Science to Practice:

Liane E. Philpotts

**Will Improved Vascular Mapping Achieved with Gadobenate Dimeglumine Aid in Interpretation of Breast MR Images?**

The findings of Sardanelli et al show not only that gadobenate dimeglumine can be used to obtain vascular maps that are superior to those obtained with gadopentetate dimeglumine but also that a difference in vascularity between the two breasts is an indirect sign of malignancy.

Radiology 2005 235: 717-718

Perspectives:

Robert T. Bramson and George A. Taylor

**SOS: Can We Save Pediatric Radiology?**

Do we want to expend the energy, try the experiments, suffer the failures, and find a way to avoid the tipping point that spells the demise of pediatric radiology?

Radiology 2005 235: 719-722

Special Reports:


**American College of Radiology Clinical Statement on Noninvasive Cardiac Imaging**

This clinical statement of the American College of Radiology (ACR) discusses various technical and patient safety issues related to cardiac CT and MR imaging, and it suggests appropriate qualifications for radiologists until such time as ACR practice guidelines for the performance of cardiac CT and cardiac MR imaging are written and approved through the usual ACR process.

Radiology 2005 235: 723-727. Published online before print April 21 2005

**Image-guided Tumor Ablation: Standardization of Terminology and Reporting Criteria**

The intent of this standardization of terminology is to provide an appropriate vehicle for reporting the various aspects of image-guided ablation therapy. Radiology 2005 235: 728-739. Published online before print April 21 2005,

Denise R. Aberle, Caroline Chiles, Constanine Gatsonis, Bruce J. Hillman, C. Daniel Johnson, Bruce L. McClennan, Donald G. Mitchell, Etta D. Pisano, Mitchell D. Schnall, and A. Gregory Sorensen

**Imaging and Cancer: Research Strategy of the American College of Radiology Imaging Network**

The individual disease site committee strategies make up the comprehensive American College of Radiology Imaging Network research agenda, which we believe also circumscribes the agenda for imaging and cancer. Radiology 2005 235: 741-751

**Editorials:**

Richard W. Katzberg

**Contrast Medium–induced Nephrotoxicity: Which Pathway?**

The current article by Heinrich et al in *Radiology* provides a renewed focus on the direct tubular cell toxicity pathway of contrast medium–induced acute renal failure and, in my opinion, a promising pathway to solve a growing clinical crisis. Radiology 2005 235: 752-755

**Breast Imaging:**

Edward A. Sickles, Diana L. Miglioretti, Rachel Ballard-Barbash, Berta M. Geller, Jessica W. T. Leung, Robert D. Rosenberg, Rebecca Smith-Bindman, and Bonnie C. Yankaskas

**Performance Benchmarks for Diagnostic Mammography**

We have presented a very extensive set of data on diagnostic mammography outcomes and performance benchmarks, among a patient population judged to be representative of the population examined in general radiology practice in the United States, with data designed to be used by mammography facilities and individual radiologists to evaluate their own performance for diagnostic mammography as determined by periodic comprehensive audits. Radiology 2005 235: 775-790
Francesco Sardanelli, Andrea Iozzelli, Alfonso Fausto, Alessandro Carriero, and Miles A. Kirchin

**Gadobenate Dimeglumine–enhanced MR Imaging Breast Vascular Maps: Association between Invasive Cancer and Ipsilateral Increased Vascularity**

Our results suggest that gadobenate dimeglumine, as compared with gadopentetate dimeglumine, may have preferential properties for MR imaging evaluation of breast vascularity and that one-sided increased breast vascularity is frequently associated with ipsilateral invasive breast cancer.


---

**Cardiac Imaging:**

Rutger W. van der Meer, Peter M. T. Pattynama, Marco J. L. van Strijen, Annette A. van den Berg-Huijsmans, Ieneke J. C. Hartmann, Hein Putter, Albert de Roos, and Menno V. Huisman

**Right Ventricular Dysfunction and Pulmonary Obstruction Index at Helical CT: Prediction of Clinical Outcome during 3-month Follow-up in Patients with Acute Pulmonary Embolism**

Both the right ventricle-to-left ventricle short-axis diameters ratio and the pulmonary vascular obstruction index as assessed at helical CT are potentially useful tools to predict mortality in patients with initially hemodynamically stable pulmonary embolism at clinical presentation.


Yasushi Koyama, Hiroshi Matsuoka, Teruhito Mochizuki, Hiroshi Higashino, Hideo Kawakami, Shigeru Nakata, Jun Aono, Taketoshi Ito, Makiko Naka, Yasuo Ohashi, and Jitsuo Higaki

**Assessment of Reperfused Acute Myocardial Infarction with Two-Phase Contrast-enhanced Helical CT: Prediction of Left Ventricular Function and Wall Thickness**

In patients with acute myocardial infarction, the myocardial enhancement pattern at two-phase contrast-enhanced CT performed after reperfusion therapy can serve as a predictor of left ventricular functional recovery and wall thickness.


Jaydip Datta, Charles S. White, Robert C. Gilkeson, Cristopher A. Meyer, Sarita Kansal, Manish L. Jani, Ronald C. Arildsen, and Katrina Read

**Anomalous Coronary Arteries in Adults: Depiction at Multi–Detector Row CT Angiography**

Multi–detector row CT angiography provided accurate depiction of vessel origin and course in this review of 20 anomalous coronary arteries.

**Contrast Media:**

Martine Remy-Jardin, Philippe Dequiedt, Olivier Ertzbischoff, Isabelle Tillie-Leblond, John Bruzzi, Alain Duhamel, and Jacques Remy

*Safety and Effectiveness of Gadolinium-enhanced Multi–Detector Row Spiral CT Angiography of the Chest: Preliminary Results in 37 Patients with Contraindications to Iodinated Contrast Agents*

Gadolinium-based contrast agents can provide diagnostic CT angiograms of the pulmonary circulation, but high-quality examinations require the use of 16–detector row CT technology.


**Diagnosis Please:**

Michelle J. Naidich and Sam U. Ho

*Case 87*

Radiology 2005 235: 827-828

Gilberto Szarf and David A. Bluemke

*Case 83: Multifocal Fibrosclerosis with Mediastinal-Retroperitoneal Involvement*

Multifocal fibrosclerosis is a rare syndrome of unknown cause that is characterized by fibrosis involving multiple organ systems.

Radiology 2005 235: 829-832

**Evidence-based Practice:**

Sandra Spronk, Johanna L. Bosch, Hermanus F. Veen, Pieter T. den Hoed, and M. G. Myriam Hunink

*Intermittent Claudication: Functional Capacity and Quality of Life after Exercise Training or Percutaneous Transluminal Angioplasty—Systematic Review*

There is still no evidence that quality of life for patients with intermittent claudication is significantly better after percutaneous transluminal angioplasty than after exercise training after more than 6 months.

Experimental Studies:

Marc C. Heinrich, Martin K. Kuhlmann, Aleksandar Grgic, Martina Heckmann, Bernhard Kramann, and Michael Uder

**Cytotoxic Effects of Ionic High-osmolar, Nonionic Monomeric, and Nonionic Iso-osmolar Dimeric Iodinated Contrast Media on Renal Tubular Cells in Vitro**
The iso-osmolar dimeric contrast media showed strong cytotoxic effects, which were significantly \((P < .001)\) more pronounced when they were compared with those of the low-osmolar monomeric contrast media at equimolar concentrations. Radiology 2005 235: 843-849.

Jin Mo Goo, Trongtum Tongdee, Ranista Tongdee, Kwangjae Yeo, Charles F. Hildebolt, and Kyongtae T. Bae

**Volumetric Measurement of Synthetic Lung Nodules with Multi–Detector Row CT: Effect of Various Image Reconstruction Parameters and Segmentation Thresholds on Measurement Accuracy**
The results of the present study suggest that we should be more attentive to the selection of section thickness and segmentation threshold than to selection of field of view and reconstruction interval for an accurate measurement of nodule volume. Radiology 2005 235: 850-856

Wouter J. H. Veldkamp, Lucia J. M. Kroft, Bart J. A. Mertens, and Jacob Geleijns

**Digital Slot-Scan Charge-coupled Device Radiography versus AMBER and Bucky Screen-Film Radiography: Comparison of Image Quality in a Phantom Study**
Compared with conventional techniques, the charge-coupled device digital radiography system has potential for improved chest imaging. Radiology 2005 235: 857-866

Gao-Jun Teng and Qin Lu

**Bile Leakage during Transjugular Intrahepatic Portosystemic Shunt Creation: In Vitro Effect of Bile on Growth and Function of Human Umbilical Vein Endothelium**
The combination of data from the present study and our previous results suggests that bile leakage is an important factor in transjugular intrahepatic portosystemic shunt restenosis and that bile leakage contributes to restenosis apparently by inhibiting the growth and activity of endothelial cells rather than by promoting the proliferation of smooth muscle cells. Radiology 2005 235: 867-871.
Hanan I. Khalil, Stacey A. Patterson, and David M. Panicek

**Hepatic Lesions Deemed Too Small to Characterize at CT: Prevalence and Importance in Women with Breast Cancer**

Although hepatic lesions deemed too small to characterize are common at initial CT performed in women with breast cancer, in our study these lesions were of benign etiology in 92.7%–96.9% of the women in whom no definite liver metastasis was evident.


Erik K. Paulson, John P. Harris, Tracy A. Jaffé, Paul A. Haugan, and Rendon C. Nelson

**Acute Appendicitis: Added Diagnostic Value of Coronal Reformations from Isotropic Voxels at Multi–Detector Row CT**

We found that combined transverse and coronal reformations from isotropic voxels obtained with 16-section multi–detector row CT rather than transverse scans alone are useful in the evaluation of patients who are suspected of having acute appendicitis.

Radiology 2005 235: 879-885. Published online before print April 15 2005

Raghu Amaravadi, Marc S. Levine, Stephen E. Rubesin, Igor Laufer, Regina O. Redfern, and David A. Katzka

**Achalasia with Complete Relaxation of Lower Esophageal Sphincter: Radiographic-Manometric Correlation**

Our experience suggests that in patients with typical radiographic findings of achalasia, the barium study can be used to guide treatment without a need for manometry.


Sridhar Shankar, Eric vanSonnenberg, Jayesh Desai, Pamela J. DiPiro, Annick Van Den Abbeele, and George D. Demetri

**Gastrointestinal Stromal Tumor: New Nodule-within-a-Mass Pattern of Recurrence after Partial Response to Imatinib Mesylate**

Several patterns of disease recurrence can be identified at imaging in patients with gastrointestinal stromal tumor after an initial response to imatinib mesylate treatment, and we report the pattern of a nodule within a mass, a finding that is important for the diagnosis of progressive disease because conventional unidimensional, bidimensional, or volume measurements of the tumor mass do not enable the detection of nodular recurrence.

Zhen J. Wang, Benjamin M. Yeh, John P. Roberts, Richard S. Breiman, Aliya Qayyum, and Fergus V. Coakley

**Living Donor Candidates for Right Hepatic Lobe Transplantation: Evaluation at CT Cholangiography—Initial Experience**

CT cholangiography accurately depicts biliary tract anatomy in living donor candidates for right hepatic lobe transplantation, and donors who undergo preoperative CT cholangiography are unlikely to need conventional intraoperative cholangiography.


---

**Genitourinary Imaging:**

Dushyant V. Sahani, Neeraj Rastogi, Alan C. Greenfield, Sanjeeva P. Kalva, Dicken Ko, Sanjay Saini, Gordon Harris, and Peter R. Mueller

**Multi–Detector Row CT in Evaluation of 94 Living Renal Donors by Readers with Varied Experience**

Multi–detector row CT used as the sole imaging technique in the preoperative evaluation of living renal donors provides high accuracy even when images are read by multiple readers with varied levels of expertise.


