothesis testing and sample size estimation using an analysis of variance approach with dependent observations. Acad Radiol 1995;2(Suppl 1):S22–S29; discussion S57–S64, S70–S71 pas.
27. Kellenberger CJ, Miller SF, Khan M, Gilday DL, Weitzman S, Babyn PS. Initial experience with FSE STIR whole-body MR imaging for staging lymphoma in children. Eur Radiol 2004;14(10):1829–1841.
28. Punwani S, Taylor SA, Bainbridge A, et al. Pediatric and adolescent lymphoma: comparison of whole-body STIR half-Fourier RARE MR imaging with an enhanced PET/CT reference for initial staging. Radiology 2010;255(1):182–190.
29. Frericks BB, Meyer BC, Martus P, Wendt M, Wolf KJ, Wacker F. MRI of the thorax during whole-body MRI: evaluation of different MR sequences and comparison to thoracic multidetector computed tomography (MDCT). J Magn Reson Imaging 2008;27(3):538–545.
30. Brennan DD, Gleeson T, Coate LE, Cronin C, Carney D, Eustace SJ. A comparison of whole-body MRI and CT for the staging of lymphoma. AJR Am J Roentgenol 2005;185(3):711–716.
31. Padhani AR, Koh D-M, Collins DJ. Whole-body diffusion-weighted MR imaging in cancer: current status and research directions. Radiology 2011;261(3):700–718.
32. MacKenzie JD, Vasanawala SS. Advances in pediatric MR imaging. Magn Reson Imaging Clin N Am 2008;16(3):385–402, v.