Harriet C. Thoeny, Frederik De Keyzer, Raymond H. Oyen, and Ronald R. Peeters

**Diffusion-weighted MR Imaging of Kidneys in Healthy Volunteers and Patients with Parenchymal Diseases: Initial Experience**

Diffusion-weighted MR imaging of the kidneys gives reproducible noninvasive information on renal function in healthy volunteers, and it is also feasible in severely ill patients.


Russell N. Low, Bridgette Duggan, Robert M. Barone, Fred Saleh, and S. Y. Thomas Song

**Treated Ovarian Cancer: MR Imaging, Laparotomy Reassessment, and Serum CA-125 Values Compared with Clinical Outcome at 1 Year**

The results of MR imaging combined with serum CA-125 values are now used in lieu of laparotomy reassessment to assess tumor response to chemotherapy and to direct clinical management decisions.

Head and Neck Imaging:

Christopher Wu, Jingbo Zhang, Christopher J. Ladner, James S. Babb, Patrick J. Lamparello, and Glenn A. Krinsky

Subclavian Steal Syndrome: Diagnosis with Perfusion Metrics from Contrast-enhanced MR Angiographic Bolus-timing Examination—Initial Experience
A bolus-timing examination in the neck vessels during which a difference of 2 seconds or more is observed between the times to peak enhancement in the vertebral arteries may indicate subclavian steal syndrome in the appropriate clinical setting.

Health Policy and Practice:

Molly T. Beinfeld and G. Scott Gazelle

Diagnostic Imaging Costs: Are They Driving Up the Costs of Hospital Care?
Although increases in imaging costs contribute to increases in hospital costs, so too do many other factors, and imaging costs cannot fully explain the observed trends.

Medical Physics:


Comparative Scatter and Dose Performance of Slot-Scan and Full-Field Digital Chest Radiography Systems
Compared with conventional chest imaging systems with antiscatter grids, the digital slot-scan imaging system offers a marked advantage in terms of scatter reduction and an associated improvement in the effective detective quantum efficiency, which leads to improved image quality.

Molecular Imaging:

Dan Yang, Lin Han, and Vikas Kundra

Exogenous Gene Expression in Tumors: Noninvasive Quantification with Functional and Anatomic Imaging in a Mouse Model
Data suggest that noninvasive imaging criteria can be substituted for invasive methods for following gene expression in tumors.
Radiology 2005 235: 950-958
Charlotte Rivière, Frank P. Boudghène, Florence Gazeau, Jacky Roger, Jean-Noël Pons, Jean-Pierre Laissy, Eric Allaire, Jean-Baptiste Michel, Didier Letourneur, and Jean-François Deux

Iron Oxide Nanoparticle–labeled Rat Smooth Muscle Cells: Cardiac MR Imaging for Cell Graft Monitoring and Quantitation
Cell labeling with anionic magnetic nanoparticles demonstrated stable and long-lasting MR contrast properties without cell toxicity, allowing detection of iron-loaded cells in myocardium at MR imaging.

Musculoskeletal Imaging:

Christian W. A. Pfirrmann, Hubert P. Notzli, Claudio Dora, Juerg Hodler, and Marco Zanetti

Abductor Tendons and Muscles Assessed at MR Imaging after Total Hip Arthroplasty in Asymptomatic and Symptomatic Patients
Defects of the abductor tendons and fatty atrophy of the gluteus medius muscle and the posterior part of the gluteus minimus muscle are uncommon in asymptomatic patients after total hip arthroplasty and therefore appear to be clinically relevant.

Bernard Mengiardi, Christian W. A. Pfirrmann, Patrick Vienne, Hans-Peter Kundert, Pascal F. Rippstein, Hans Zollinger, Jürg Hodler, and Marco Zanetti

Anterior Tibial Tendon Abnormalities: MR Imaging Findings
Characteristic findings of anterior tibial tendon abnormalities include tendon thickening (>5 mm) and diffuse or posterior signal intensity abnormalities of the tendon within 3 cm from the distal point of insertion.

Neuroradiology:

Fumiyuki Yamasaki, Kaoru Kurisu, Kenichi Satoh, Kazunori Arita, Kazuhiko Sugiyama, Megu Ohtaki, Junko Takaba, Atushi Tominaga, Ryousuke Hanaya, Hiroyuki Yoshioka, Seiji Hama, Yoko Ito, Yoshinori Kajiwara, Kaita Yahara, Taiichi Saito, and Muhamad A. Thohar

Apparent Diffusion Coefficient of Human Brain Tumors at MR Imaging
We conclude that the apparent diffusion coefficient is useful for the differentiation of some human brain tumors.
Dong Gyu Na, Eung Yeop Kim, Jae Wook Ryoo, Kwang Ho Lee, Hong Gee Roh, Sam Soo Kim, In Chan Song, and Kee-Hyun Chang

CT Sign of Brain Swelling without Concomitant Parenchymal Hypoattenuation: Comparison with Diffusion- and Perfusion-weighted MR Imaging

This CT sign of early ischemic change does not represent severe ischemic damage, but it does suggest ischemic penumbral or oligemic tissue.


Yukio Miki, Milliam L. Kataoka, Toshiya Shibata, Tabassum L. Haque, Mitsunori Kanagaki, Taro Shimono, Tomohisa Okada, Akira Hiraga, Sadahiko Nishizawa, Hiroyuki Ueda, Mahbubur Rahman, and Junji Konishi

The Pituitary Gland: Changes on MR Images During the 1st Year after Delivery

The results of our longitudinal in vivo study revealed that both the volume of the pituitary gland and the signal intensity on T1-weighted MR images of the anterior lobe decrease until 8 months after delivery; the termination of lactation had no statistically significant effect on volume and T1-weighted signal intensity.


Obstetric Imaging:

Lee J. Brewerton, Radha S. Chari, Yuanyuan Liang, and Ravi Bhargava

Fetal Lung-to-Liver Signal Intensity Ratio at MR Imaging: Development of a Normal Scale and Possible Role in Predicting Pulmonary Hypoplasia in Utero

Data from this study have derived a normal range with a 95% prediction interval for lung-to-liver signal intensity ratio that will serve as a reference for future studies and aid in the assessment of fetal lung appearance.


Pediatric Imaging:

Nancy Rollins, Claro Ison, Tony Reyes, and Jon Chia

Cerebral MR Venography in Children: Comparison of 2D Time-of-Flight and Gadolinium-enhanced 3D Gradient-Echo Techniques

In this prospective nonblinded study, gadolinium-enhanced three-dimensional gradient-echo MR venography performed by using segmented central k-space ordering appeared to be superior to two-dimensional time-of-flight MR venography in the depiction of major draining veins in infants and children.


Fraukje Wiersma, Alexandr Srámek, and Herma C. Holscher

US Features of the Normal Appendix and Surrounding Area in Children

The results of this study show that in 82% of the children without clinical suspicion of appendicitis, the normal appendix can be depicted with US.

**Signs in Imaging:**

Tatum Simon Johnson

**The Spur Sign**

The spur sign is indicative of fracture in both the anterior and posterior acetabular columns (both-column fracture).

Radiology 2005 235: 1023-1024

---

**Special Reviews:**

Thomas G. Flohr, Stefan Schaller, Karl Stierstorfer, Herbert Bruder, Bernd M. Ohnesorge, and U. Joseph Schoepf

**Multi–Detector Row CT Systems and Image-Reconstruction Techniques**

In this article, we will review the general technical principles of multi–detector row CT as they apply to the established four– and eight–detector row systems, the more recent 16–detector row scanners, and generations of CT systems yet to come.

Radiology 2005 235: 756-773. Published online before print April 15 2005

---

**Thoracic Imaging:**

Conrad Wittram, Michael M. Maher, Elkan F. Halpern, and Jo-Anne O. Shepard

**Attenuation of Acute and Chronic Pulmonary Emboli**

This study demonstrates that chronic pulmonary embolism has a higher mean attenuation than acute pulmonary embolism, and we believe this difference is likely due to enhancement of organizing thrombus.


Patrick Berger, Vincent Perot, Pascal Desbarats, José Manuel Tunon-de-Lara, Roger Marthan, and François Laurent

**Airway Wall Thickness in Cigarette Smokers: Quantitative Thin-Section CT Assessment**

We validated an algorithm that can be used clinically to measure airway lumen and wall area on thin-section CT images and determined a new parameter that reflects airway wall thickness and that allows discrimination of healthy subjects from smokers with chronic obstructive pulmonary disease (COPD) or smokers without COPD.

Vascular and Interventional Radiology:

Elmar M. Merkle, Sherif Gamal Nour, and Jonathan S. Lewin

**MR Imaging Follow-up after Percutaneous Radiofrequency Ablation of Renal Cell Carcinoma: Findings in 18 Patients during First 6 Months**

Although the signal intensity characteristics of the kidneys on T2-weighted MR images obtained after radiofrequency (RF) thermal treatment are the same as those of the liver—both organs appear hypointense—the appearance of the renal RF thermal ablation zone on T1-weighted MR images is brighter than that of the hepatic RF thermal ablation zone.

Radiology 2005 235: 1065-1071

Byung Seop Shin, Young Soo Do, Byung Boong Lee, Dong Ik Kim, Ik Soo Chung, Hyun Sung Cho, Myung Hee Kim, Gaab Soo Kim, Chung Su Kim, Hong Sik Byun, Sung Wook Shin, and Kwang Bo Park

**Multistage Ethanol Sclerotherapy of Soft-Tissue Arteriovenous Malformations: Effect on Pulmonary Arterial Pressure**

A high incidence of acute pulmonary hypertension was observed in each session without lasting effect on pulmonary arterial pressure during multistage ethanol sclerotherapy.


Toyomichi Shibata, Kyo Itoh, Takeshi Kubo, Yoji Maetani, Toshiya Shibata, Kaori Togashi, and Koichi Tanaka

**Percutaneous Transhepatic Balloon Dilation of Portal Venous Stenosis in Patients with Living Donor Liver Transplantation**

Our findings suggest that percutaneous transhepatic balloon dilation is a safe and effective method for the treatment of portal venous stenosis after living donor liver transplantation.


Technical Developments:

Thomas K. F. Foo, Vincent B. Ho, Manojkumar Saranathan, Liu-quan Cheng, Hajime Sakuma, Dara L. Kraitchman, Katherine C. Wu, and David A. Bluemke

**Feasibility of Integrating High-Spatial-Resolution 3D Breath-hold Coronary MR Angiography with Myocardial Perfusion and Viability Examinations**

It seems feasible to complete a cardiac MR imaging examination that facilitates myocardial perfusion, myocardial viability, and proximal and middle coronary artery vessel assessments within 30–45 minutes.

Radiology 2005 235: 1025-1030
Richard D. Kacere, Mercedes Pereyra, Margit A. Nemeth, Raja Muthupillai, and Scott D. Flamm

**Quantitative Assessment of Left Ventricular Function: Steady-State Free Precession MR Imaging with or without Sensitivity Encoding**

The results of this study suggest that sensitivity encoding–assisted cine acquisitions with steady-state free precession (SSFP) can help to reduce examination time while producing left ventricular functional assessment data comparable to those achieved with conventional cine SSFP acquisitions.


Daniel Goldberg-Zimring, Bruria Shalmon, Kelly H. Zou, Haim Azhari, Dvora Nass, and Anat Achiron

**Assessment of Multiple Sclerosis Lesions with Spherical Harmonics: Comparison of MR Imaging and Pathologic Findings**

Spherical harmonics provided unbiased lesion volume and smaller variable biases, along with three-dimensional shape information.


Frank A. Morello, Jr, Kenneth C. Wright, and Thomas M. Lembo

**New Suction Guide Needle Designed to Reduce the Incidence of Biopsy-related Pneumothorax: Experimental Evaluation in Canine Model**

The results of the present experiment are promising because we significantly reduced the incidence of intraprocedural pneumothoraces by using the suction guide needle.

Radiology 2005 235: 1045-1049

**Letters to the Editor:**

Ukihide Tateishi, Kyung Soo Lee, and Joungho Han

**Vascular Endothelial Growth Factor–related Angiogenesis • Drs Lee and Han respond:**

Radiology 2005 235: 1084-1085

Flavio T. Braga, Antonio J. da Rocha, Ricardo B. Fonseca, Chantal Frigon, Dennis W. W. Shaw, Susan Heckbert, and Ed Weinberger

**Supplemental Oxygen Concentration and Increased Signal Intensity of Cerebrospinal Fluid on FLAIR MR Images • Dr Frigon and colleagues respond:**

Radiology 2005 235: 1085-1086

**Book Reviews:**

**PACS and Imaging Informatics: Basic Principles and Applications**

Radiology 2005 235: 740

**Essentials of Musculoskeletal Imaging**

Radiology 2005 235: 968
Science to Practice

Will Improved Vascular Mapping Achieved with Gadobenate Dimeglumine Aid in Interpretation of Breast MR Images?

The Setting

Use of gadopentetate dimeglumine (Magnevist; Schering, Berlin, Germany), the contrast agent most commonly used for breast magnetic resonance (MR) imaging, yields information about the kinetic characteristics (uptake and washout) of breast lesions. As with all gadolinium-based contrast agents, gadopentetate dimeglumine has a two-compartment biodistribution, with rapid diffusion from the vascular to the extracellular space. Another gadolinium-based contrast agent, gadobenate dimeglumine (MultiHance; Bracco Imaging, Milan, Italy), has similar two-compartment characteristics, but it also has transient protein binding such that its T1 relaxivity is approximately two times that of gadopentetate dimeglumine. Gadobenate dimeglumine has been tested in extensive trials in Europe and has recently received Food and Drug Administration approval in the United States.

In this issue of Radiology, Sardanelli et al (1) report the findings and data obtained in a European multicenter phase II trial that was performed by using three different doses of gadobenate dimeglumine—0.05, 0.10, and 0.20 mmol per kilogram of body weight—and a standard 0.10 mmol/kg dose of gadopentetate dimeglumine. They assessed the vascular mapping of the breast achieved with these two contrast agents and the relationship between ipsilateral increased vascularity and breast cancer.

The Science

Vascularity can increase not only within a breast cancer lesion but also in the ipsilateral breast as a whole. Sardanelli et al (1) obtained vascular maps by using the maximum intensity projections from subtracted MR images. These images were read by readers who were blinded to the final histopathologic diagnosis of the lesion. To simulate normal clinical interpretation conditions, however, the lesions were not masked. The readers rated the depicted vascularity of each breast on a scale of 0–3. Compared with the vascular maps obtained with gadopentetate dimeglumine enhancement, markedly improved vascular maps were obtained with use of all doses of gadobenate dimeglumine and enabled assessment of increased ipsilateral vascularity. This finding of increased ipsilateral vascularity correlated significantly with malignancy: Sensitivity, specificity, accuracy, positive predictive, and negative predictive values were 88% (44 of 50 patients), 82% (14 of 17 patients), 87% (58 of 67 patients), 94% (44 of 47 patients), and 70% (14 of 20 patients), respectively, suggesting that observed differences in overall breast vascularity could be used as a sign of malignancy.

The assessment of vascularity, however, is somewhat subjective. The investigators strived for objectivity by using specific criteria such as the size (width and length) and number of vessels observed. Given that the lesions were not masked during readings, there was the potential for bias in the interpretations. Nonetheless, the findings were intuitive. Neoangiogenesis is known to occur in malignancies, and contrast agents that yield vascular maps may yield visible differences that are clinically important.

Clinical utility will depend on whether the combination of breast vascular mapping and the traditional criteria used to determine the likelihood of malignancy (morphologic and kinetic features) will improve the interpretation of breast MR images.

The Practice

Clinical use.—Gadobenate dimeglumine has been shown to have a safety profile similar to that of gadopentetate dimeglumine; only nonserious adverse reactions have been encountered (2). Thus, safety considerations should not hamper the adoption of this agent into clinical use.

While MR imaging is very sensitive for the detection of invasive cancer, its specificity is more variable. The sensitivity of MR imaging for the detection of intraductal carcinoma is generally lower than that for the detection of invasive disease. Clinically valuable advancements in breast MR imaging will be based on improvements in intraductal carcinoma detection (ie, sensitivity) and overall specificity (ie, differentiation of benign lesions with enough certainty to avoid biopsy).

Future opportunities and challenges.—While the re-
sults reported in the Sardanelli et al article (1) are interesting, it is uncertain whether the findings will influence breast MR imaging. Clinical utility will depend on whether the combination of breast vascular mapping and the traditional criteria used to determine the likelihood of malignancy (morphologic and kinetic features) will improve the interpretation of breast MR images. Further studies will be necessary to determine this, along with whether the finding of increased vascularity is observed with intraductal tumors also.

Summary

The findings of Sardanelli et al (1) show not only that gadobenate dimeglumine can be used to obtain vascular maps that are superior to those obtained with gadopentetate dimeglumine but also that a difference in vascularity between the two breasts is an indirect sign of malignancy. Further examinations will be necessary to assess whether the finding of ipsilateral vascularity, when combined with other interpretation criteria, improves overall diagnostic accuracy.

References


Liane E. Philpotts, MD
Department of Radiology, Yale University School of Medicine, 333 Cedar St, PO Box 208042, New Haven, CT 06520-8042
Liane.Philpotts@yale.edu
In November 1972, John Holt wrote an editorial in *American Journal of Roentgenology*, *Radium Therapy*, and *Nuclear Medicine* (1) in which he said there were too many pediatric radiologists in practice, there were too many radiology residents entering pediatric fellowship training, and radiology at that time had a crop of pediatric radiologists that was too young to make it a wise career choice for large numbers of radiology residents. What a difference a generation makes!

Try to hire a pediatric radiologist. They are in short supply. One hundred one positions for pediatric radiologists went unfilled in 2003, and by the end of 2004 the number had reached 117 (S. Royal, Society of Pediatric Radiology Committee on Physician Workforce 2004, personal communication, January 2005). That number represents about 15% of the total number of pediatric radiologists who practice in North America. For the past 5 years, the number of new radiologists has declined between 20 and 35 (S. Royal, Society of Pediatric Radiology Report of Membership Committee, personal communication, January 2005). At that rate, it will take 3–5 years just to fill the available positions if no one retires. The mean age of pediatric radiologists is the oldest of the radiology subspecialists—by 10 years in some cases (2). Retirement looms closer for a large percentage of pediatric radiologists with each passing year.

On the basis of our own experience and what we have learned from others, the number of applicants for pediatric radiology fellowship positions either remains stable or, in recent years, has had several periods of decline. No new fellowship training programs are being added, and many existing programs risk losing their accreditation because they have had no trainees in several years. In addition, an increasing percentage of pediatric radiologists are women. The percentage of women who work part time is greater than the percentage of men who work part time, resulting in fewer “full-time equivalents” in the active workforce (2,3).

Most radiology groups, including our own, have seen an annual workload increase of about 4%–6%. Academic institutions, where the majority of pediatric radiologists practice, also report that the workload for their radiologists continues to increase annually by about 5%–6%. Pediatric hospitals report a similar steady increase in radiology workload, as measured in relative value units (4,5).

At the same time, many third-party payers are contracting with companies that manage radiology benefits and costs to health plans. According to one of these companies, National Imaging Associates, “about one third of advanced imaging tests are either inappropriate for the medical problem or don’t contribute to a doctor’s diagnosis or a patient’s outcome” (6). These efforts in controlling medical imaging costs have resulted in complex requirements for preauthorization and the diversion of more expensive imaging examinations, such as computed tomography (CT) and magnetic resonance (MR) imaging, away from academic or specialty centers toward less expensive providers. That, however, does not always reduce work demands for pediatric radiologists, although it may decrease costs to the third-party payer. For example, at our pediatric institution we interpret or provide a second opinion on CT or MR images obtained outside our institution approximately 3000 times each year. The studies are performed at other locations either because third-party payers refer them there, as a part of a program like National Imaging Associates, or because a pediatric radiologist was not available at that location to obtain and interpret the images. These two trends reinforce the concern about an absence of pediatric radiologists and also point out how an overworked staff of pediatric radiologists ends up doing extra work with little or no financial compensation because of the quirks in the system. We will say more about financial inequities later in this article.

Great! Rejoice! Fewer workers and more work means more money for the remaining supply of radiologists. Alas, the workplace does not work that way. In 1994, Apple Computer controlled 14% of the marketplace with its Macintosh operating system (7). The Macintosh arguably was a very user-friendly computer operating system. Its future potential to capture the majority of customers and businesses buying computer systems was enormous. Software developers wrote new software for both the Macintosh operating system and its rival the personal computer, or PC, operating system. For a variety of reasons that had nothing to do with the quality of the Macintosh operating system, by 1996 the percentage of market share owned by the Macintosh had dropped below 6% (7). Apple Computer had crossed a critical threshold; it had reached a “tipping point.” As the market share of the Macintosh operating system dropped below this critical level, a chain of events occurred that made the demise of the organization imminent. In Apple Computer’s case, independent software developers wrote their software for the 92% of the market share that the Macintosh operating system did not control (7). Software engineers did not write...
software for the Macintosh system or wrote it only as an afterthought. Customers who wanted to use Macintosh operating systems could not get the type and quality of software for their desktop computers that they wanted. The end for the Macintosh loomed near.

Apple Computer did not die, but it had to reinvent itself, and it lost the battle for who controls personal computer systems. Macintosh computers are still sold throughout the country, but when the percentage of market share dropped below the tipping point, Macintosh lost the ability to dominate and control the world of computers. Similar examples exist throughout the business world.

When a company or organization crosses a certain threshold and its market share drops below a certain percentage, it loses its relevance, and the marketplace looks for alternative solutions to the problems that the organization was solving. The exact percentage for any one type of organization is not known; for Apple Computer and Macintosh it was somewhere below 10%. For other types of businesses and organizations it may be another number, but the concept is recognized and makes perfectly good sense. When a company cannot provide the services that its customer base needs, customers find a different way to satisfy their needs.

If pediatric radiology is not able to supply the needs of its referring physicians because of an inadequate workforce supply, pediatric radiology could cross its tipping point and become irrelevant. Customers would find a different method to satisfy their needs. The truth of this is apparent to almost all radiologists. Those locations that do not have pediatric radiologists—and the numbers are increasing—today find the workload shifts to radiologists who interpret adult images and who may be uncomfortable handling pediatric patient cases. That is one potential scenario as the demand for pediatric radiology services continues to increase faster than the supply. The authors are not screaming that the end is near. The authors think pediatric radiology can be saved. But radiology leaders must recognize the impending crisis and take actions to preserve an important subspecialty of radiology. If radiology does not meet the demand for pediatric radiologists, pediatric radiology could quickly reach its tipping point. In the best of the bad-case scenarios, the marketplace will force adult radiologists to perform pediatric radiology studies. In the worst-case scenarios, the various pediatric clinical specialists will fill the vacuum and perform the pediatric imaging. Even today in pediatric radiology there are nephrologists and urologists who want to perform their own renal ultrasonographic (US) studies, orthopedic surgeons who want to interpret their own musculoskeletal MR images, and emergency room physicians advocating their use of US for abdominal trauma. Radiology’s history of holding on to specialized turf has not been good. Cardiologists, obstetricians, and others have slowly captured market share that once belonged exclusively to the radiologist. Why? The inability to provide the services wanted by the referring base ranks as one of the leading reasons. Others might argue that money also is a factor, but let’s be generous and say that perhaps the absence of radiology services available around the clock might at least be a contributing factor.

How can pediatric radiologists retain control of their market share? How can they expand their ability to do the increasing amount of work? How can they capture new market share and do more of the pediatric work that, by default, their adult radiology colleagues are performing and may not want to perform? We propose several strategies. Although we make no claim about the originality of these various strategies, all of them probably need to be implemented, and some of them may even help. By calling the impending crisis to the attention of the radiology community, we hope that there are additional ideas and strategies that others can conceive that we have not even dreamt possible. The critical point is that all of organized radiology must recognize the impending crisis and take actions to remedy the situation prior to reaching that critical level where market share rapidly disappears.

We will discuss several potential strategies that, if implemented either individually or together, may have a substantial positive effect on pediatric radiology.

**Increase the Number of Pediatric Radiologists**

Radiology must demonstrate that pediatric radiology is an exciting, attractive career choice for young physicians in training—the earlier in their training they get this message, the better. Pediatric radiologists need to aggressively move into the lecture circuit for medical students and radiology residents. They should participate in the development of a national medical student curriculum in radiology (8). The frequent exposure to pediatric radiologists keeps the topics of pediatric radiology in front of the potential recruits on a daily or weekly basis. It allows them to see the effect pediatric radiologists have in the care of children, the ways in which they are useful and respected by their colleagues, and the wide array of interesting problems that they encounter on a daily basis.

The younger generation of physicians makes career choices with a value system and set of goals that may not be exactly the same as those of their predecessors. The younger generation expresses concern over the effect that career choices have on family life. These trends and value choice differences are nicely discussed in Geeks and Geezers: How Era, Values, and Defining Moments Shape Leaders, an insightful book by Warren Bennis and Robert Thomas (9). It is important for the older generations to understand these differences in motivation, since many of the leaders qualify as older members of the radiology profession. The key is for the leaders to recognize that younger members of the medical profession may have very different goals and ideas about what they want to do with their career than did the current leaders when they were just starting in the field of medicine. Engaging trainees in a meaningful discussion of lifestyle and financial and work flexibility can have a very positive effect on the career choice of a physician in training at several different time points. It is especially important to address any potential misconceptions about pediatric radiology, including the need to practice in a rigorous academic environment, the limited supply of positions, and financial discrepancies (10). At worst, physicians in training will learn more about pediatric radiology. At best, they will learn the unique and exciting things that make pediatric radiology fun. Hopefully, some will decide that a career in pediatric radiology excites them.

**Reduce Barriers to Entry in the Subspecialty**

The educational pathways leading to a career in pediatric radiology need to be carefully examined, with the goal of reducing the amount of training required to reach certification. Is each step necessary? Is there a better way? This may require some radical thinking rather than adherence to a tradition that has, historically, simply added years to the training process at each step. Some thought needs to be focused on what it is in the way of training we want to require for physicians to be
pediatric radiologists. Changes might be accomplished in a number of ways, such as the development of a focused “pediatric track” during radiology residency. The American Board of Radiology has approved an experimental pathway called Pediatric Emphasis Diagnostic Radiology Alternative Pathway, or PEDRAP. With this plan, radiology residents would spend half of their time during residency training focused on pediatric radiology and the other half focused on adult radiology. The radiology resident who follows this training pathway will likely have come from a background in pediatrics (11). Some leaders have suggested a variation on the procedure, such as identifying pediatricians interested in radiology and building a special training program for them.

It is not clear how the problem of board certification with special qualification in pediatric radiology will be handled with either of these two suggestions, but that is a topic that needs to be solved. If the fellowship year could be eliminated, that might make the field attractive to some pediatricians who discover that they really prefer practicing radiology to pediatrics but abhor the long training required by both residency and fellowship in radiology.

In response to marketplace pressures, some institutions have changed the length of time required for a pediatric radiology fellowship. The Children’s Hospital in Boston, Massachusetts, once required each radiology fellow to commit to a 2-year fellowship. Now many of their fellows spend only a single year in the fellowship and move on to positions elsewhere. This change met considerable resistance on the part of some staff radiologists at Children’s Hospital, but it was implemented with the intention of trying to meet market demands and the rising shortage of pediatric radiologists.

**Increase the Workforce in Pediatric Radiology**

In any industry, a disruptive innovation sneaks in from below. While the dominant players are focused on improving their products or services to the point where the average consumer doesn’t even know what [he or] she’s using (think overengineered computers), they miss simpler, more convenient, and less costly offerings initially designed to appeal to the low end of the market. Over time, the simpler offerings get better—so much better that they meet the needs of the majority of users. (12)

This quotation from the *Harvard Business Review* succinctly summarizes one viewpoint that applies equally well to radiology as it does to other forms of specialization. People working in a specialized field tend to develop a very narrow focus on how to improve their products or services. Radiology needs to take a serious look at the ways in which pediatric radiologists add value to the diagnostic process. Does it really take someone with full residencies in pediatrics and radiology, a 2-year fellowship in pediatric radiology, and 20 years of clinical experience to interpret follow-up fracture images or the fifth scoliosis series? If we add no meaningful value to the care of the child, then let’s not do it, or let’s delegate the activity to a more appropriate level of training or expertise. A provocative article by Christensen et al (12) in the *Harvard Business Review* proposes that the health care industry focus on “enabling less expensive professionals to do progressively more sophisticated things in less expensive settings” as a potential solution to the current health care crisis. This radical proposal touches on the controversy of the potential role of a physician extender. Other specialties have used physician extenders or physician assistants (13). Why not radiology?

For example, many technologists have years of experience but have a limited career pathway beyond promotion to a higher grade level or into more complex imaging techniques. We should explore the possibility of using technologists in expanded practice models. Specific sets of tasks would be identified as appropriate for a technologist to perform. A course of instruction would be organized. Selected technologists would complete the course, be tested initially and on a periodic basis, and then perform those tasks under the supervision of a radiologist. This would not be an independent practice of radiology, but merely the use of a highly skilled and experienced technologist to perform more routine tasks in a highly supervised environment. The American College of Radiology is currently discussing these possibilities. Perhaps pediatric radiology is the right environment and gives us the opportunity to test this experiment (14).

**Increase Efficiency**

All radiologists need to leverage the power of technology (eg, picture archiving and communications systems [PACS], voice recognition systems) to make work more efficient. The data do not yet convince everyone that PACS make radiologists more efficient, but the evidence is accumulating. There is a suggestion that PACS allow subspecialists like pediatric radiologists to extend their expertise beyond the geographic borders of their institutions. This could, in theory, supply needed services to some remote locations or to those without a pediatric radiologist available. Clearly, this will only help with the maldistribution of pediatric radiologists. There still needs to be a sufficient number of pediatric radiologists available to actually perform the work. New technology systems will evolve that may allow radiologists to operate more efficiently. Leaders in radiology need to become especially alert to experimenting with new methods to leverage technology to improve workflow processes. This will require leaders with the mind-set to experiment and the ability to fail.

**Economic Incentives**

Compensation for pediatric radiologists has historically lagged behind that for their adult radiologist counterparts. Third-party payers have paid less for pediatric admissions and procedures than for similar adult admissions and procedures. Pediatric radiologists do not perform the same percentage of high-technology studies, such as CT and MR imaging, and highly compensated studies as their colleagues in adult radiology. It generally takes longer to perform the same type of procedure on a child than on an adult. Therefore, pediatric radiologists usually generate a lower volume of images than do their adult colleagues, and their work consists of lower relative value units. All of these factors account for less pay to pediatric radiologists than adult radiologists.

Radiology needs to make a convincing case to third-party payers that a premium should be paid for subspecialty work done by pediatric radiologists. This may require a joint educational effort to persuade parents that a study performed by a pediatric radiologist is more likely to satisfy the needs and consider the peculiar demands of children and their parents, intuitively resulting in a better study for the child. Third-party payers need to understand this. Parents and pediatricians jointly may need to help persuade insurance companies that pediatric radiologists add value to the care of children (15).

Improving customer service to referring physicians will make the skills and talents of subspecialists more attractive to clinicians. All radiologists could benefit by this, but pediatric radiologists have a unique relationship with their referring physician base. In general, the desire for consultation with radiologists is greater among pediatricians than among adult...
Internists and surgeons. Pediatric radiologists should therefore make a compelling case to parents as to how much value the radiologist adds to the care of the pediatric patient. Showing pediatric specialists the new technology available to help them diagnose and care for their patients should be a coordinated marketing strategy of the pediatric radiology community. It may be necessary to increase ties with subspecialty pediatricians to provide new technology to help them.

**Geographic Maldistribution of Pediatric Patients Needing Radiologists**

Most pediatric radiologists practice in academic institutions. The opportunities for pediatric radiologists to practice their skills in rural locations or even private practice groups are few. As with any specialty, a certain critical mass of patients is needed to generate enough business to make it financially worthwhile. There are many radiology groups where there are not a sufficient number of pediatric patients for one or two radiologists to focus entirely on pediatric patients during the workday. At those locations, the pediatric work is distributed among the various radiologists, or a single radiologist spends 50%–60% of the time doing pediatric work and the remaining workday is spent doing adult radiology work. Several groups in the same geographic location may have the same situation, but because those groups are competitive rivals they do not share their pediatric workload. Thus, the efficiency of caring for pediatric patients in an optimum manner takes a backseat to the practicalities of the workplace and competitive marketplace. Can we somehow solve this problem? Can we develop methods to transmit images for specialty interpretation without the fear that the sending organization will lose revenue or that the receiving organization is doing work for no compensation?

Leaders need to acknowledge that a problem exists and take bold innovative steps. We need to try different, unique, and innovative methods. As mentioned, we need to not be afraid to fail when trying out new ideas. The end result is too critical to simply wait for a solution to evolve. Letting things evolve on their own represents a conscious decision not to do something with a looming crisis. Radiology needs to learn from other business and management failures and from its own failure in retaining market share; experience with cardiology and some obstetric imaging comes to mind. Unfortunately, leaders are often the most resistant to trying new experiments in how they practice and administer radiology. As Hamel succinctly phrased it, “most management processes are controlled by defenders of the past” (16). Experience in businesses demonstrates that to survive, an organization must adapt. An organization will not adapt unless someone steps up and provides some direction that the organization can take. Who is going to provide that leadership? Can we save pediatric radiology? Do we want to expend the energy, try the experiments, suffer the failures, and find a way to avoid the tipping point that spells the demise of pediatric radiology?

**References**

American College of Radiology Clinical Statement on Noninvasive Cardiac Imaging

EDITOR’S NOTE: This article is simultaneously published in the June 2005 printed issues of Radiology and the Journal of the American College of Radiology. The copyright is held by the American College of Radiology.

GENERAL INTRODUCTION

Coronary artery disease (CAD) and other acquired and congenital cardiac diseases are major medical and socioeconomic problems. CAD affects 13.2 million Americans and was responsible for 502,189 deaths in 2001. In 2004, the direct and indirect economic impact of CAD was in excess of $120 billion, which was about one-third of the total costs attributable to cardiovascular diseases (1).

Historically, imaging has had a critical role in the diagnosis and evaluation of acquired and congenital cardiac disease, beginning with chest radiography and fluoroscopy and progressing to coronary angiography and cardiac catheterization, ultrasonography (echocardiography), and nuclear medicine. All of these modalities have a well-established role in patient care. Computed tomography (CT), with multidetector CT and electron-beam technology, and magnetic resonance (MR) imaging, with appropriately equipped imagers, now can image the coronary arteries, cardiac chambers, valves, myocardium, and pericardium and can help assess cardiac function. Thus, CT and MR imaging will have an increasing role in comprehensive cardiac imaging.

While the technical parameters and field of view of a cardiac CT or MR examination will appropriately be tailored to help evaluate the cardiac anatomy and/or function in question, the images obtained will demonstrate adjacent anatomy, often including portions of the lungs, mediastinum, spine, and upper abdomen. It has been documented that these studies often demonstrate clinically significant noncardiac findings (2,3). In addition to examining the cardiac structures of interest, the interpreting physician is responsible for examining all the visualized noncardiac structures and must report any clinically relevant abnormalities of these adjacent structures. In some cases, these structures may be seen only on localizing (scout) images.

Cardiac CT and cardiac MR imaging each present potential patient safety issues. Cardiac CT safety issues are related to radiation exposure and to administration of intravascular (IV) contrast media. The safety concerns for cardiac MR imaging are primarily related to the strong magnetic field and its potential effect on implanted devices, but MR imaging contrast agents and patient sedation also present potential safety issues. In addition, pharmacologic agents may be administered for either CT or MR imaging examinations.

Radiologists, because of their extensive experience in CT and MR imaging, have an important role in imaging of cardiac patients with these modalities. Most radiologists already supervise the performance of CT and interpret CT scans of the chest (including basic evaluation of the pericardium, heart size, and cardiac masses) and CT angiographic and MR angiographic images. Their knowledge of structures beyond the heart provides added value in cardiac imaging. They already oversee CT and MR imaging equipment and personnel. Their experience with the techniques now being applied to the heart provides expertise to develop specific cardiac applications of CT and MR imaging.
MR imaging, and it shortens their learning curve for these cardiac applications.

The American College of Radiology (ACR) develops and revises practice guidelines and technical standards that address a wide range of imaging applications. Existing practice guidelines address many areas related to cardiac CT and MR imaging. These include the following: “ACR Practice Guideline for Performing and Interpreting Diagnostic Computed Tomography (CT),” “ACR Practice Guideline for Performing and Interpreting Magnetic Resonance Imaging (MRI),” “ACR Practice Guideline for the Performance and Interpretation of CT Angiography (CTA),” “ACR Practice Guideline for the Performance and Interpretation of Pediatric and Adult Body Magnetic Resonance Imaging (MRA),” “ACR Practice Guideline for the Performance of Cardiovascular Magnetic Resonance Imaging (MRI),” “ACR Practice Guideline for the Use of Intravascular Contrast Media,” “ACR Practice Guideline for Adult Sedation/Analgesia,” and “ACR Practice Guideline for Pediatric Sedation/Analgesia.”

This clinical statement of the ACR discusses various technical and patient safety issues related to cardiac CT and MR imaging, and it suggests appropriate qualifications for radiologists until such time as ACR practice guidelines for the performance of cardiac CT and cardiac MR imaging are written and approved through the usual ACR process. Issues related to vascular CT and MR are addressed in documents listed in the preceding paragraph.

CARDIAC CT
Introduction

CT is a proven and important imaging modality for the detection and characterization of cardiac disease (4). CT may be used as either the primary modality for detecting disease or as an adjunct to other imaging modalities to better characterize disease and help assess change over time. CT can be used to assess both cardiac structure and function (5,6), as well as evaluate disease processes within the field of view but outside of the heart and pericardium (7,8).

Applications of cardiac CT include but are not limited to the following (5,9–17):

(a) detection and characterization of coronary artery occlusive lesions secondary to atherosclerosis, transplant arteriopathy, intimal dissection, and vasculitis;
(b) detection and characterization of coronary artery anomalies;
(c) detection and characterization of coronary artery aneurysms;
(d) coronary vein mapping;
(e) characterization of cardiac chamber morphology and function;
(f) characterization of native and prosthetic cardiac valves;
(g) detection and characterization of congenital heart diseases;
(h) characterization of cardiac masses;
(i) diagnosis of pericardial diseases; and
(j) detection and characterization of postoperative abnormalities.

Qualifications of Personnel

The Radiologist

The radiologist who supervises and interprets cardiac CT examinations should meet the following criteria for calcium scoring:

By virtue of experience and residency training, which has included cardiac anatomy and CT physics, a board-certified radiologist is qualified to perform calcium scoring of coronary arteries. It is expected that board-certified radiologists will be familiar with the indications and techniques for, as well as the interpretation of, coronary artery calcium scoring.

The radiologist should meet the following criteria for cardiac CT (not including examinations performed exclusively for calcium scoring):

1. Certification in radiology or diagnostic radiology by the American Board of Radiology, the American Osteopathic Board of Radiology, the Royal College of Physicians and Surgeon of Canada, or Le College des Medecins du Quebec and have supervised and interpreted 75 cardiac CT cases, excluding those performed exclusively for calcium scoring, in the past 36 months.

OR

Completed an Accreditation Council for Graduate Medical Education (ACGME)-approved residency program and have supervised and interpreted 75 cardiac CT cases, excluding those performed exclusively for calcium scoring, in the past 36 months.

AND

2. Completed at least 40 hours of category I continuing medical education in cardiac imaging, including cardiac CT, anatomy, physiology, and/or pathology or documented equivalent supervised experience in a center actively performing cardiac CT.

Maintenance of competence.—All radiologists’ performing cardiac CT examinations should demonstrate evidence of continuing competence in the interpretation and reporting of those examinations. If competence is assured primarily on the basis of continuing experience, a minimum of 75 examinations, excluding those performed exclusively for calcium scoring, every 3 years is recommended in order to maintain the radiologist’s skills.

Continuing medical education.—The radiologist’s continuing medical education should be in accordance with the “ACR Practice Guideline for Continuing Medical Education (CME)” of 150 hours of approved education every 3 years, and should include continuing medical education in general and in cardiac CT as is appropriate to the radiologist’s practice needs.

The Technologist

Technologists performing CT examinations should be certified by the American Registry of Radiologic Technologists (ARRT) or have an unrestricted state license with documented training and experience in cardiac imaging procedures. It is recommended that the technologist performing cardiac CT have advanced certification in CT. Each technologist should have supervised experience in the performance of cardiac CT examinations and in the intravenous administration of conventional CT contrast agents. If intravenous contrast material is to be administered, qualifications for technologists performing intravenous injections should be in compliance with current ACR policy statements and existing operating procedures or manuals at the imaging facility. (See the “ACR Practice Guideline for the Use of Intravascular Contrast Media.”) The American College of Radiology approves of the injection of contrast material and diagnostic levels of radiopharmaceuticals by certified and/or licensed radiologic technologists and radiologic nurses under the direction of a radiologist or his or her physician-designee who is personally and immediately available, if the practice is in compliance with institutional and state regulations. There must be prior written approval by the medical director of the radiology department or service of such individuals, such approval process having followed established policies and procedures, and the radiologic technologists and radiologic nurses who have been so approved maintain documentation of continuing medical education related to the materials being injected and to the proce-
Cardiac CT Safety Issues

Safety issues in cardiac CT relate to radiation exposure, IV contrast material administration, and β-blocker and nitrate administration.

With regard to radiation exposure, the supervising physician should be familiar with the various technical parameters of the examination that affect radiation dosage, including milliamperes-seconds (mA) and peak voltage settings (kVp) and scan pitch. Moreover, automated x-ray dose-shaping algorithms and x-ray tube pulsing, when available, should be applied to minimize radiation exposure while allowing diagnostic image quality. As with all examinations that use ionizing radiation, cardiac CT should be performed with a radiation dose that is as low as reasonably achievable, or ALARA, without compromise to the resulting images. This is especially important for cardiac CT patients, since they may undergo many radiographic examinations, including fluoroscopically guided interventional cardiac procedures that may require a high radiation dose. Particular attention to radiation dose is needed for children and young adults, who are more susceptible to the effects of radiation, and especially for young female patients, since the breasts will likely be within the area of scanning. As a general rule a multi-detector CT scan encompassing the heart should not result in a volume CT dose index (CTDI) greater than 60 mGy or an effective dose of greater than 13 mSv (18,19).

With regard to the administration of IV contrast media, the physician should supervise patient selection to identify those patients for whom IV contrast media may present an increased risk or be contraindicated, particularly in those patients with renal insufficiency and/or a history of reaction to contrast media. Some of these patients may require pretreatment to allow safe contrast agent administration. The physician should also be available to treat adverse reactions to IV contrast media. The ACR Practice Guideline for the Use of Intravascular Contrast Media (20) and the ACR Manual on Contrast Media (21) are helpful resources in this area.

β-Blockers and nitrates are commonly used in conjunction with cardiac CT studies. Physicians performing cardiac CT should be knowledgeable about the administration, risks, and contraindications of these drugs. Blood pressure and heart rate should be monitored.

Cardiac CT Equipment Recommendations

The availability of a multi-detector row helical CT or an electron-beam CT scanner is a requirement for cardiac CT applications, especially for coronary artery calcium scoring and CT angiography. For multi-detector row CT, at least four detector rows are preferred for calcium scoring and at least 16 are preferred for CT coronary angiography. The temporal imaging capability should be 500 msec or less, and spatial resolution should be such that in-plane voxels that approach 0.5 mm³ are obtainable. The capability to image a section thickness of less than 1.0–1.5 mm is also necessary for coronary imaging, as is electrocardiographic gating and the ability to acquire images in both prospective and retrospective modes.

A powered contrast medium injector that allows programming of both the volume and flow rate of the contrast agent must be used for many contrast medium-enhanced cardiac CT examinations.

A workstation capable of creating multiplanar reformations, maximum intensity projections, and volume renderings or shaded surface displays should be available.

CARDIAC MR IMAGING

Introduction

Clinical application.—Cardiac MR imaging represents the specialized application of MR to imaging the heart to help diagnose both acquired and congenital disease. Applications of cardiac MR include, but are not limited to, the following (36–49): (a) assessment of myocardial scar, infiltrative processes, and inflammation; (b) assessment of myocardial ischemia; (c) assessment of ventricular function; (d) characterization of cardiac chamber morphology and function; (e) detection and characterization of congenital heart disease; (f) characterization of cardiac masses; (g) diagnosis of pericardial disease; (h) quantification of valvular disease and shunt physiology; (i) detection of coronary artery atherosclerosis; (j) detection and characterization of coronary artery anomalies; and (k) detection and characterization of coronary artery aneurysms.

Technical specifications.—A physician who performs cardiac MR should be familiar with all aspects of the MR examination. This includes not only clinical indications, but also technical specifications. This should include a thorough knowledge of cardiac MR pulse sequences, which include, but are not limited to, gradient-recalled-echo sequences including steady-state balanced methods (fast imaging with steady-state precession, or true FISP; fast imaging employing steady-state acquisition, or FIESTA; balanced fast field echo), fast spin-echo and half-Fourier spin-echo sequences, phase-contrast and flow-quantification methods, and contrast-enhanced MR angiography techniques (39,41,43,50–55). The cardiac MR physician should also have a background in MR physics as related to cardiac MR to include MR parameters, artifacts, k-space, and image formation, along with knowledge of hardware components such as electrocardiography leads, methods of gating, and basic coil function and design. MR physics training is part of the core curriculum of the radiology residency, and completion of the radiology residency is accepted as evidence of MR physics training.

Qualifications of Personnel

The Radiologist

The radiologist who supervises and interprets cardiac MR examinations should meet the following criteria for cardiac MR:

1. Certification in radiology or diagnostic radiology by the American Board of Radiology, the American Osteopathic Board of Radiology, the Royal College of Physicians and Surgeons of Canada, or Le College des Medecins du Quebec and have supervised and interpreted 75 cardiac MR cases in the past 36 months.

OR

Completed an ACGME-approved radiology residency program and have super-
vised and interpreted 75 cardiac MR cases in the past 36 months. AND

2. Completed at least 40 hours of category I continuing medical education in cardiac imaging, including cardiac MR, anatomy, physiology, and/or pathology, or have documented equivalent supervised experience in a center where cardiac MR is actively performed.

For pharmacologic stress testing, the radiologist should meet the following criteria:

Radiologists performing pharmacologic stress testing as part of cardiac MR imaging should be knowledgeable about the administration, risks, and contraindications of pharmacologic agents used for stress testing.

Personnel monitoring stress-induced studies should have current advanced cardiac life support certification.

Maintenance of competence.—All radiologists who perform cardiac MR should demonstrate evidence of continuing competence in the interpretation and reporting of those examinations. If competence is assured primarily on the basis of continuing experience, a minimum of 75 examinations every 3 years is recommended in order to maintain the radiologist’s skills.

Continuing medical education.—The radiologist’s continuing medical education should be in accordance with the “ACR Practice Guideline for Continuing Medical Education (CME)” of 150 hours of approved education every 3 years and should include education in general and cardiac MR as appropriate to the radiologist’s practice needs.

The Technologist

The technologist who performs cardiac MR should be certified by the ARRT or have an unrestricted state license with documented training and experience in cardiac imaging procedures. It is recommended that the technologist performing cardiac MR have advanced certification in MR. Each technologist should have supervised experience in the performance of cardiac MR and in the IV administration of conventional MR contrast agents. If IV contrast material is to be administered, qualifications for technologists who performing such injections should be in compliance with current ACR policy statements (20,21) and existing operating procedures or manuals at the imaging facility. (See the ACR Practice Guideline for the Use of Intravascular Contrast Media. The American College of Radiology approves of the injection of contrast material and diagnostic levels of radiopharmaceuticals by certified and/or licensed radiologic technologists and radiologic nurses under the direction of a radiologist or his or her physician-designee who is personally and immediately available, if the practice is in compliance with institutional and state regulations. There must be prior written approval by the medical director of the radiology department or service of such individuals, such approval process having followed established policies and procedures, and the radiologic technologists and radiologic nurses who have been so approved maintain documentation of continuing medical education related to the materials being injected and to the procedures being performed [resolution 1-H; 1987, 1997].)

In addition, each technologist should maintain 24 hours of continuing education every 2 years, as stipulated by the ARRT. It is recommended that these credits include those from activities that provide education in the performance of cardiac MR. It is also recommended that technologists performing CMR examinations maintain basic life support certification and be capable of using an automatic external defibrillator.

To ensure competence, all technologists must be evaluated by the supervising radiologist.

Cardiac MR Safety Issues

The cardiac MR physician should have thorough knowledge of patient safety to include specific-absorption-rate, or SAR, limits, possible neurologic effects, tissue heat deposition, and contraindications to MR imaging, such as implantable devices (56).

With regard to the administration of IV contrast media, the physician should supervise patient selection to identify those patients for whom IV contrast medium administration may present an increased risk or be contraindicated. Although contrast agent reactions occur less frequently with gadolinium-based contrast agents than with iodinated agents, some patients may require pretreatment to allow safe contrast medium administration. The physician should also be available to treat adverse reactions to IV contrast media. The “ACR Practice Guideline for the Use of Intravascular Contrast Media” (20) and the ACR Manual on Contrast Media (21) are helpful resources in this area.

Cardiac MR Equipment Recommendations

Imagers for clinical cardiac MR should be accredited by the ACR, with equipment performance monitoring in accordance with the ACR Technical Standard for Diagnostic Medical Physics Performance Monitoring of MR imaging Equipment (57).

The MR imager should be capable of fast three-dimensional gradient-echo imaging, steady-state imaging with free precession, phase-contrast flow quantification, fast multisection myocardial perfusion imaging, and delayed contrast-enhanced myocardial imaging. Commercial Food and Drug Administration–approved software for data processing (calculation of ejection fractions, reformatting of angiographic data) should be available either as part of the MR system or on a separate workstation. Postprocessing performed by a technologist should be supervised by the cardiac MR physician.

Acknowledgments: Ad Hoc Committee on Cardiac Imaging: Jeffrey C. Weinreb, MD, chair; David A. Bluemke, MD, PhD; Andre J. Duerinckx, MD, PhD; N. Reed Dunnick, MD; Paul A. Larson, MD; Geoffrey D. Rubin, MD; Geoffrey G. Smith, MD; William Stanford, MD; Arthur E. Stillman, MD, PhD; Pamela K. Woodard, MD.

References

7. American College of Radiology. ACR practice guidelines for cardiac MR.


Image-guided Tumor Ablation: Standardization of Terminology and Reporting Criteria

The field of interventional oncology with use of image-guided tumor ablation requires standardization of terminology and reporting criteria to facilitate effective communication of ideas and appropriate comparison between treatments that use different technologies, such as chemical (ethanol or acetic acid) ablation, and thermal therapies, such as radiofrequency, laser, microwave, ultrasound, and cryoablation. This document provides a framework that will hopefully facilitate the clearest communication between investigators and will provide the greatest flexibility in comparison between the many new, exciting, and emerging technologies. An appropriate vehicle for reporting the various aspects of image-guided ablation therapy, including classification of therapies and procedure terms, appropriate descriptors of imaging guidance, and terminology to define imaging and pathologic findings, are outlined. Methods for standardizing the reporting of follow-up findings and complications and other important aspects that require attention when reporting clinical results are addressed. It is the group’s intention that adherence to the recommendations will facilitate achievement of the group’s main objective: improved precision and communication in this field that lead to more accurate comparison of technologies and results and, ultimately, to improved patient outcomes.

EDITOR’S NOTE: This report is an update of an article published previously in this Journal (Radiology 2003; 228:335–345).

Recently, the International Working Group on Image-Guided Tumor Ablation published a document entitled “Image-guided Tumor Ablation: Proposal for Standardization of Terms and Reporting Criteria” (1). The main objective was “improved precision and communication in this field that leads to more accurate comparison of technologies and results and ultimately to improved patient outcomes” (1). It was acknowledged by the members of the working group that the new field of image-guided tumor ablation (a branch of interventional oncology) required standardization of terminology and reporting criteria to facilitate effective communication of ideas and appropriate comparison between treatments that use different technologies. On the basis of this premise, a committee was established to author proposed standards, with the proposal unanimously adopted by the committee and ratified by the International Working Group on Image-Guided Tumor Ablation.

The initial goals of the working group’s proposal for standardization fall in line with the new initiative of the Society of Interventional Radiology (SIR), which promotes interventional oncology. Along these lines, a Technology Assessment Committee of the SIR has been charged with reviewing and commenting on the standardization of terminology and reporting criteria. Accordingly, the document has been modified in an attempt to align the contents with prior SIR standards and to address additional issues that have been raised by the Technology Assessment Committee. Additionally, we have attempted to respond to several recommendations of the Food and Drug Administration’s Center for Devices and Radiological Health in this version of the document. In essence, this independent review and ratification by the SIR Technology Assessment Committee of the prior report (1)
Image-guided Tumor Ablation

The term tumor ablation is defined as the direct application of chemical or thermal therapies to a specific focal tumor (or tumors) in an attempt to achieve eradication or substantial tumor destruction (2-6). The term “direct” aims to distinguish these therapies from others that are applied orally or via an intravascular or peripheral venous route. We stress the concept of image guidance in the title of our field to reflect our radiologic perspective and to highlight that image guidance is critical to the success of these therapies (2-6). Given that most of these therapies can be performed by using a host of imaging modalities (ie, ultrasonography [US], computed tomography [CT], magnetic resonance [MR] imaging, and fluoroscopy), the more general term image guidance is preferred, unless a particular imaging modality is mandated as part of the technique. However, virtually all available ablation techniques can theoretically be used with more than one modality.

While previously, some authors have referred to these procedures as “minimally invasive” or “percutaneous” therapies, these terms should be used only where appropriate. Minimally invasive therapies refer to all therapeutic procedures that are less invasive than conventional open surgery. All percutaneous procedures are therefore minimally invasive; however, not all minimally invasive therapies are performed or applied percutaneously. Indeed, the term “minimally invasive” is often used by surgeons to refer to procedures performed with minimal laparotomy or laparoscopy (7). Although less invasive than open surgery, these procedures are clearly more invasive than are percutaneous image-guided tumor ablation procedures. Inclusion of the term “percutaneous” as a prefix to “image-guided tumor ablation” is often too limiting because it does not reflect the fact that tumor ablation procedures can also be performed at laparoscopy, endoscopy, or surgery (8,9).

Individual procedures and therapies have often been given multiple different names by various investigators, which can potentially lead to confusion. Hence, we propose and recommend a unified approach to the terminology regarding these therapies. The primary aim of this classification is to provide simplicity and clarity, most notably by eliminating extraneous detail and many acronyms. The committee acknowledges that some acronyms (such as RF and RFA for RF ablation and HIFU for high intensity focused ultrasound) have gained widespread international acceptance. Nevertheless, niche application acronyms should be avoided.

When discrimination between the ablation of malignant versus nonmalignant tissue is needed, the descriptive term “ablation” should still be used, with the type of ablated tissue stated afterward (eg, acetic acid ablation of renal cell carcinoma or RF ablation of atherosclerotic plaque). In other words, the term “thermal (or laser, microwave, etc) ablation” should be used regardless of what is being ablated.

The methods of tumor ablation most commonly used in current practice are divided into two main categories: (a) chemical ablation and (b) thermal ablation. These categories require further definition and standardization of terminology as outlined below. Other interventional oncologic therapeutic approaches, including the percutaneous delivery of genetic material and radioactive seeds and the transcatheter delivery of chemoembolization agents (10,11), may ultimately require better definition and standardization of terminology as outlined below. Other interventional oncologic therapeutic approaches, including the percutaneous delivery of genetic material and radioactive seeds and the transcatheter delivery of chemoembolization agents (10,11), may ultimately require better definition and standardization of terminology as outlined below.

Energy Sources and Applicators

Although the devices are often referred to as “needles” or other nonspecific terms, they do not always conform to these precise classifications. Hence, the term applicator should be used generally to describe all devices. For precision, RF applicators are electrodes, microwave applicators are antennas, and laser applicators are fibers. On the basis of convention and consensus, cryoprobes are used to freeze tissue during cryoablation. For reporting completeness, a reference describing the appropriate applicator is encouraged as part of the report. The report describes a new prototype device, in which case an appropriate figure and/or schematic should be provided.

RF Ablation

This term applies to coagulation induction from all electromagnetic energy sources with frequencies less than 30 MHz, although most currently available devices function in the 375-500 kHz range (15). The term “radiofrequency” should be written as a single nonhyphenated word. Most devices currently used...
are monopolar in that there is a single “active” electrode, with current dissipated at a return grounding pad. Bipolar devices have two “active” electrode applicators, which are usually placed in proximity to achieve contiguous coagulation between the two electrodes (16). Additionally, many electrode modifications are now available, as classified below. The type of device and electrode used clearly influences the extent of ablation. Hence, clarity and standardization of terminology are required.

**Multitined expandable electrodes.**—This standard term refers to a family of electrodes that are currently available from several manufacturers (8,9,17–20). The usual embodiment of this type of device is an array of multiple electrode tines that expand from a single centrally positioned larger needle cannula. Currently, these are referred to as “umbrella electrodes,” “multitined electrodes,” “Christmas tree electrodes,” “multiple hooked electrodes,” or “arrays,” but this has led to confusion. Given the number of electrode types that have recently become available and the fact that several multitined devices are now available with variable deployment lengths, the exact electrode model and diameter of the electrode array used must be specified. Also, if a stepped deployment was performed with a multitined device, this too needs to be explained in detail regarding the length and time of deployment.

**Internally cooled electrodes.**—Some devices have a perfusate (such as saline or water) that flows in internal lumina that do not come in direct contact with patient tissues (21–23). These should be referred to as “internally cooled” (single or cluster [not “clustered”]) and not confused with perfusion electrodes, as described below. The term **cluster electrode** is most appropriate to describe internally cooled electrode devices in which three or more closely spaced (<1 cm) electrodes are used simultaneously to approximate an electrode with a larger diameter (24). Many investigators refer to these electrodes as “an array,” which may not adequately reflect the true underlying mechanism for enhanced energy deposition and ablation.

**Perfusion electrodes.**—Electrodes that have small apertures at the active tip that allow fluids (ie, normal or hypertonic saline [see “Adjuvant therapies” below]) to be infused or injected into the tissue before, during, or after the ablation procedure should be referred to as **perfusion electrodes** (25,26). The term replaces descriptions such as “cool-wet,” “wet,” or “saline-enhanced” electrodes.

**Algorithm of energy deposition.**—The methods used for applying energy have undergone continuous modification and improvement, which has led to substantial confusion and difficulty when the results of studies performed by different groups of investigators are compared. When reporting results, pulsing techniques and other methods for amplying energy deposition should be succinctly elaborated in the Materials and Methods section. Whenever possible, a reference for the precise algorithm used (eg, ramped energy deposition [18] or impedance regulated [27]) and the model number of the generator should be cited. Additionally, other parameters, including the use of monopolar or bipolar systems, the amount of energy applied (current [milliamperes] or power [watts]), and the duration of ablation should be provided. Adjuvant therapies.—Increased use of adjuvant therapies, such as concomitant percutaneous instillation of sodium chloride solutions to alter electric and thermal conductivity during ablation, are being reported with many variations in technique (28,29). Hence, specific details of the adjuvant used (ie, drug concentration, route and rate of administration, timing in relation to ablation therapy) must be provided. Whenever possible, a reference for the precise algorithm and the rationale for the selected adjuvant therapy used should be provided.

**Laser Ablation**

The term **laser ablation** should replace terms such as “laser interstitial tumor therapy,” “laser coagulation therapy,” and “laser interstitial photoocoagulation” (30–34). This term should be used for all types of ablation with light energy. Given multiple laser technologies and application methods, including superficial therapy (contact or noncontact mode) or transcutaneous ablation, the term “interstitial” or “direct” can be reported to clarify that laser energy is applied via fibers directly inserted into the tissue.

In addition to the laser source (eg, Nd: YAG, erbium, holmium) and precise wavelength, additional device characteristics must be specified, including (a) type of laser fiber (flexible or glass dome); (b) modifications to the tip (ie, flexible diffusor tip or scattering dome), with dimensions and materials specified; (c) length of applicator and diameter of the optic fiber; and (d) number of laser applicators used (ie, single vs multiple applicators). Similar to the reporting requirements for RF ablation, additional details of device modification, such as pulsing algorithms and internal cooling of the applicator, should be provided. The following technical parameters also should be provided: (a) laser power, reported as watts per centimeter of active length of laser applicator; (b) total duration of energy application; (c) total amount of energy applied per tumor (mean and range); and (d) sequential or simultaneous energy application to multiple fibers. For energy applied, in addition to the energy measured before the laser enters the fiber, ideally the actual energy output of the fiber or dome prior to the ablation and/or at the end of the procedure should be measured.

**Microwave Ablation**

This term should be used for all electromagnetic methods of inducing tumor destruction by using devices with frequencies from 30 MHz to 30 GHz (35–37). The term “microwave ablation” should replace the less succinct terms “percutaneous microwave coagulation therapy” and “microwave coagulation therapy.” Additionally, the precise frequency of the device and the type of applicator(s) should be provided.

**Ultrasound Ablation**

There are currently two methods for the application of ultrasound energy: extracorporeal (or transcutaneous) (38) and direct for percutaneous application with a needle-like applicator (39) and for intracavitary (and intracardiac) devices. Hence, the additional nomenclature of “extracorporeal” or “direct” is required prior to focused ultrasound ablation. The term “high intensity” (as commonly found in “high-intensity focused ultrasound”) is not essential because it is vague, imprecise, and implied by the proposed terminology.

**Cryoablation**

This term should be used to describe all methods of destroying tissue by means of the application of low-temperature freezing (40–45). The term “cryotherapy” is a suitable alternative because it has been used for many years to describe these methods, and it may also be useful when a literature search on this subject is conducted (44). The use of “cryo” as a free-standing term is to be avoided because “cryo” is a prefix and not a word. The archaic term “cryosurgery” is also to be
The freezing of tissue with rapid thawing leads to the disruption of cellular membranes and induces cell death (45). In the past, liquid nitrogen was placed directly on tissue, but with the exception of dermatologic applications, this method is no longer used. In the neck, chest, abdomen, pelvis, and extremities, cryoablation is performed by using a closed cryoprobe that is placed on or inside a tumor. In the two main types of systems, argon gas and either gas or liquid nitrogen are used. Temperatures are measured either at the tip of the cryoprobe or in the handle. In the past, temperature readings from cryoprobes have been a source of controversy because some devices of manufacturers measure the temperature of the coolant as it enters the distal probe tip and others measure at the probe tip itself. Hence, the temperatures at which cryoablation is performed should be specified. For publication purposes, the type of cryoablation system, the gases used, probe dimensions, and length and number of freeze-thaw cycles (active or passive thawing) should also be specified.

**Terminology for Describing the Effects of Blood Flow**

All of the thermal methods are negatively influenced by blood flow because it can potentially remove heat before complete tumor ablation is achieved (1–6). (This is also true in reverse for cryoablation, where the premature warming of tissue by blood can limit the effects of freezing on tissue.) The term heat sink effect refers to cooling by adjacent visible (>1-mm-diameter) blood vessels when ablated tissues are heated (46–48). In effect, the shape of the thermal zone of ablation is altered away from the blood vessel, and the overall ablation size is diminished due to removal of heat by flowing blood (46,47), or in the case of cryoablation, due to addition of heat. Although these phenomena serve to protect blood vessels and prevent bleeding from large vessels, they are also a major source of incomplete tumor ablation in many studies involving both thermal and cryoablation. Perfusion-mediated tissue cooling (or heating) is a more encompassing term that refers to both the effects of the larger heat sink vessels and the substantial effects of capillary level microperfusion (48). Several strategies have been developed to overcome this problem: pharmacologically decreased blood flow (49), temporary vascular balloon occlusion of a specific vessel during ablation (ie, hepatic artery, hepatic vein, and/or portal vein during intrahepatic ablation) (50), intraarterial embolization and chemoembolization (36,51,52), and a Pringle maneuver (ie, temporary hepatic arterial and portal venous occlusion by means of direct compression of the vessels) during RF ablation at laparotomy (9,47).

**IMAGE GUIDANCE**

While all procedures mentioned in this article refer to tumor ablations guided by imaging, it is important to understand what is meant by the term “image guidance.” First, “guidance” refers to procedures in which imaging techniques (eg, fluoroscopy, US, CT, and MR imaging) are used during the procedure. Imaging is used in five separate and distinct ways: planning, targeting, monitoring, controlling, and assessing treatment response (53). Treatments are planned before the procedure, and the assessment of treatment response occurs after the procedure is completed. Targeting, monitoring, and controlling are all performed during the procedure.

**Planning**

Imaging techniques, including US, CT, MR imaging, and more recently positron emission tomography (PET), are used to help determine whether patients are suitable candidates for these procedures. Imaging aspects that are particularly important include tumor size and shape, number, and location within the organ relative to blood vessels, as well as critical structures that might be at risk for injury during an ablative procedure. Modalities such as combined PET and CT and three-dimensional reconstructions of cross-sectional imaging data may be used more often in the planning of image-guided tumor ablations in the future.

**Targeting**

This term is used to describe the step during an ablation procedure that involves placement of an applicator (eg, an RF electrode or cryoprobe) into the tumor. While much of the current image-guided tumor ablation literature describes the use of techniques such as US and CT to target tumors for purposes of ablating them, targeting is only one aspect of intra procedural image guidance. Ideal qualities of a targeting technique include clear delineation of the tumor(s) and the surrounding anatomy, coupled with real-time imaging and multiplanar and interactive capabilities. For example, US (54) and some MR imaging (55,56) systems have all of these qualities.

**Monitoring**

Monitoring is the term that is used to describe the process with which therapy effects are viewed during a procedure. Changes in imaging that occur during a procedure can and should be used to determine treatment effects. Important aspects of monitoring include how well the tumor and/or target is being covered (ie, included and/or encompassed) by the ablation zone and whether any adjacent normal structures are being affected at the same time. Not all image-guided techniques provide the same degree and types of monitoring. For example, MR imaging is currently the only modality with well-validated techniques for real-time temperature monitoring (40,57–59). The term “monitoring” should not be used to describe response to treatment; for this, “treatment assessment” or “follow-up” is used.

**Controlling**

This term is used to describe the intra-procedural tools and techniques that are used to control the treatment. To control an image-guided ablation procedure, the treatment should be monitorable, such that the operator can utilize the image-based information obtained during monitoring to control it. This may simply be repositioning of a therapy applicator on the basis of physician experience, imaging findings, and thermal feedback, or it could be as sophisticated as an automated system that automatically terminates the ablation at a critical point in the procedure (60).

**Assessment of Treatment Response**

Imaging used to assess an image-guided tumor ablation procedure occurs after the procedure is completed and is discussed below as postprocedural imaging (2–6).
stressed by the appropriate selection of terminology. Although in many cases there is a good correlation or overlap between radiologic and pathologic findings, this is not invariably the case since over- and under-reporting of the true extent of disease has occurred (61,62). The classic example of this is the assumption that imaging findings (ie, the zone of abnormality on the image) are equivalent to the pathologic findings (ie, the true zone of tumor destruction and/or treatment effect), which may not always be the case. Hence, careful differentiation between imaging and pathologic findings must be made. This distinction is critical, given that the accuracy of assessment of the extent of tumor destruction by means of imaging findings is limited by the resolution of images and the uncertainty about the viability of cells at the radiographic margins of the zone of ablation.

Zone of Cell Death at Pathologic Examination

This should be referred to as coagulation or coagulative necrosis. Given that many tumors undergo central necrosis without ablation therapy, the term “coagulation” is preferred over the use of “necrosis” alone because it denotes that the ablation intervention is actively leading to tumor destruction. The more generalized term “coagulation” is preferred over the term “coagulative necrosis” because the latter term has a well-defined meaning in the pathology literature, including the absence of visible nuclei within the dead cells. In actuality, the zone of coagulation, while predominantly consisting of coagulative necrosis, often lacks the classic well-defined histologic appearance of coagulative necrosis in the acute postablation period or even within some zones of adequately ablated tissue for many months following ablation (22,61,63). Indeed, in many cases, specialized stains are required to confirm that cellular death has been achieved after thermal ablation (61).

Another important issue is definition of the zone of ablation at gross pathologic examination. Most thermal therapies induce a central “white zone” of coagulation, a pathologic finding that is generally accepted to represent coagulated tissue, surrounded by a variable “red zone” of hyperemia that is most often absent in ex vivo specimens (64). However, there has been controversy in the measurement; hence, comparison of the “true” size of induced zones of ablation is based on the fact that some authors have reported that this more peripheral red zone also represents ablated tissue, and they include it in their measurements. To avoid confusion, both measurements (white zone alone and white plus red zones) should be provided. At a minimum, the zones included in gross pathologic measurements should be specified.

Zone of Ablation at Postprocedural Imaging

Appropriate terminology must reflect the fact that although we rely on imaging to define the gross extent of induced coagulation, our accuracy is limited by both spatial and contrast resolution to approximately 2–3 mm (depending on the imaging modality) (61). Hence, in truth, postprocedural imaging findings are only a rough guide to the success of ablation therapy, since microscopic foci of residual disease, by definition, cannot be expected to be identified. The term “ablation zone” can be used to describe the radiologic region or zone of induced treatment effect (ie, the area of gross tumor destruction visualized at imaging). The term “lesion” is to be avoided, given the potential confusion about the intended meaning, since the term “lesion” has been used to refer to both the “ablation zone” and the underlying tumor to be ablated.

There are two types of imaging findings that are identified after an ablation procedure: those related to zones of decreased perfusion and those in which the signal intensity (at MR imaging), echogenicity (at US), or attenuation (at CT) are altered (1–6). Hence, the imaging strategy and the criteria used to define ablation must be specified. For contrast material–enhanced studies, it is important to recognize that in some organ sites, particularly the kidney, minimal contrast enhancement (ie, <20 HU for CT) seen soon after ablation can be identified in areas that are subsequently proved at pathologic examination to be uniformly dead tissue (65). (This finding is not well understood but may be due to pseudoenhancement, as has recently been described for renal cysts, or to represent true minimal enhancement from leaky capillaries at the treatment margin.) Other imaging findings also require precise definition.

Transient hyperechoic zone.—This is the preferred term to describe the transient (up to 30–90 minutes) zone of increased echogenicity seen at US within and surrounding a tumor during and immediately after RF ablation (66,67). Thereafter, treated tumors often develop mixed echogenicity on follow-up scans. This finding is believed to represent microbubbles of water vapor and other cellular products that form as a result of tissue vaporization during active heating and is most often used as a rough guide as to the extent of induced tumor destruction. However, it is not a precise marker, because both under- and overestimation of the true extent of coagulation have been reported. The term “transient hyperechoic zone” should replace imprecise terms such as “ultrasound cloud,” “ultrasound storm,” “outgassing,” and “microbubble vaporization.”

Ablative margin.—For many disease processes, particularly for tumors in the liver, the ablation of appropriate margins beyond the borders of the tumor is necessary to achieve complete tumor destruction. The term “ablative margin” is proposed to describe the 0.5–1.0-cm-wide region that should ideally be ablated in these cases (68). This term is preferable to “surgical margin” (because there is no surgery). It is important to stress that this extent of treatment is not always necessary or desired, particularly during attempts to destroy focal tumors in the kidney in patients with a tendency toward the development of multiple tumors (such as those with von Hippel-Lindau disease), where nephron-sparing and more limited ablation are desired to preserve renal function and avoid dialysis (65).

For normally vascular organs such as the kidney and liver, creation of an ablative margin results in zones of low attenuation and absent perfusion that extend into the parenchyma (2–6). Increased attenuation occurs in low-density tissues such as perinephric fat (for exophytic renal tumors) (65,69) and the lungs, where the term “ground-glass opacity” is used to describe the imaging findings in the treatment zone surrounding and including the ablated lung tumor (70).

Benign peritumoral enhancement.—This finding can be seen at both pathologic examination and contrast-enhanced imaging and typically suggests a benign physiologic response to thermal injury (initially, reactive hyperemia; subsequently, fibrosis and giant cell reaction) (61). Depending on the protocol used for contrast-enhanced imaging (injection rate and scanning delay), this transient finding can be seen immediately after ablation and can last for up to 6 months after ablation. This finding
usually manifests as a penumbra, or a thin rim peripheral to the zone of ablation, that can typically measure up to 5 mm acutely but most often measures 1–2 mm. It is a relatively concentric, symmetric, and uniform process with smooth inner margins, and it needs to be differentiated from “irregular peripheral enhancement.” The finding is most readily appreciated on the arterial phase CT scans, with persistent enhancement that is often seen on delayed MR images.

Irregular peripheral enhancement.—This term represents residual tumor that occurs at the treatment margin. In contrast to “benign periablational enhancement,” residual unablated tumor often grows in a scattered, nodular, or eccentric pattern. This sign indicates incomplete local treatment (ie, residual unablated tumor). As such, if they are not subject to further therapy, these foci tend to continue to grow. Given the delayed enhancement characteristics of many hypovascular tumors, this finding is often best appreciated in a comparison of portal venous or delayed images (3 or more minutes after contrast material injection) with baseline images.

Involution of coagulation.—The term “involution” should describe the process by which the body eliminates the zone of induced coagulation over time. The term “shrinkage” should be avoided as imprecise. The term “regression” is likewise to be avoided, given that it is commonly used in the medical oncology literature to describe involution of just the tumor itself, rather than the induced coagulation that often involves both the tumor and the surrounding tissues (ie, the ablative margin). It is important to note that no or minimal involution does not imply treatment failure.

Other imaging findings.—Many other imaging findings that represent both host reaction to ablation and repair mechanisms will undoubtedly be seen and reported. Such findings include inflammatory standing in the acute period after ablation and more chronic findings, such as fibrosis, scarring, and architectural distortion. In general, despite the tendency toward creative description, previously standardized radiologic nomenclature should be used to describe these findings whenever possible. The number of new terms to describe these processes should be minimized to wherever new descriptive terminology imparts prognostic value (eg, differentiating between benign peritumoral enhancement and residual unablated tumor).

Reporting of Tumor and Ablation Sizes

Appropriate uniform guidelines and standards are needed to report the extent of induced coagulation. In the past, comparisons between technologies have been made somewhat difficult because some authors report the largest diameter of induced coagulation, some report the average diameter, and some report the short-axis diameter. Additionally, coagulation has occasionally been reported as a volume of ablated tissue without any definition of dimensional measurements. Hence, uniform standards of comparison are essential and must be adopted.

Index tumor.—This is the preferred term to describe the initially identified tumor prior to ablation. This tumor should not be referred to as a “lesion” because this term could be confused with the zone of induced coagulation or the region of ablation at imaging.

Size classification of tumors.—Actual tumor sizes (mean ± standard deviation and range if applicable) should be reported. Given that appropriate ablation of adequate margins often represents the rate-limiting step for treatment effectiveness, the maximum diameter of the original tumor must be specified. However, many investigators perform analyses of their results on the basis of stratification of tumor sizes. In this regard, there is currently too much ambiguity and variability in the categorization of tumors by size. Investigators have reported upper limits of 2.0, 2.5, 3.0, and 5.0 cm in diameter for “small tumors” and 5 or 10 cm for “large tumors.” These differences have made the direct comparison of results with different technologies challenging. We, therefore, recommend that if such categorization is performed, the tumor size classification should be standardized according to the following scale: small tumors, diameter of 3 cm or less; intermediate tumors, diameter of 3–5 cm; and large tumors, diameter of more than 5 cm. This classification was determined as the most practical because it parallels the current technical capabilities and effectiveness for most image-guided ablation therapies.

Comparing zones of coagulation among different ablation techniques.—Often, the extent of induced coagulation is reported in experimental studies as a vehicle for comparing different ablation technologies and parameter modifications (71,72). The extent of induced coagulation should include the short-axis diameter, given that this parameter influences the overall extent of necrosis that can be achieved from a single application of energy and is likely to be an important factor that influences technical success in clinical practice. Hence, while additional parameters can certainly be provided and may be potentially useful, at a minimum this should be the standard that is reported to enable honest comparison between techniques. Of course, given that the ablation of a tumor is performed in three dimensions (ie, it is a volumetric problem), ideally, all three-dimensional measurements of the ablation zone and tumor should be provided, and less ideally, both measurements of the cross-sectional area should be provided. If volume is to be used as the only reported parameter, then a rationale must be specified. Average diameters should be accepted only if the tumor or zone of ablation is truly spherical, varying not more than 2–3 mm in cross-sectional diameter. It is further well known that many devices produce irregularly shaped zones of coagulation. Hence, the degree of uniformity or irregularity in the shape of the ablation zone should be specified.

It is important to stress that reliance on minimum and maximum sizes for the zone of ablation may not be useful for predicting clinical technical effectiveness because other technical factors are likely to be equally important. For instance, depending on the orientation of the energy applicator, a 1 × 2-cm tumor may be adequately treated with a 2 × 3-cm zone of ablation but not with a 3 × 2-cm zone of ablation. Ablation diameter or volume may also not tell the entire story. Although a 3 × 3-cm zone of coagulation may completely cover a 2-cm-diameter tumor when it is correctly positioned; if the zone is off the mark, the entire tumor will not be destroyed.

STANDARDIZATION OF FOLLOW-UP

Currently, definitions of the appropriate length of follow-up and the time points for technical success are not well established. One investigator’s long-term follow-up is often another’s short-term follow-up. Hence, specific guidelines need to be adhered to that depend on the type of disease treated and the intended goal of the study. Treatment study goals are generally related to one or more of the following four categories, which usually need to be distinguished from each another: (a) technical success, or, was the tumor treated according to the protocol? (b) technique effectiveness, or, was the
tumor effectively ablated? (c) morbidity, or, were critical structures and complications avoided? and (d) outcomes, or, was there some improvement in survival, quality of life, or palliation?

**Technical Success**

This term simply addresses whether the tumor was treated according to protocol and was covered completely. Tumor coverage can be assessed either during or immediately after the procedure. For example, MR imaging can be performed to monitor thermal injury and to show that the tumor is being covered completely during the procedure. Contrast-enhanced CT can be performed immediately after ablation. A tumor that is treated according to protocol and covered completely, as determined at the time of the procedure, is “technically successful.” The importance of this term is to help investigators separate those patients in whom the protocol could not be executed completely, for either technical reasons or reasons related to comorbid disease, from those that were treated according to protocol.

**Technique Effectiveness**

Distinction between “technical success” and “technique effectiveness” must be made. Effectiveness can only be demonstrated with appropriate clinical follow-up. “Technique effectiveness” should therefore refer to a prospectively defined time point (ie, immediately after the last course of a defined ablation protocol or at 1 week or 1 month after treatment), at which point “complete ablation” of macroscopic tumor as evidenced at imaging follow-up (or another specified end point) was achieved. The number of sessions (ie, the number of interventional procedures) to achieve the specified end point should likewise be defined. Authors are encouraged to report whether or not this complete ablation included an ablative margin.

Comparison of technical success and effectiveness between various ablation protocols has been challenging because many authors have adopted different terminology or guidelines. This problem is further compounded by our ability, and often the clinical need, to ablate a tumor over many sessions and the possibility of ablating growing foci of local tumor progression months after the initial course of therapy. A window of initial therapy for each ablation technique, during which it is reasonably expected that the tumor will be completely ablated, should be defined. For percutaneous thermal ablation, this window should ideally not exceed an upper limit of either one to four procedures or a specified time frame (up to 1–3 months), depending on the size, type, and location of the tumor, as well as the rationale for therapy. We have purposefully left definition of this end point as a broad range, given evolving consensus on defining more specific parameters because each disease process may vary. If complete ablation cannot be achieved within these specified parameters, the tumor should be classified as “unsuccessfully treated.”

**Primary and Secondary Technique Effectiveness Rates**

Given that multiple treatments with image-guided tumor ablation therapy are often given over the course of the disease, primary and secondary technique effectiveness rates should be reported. The primary effectiveness rate is defined as the percentage of tumors that were successfully eradicated following the initial procedure or a defined course of treatment. The secondary or assisted effectiveness rate includes tumors that have undergone successful repeat ablation following identification of local tumor progression. The term re-treatment should be reserved for describing ablation of locally progressive tumor in cases where complete ablation was initially thought to have been achieved on the basis of imaging findings that demonstrated “adequate” ablation of the tumor.

The technical success and technique effectiveness rates are very important as we define the limitations of our technologies, ideally in a manner similar to that used in other disciplines (ie, articles about surgical resection typically report a positive margin rate). Nevertheless, for some protocols, the concepts of local technical success and local tumor progression (ie, technique effectiveness) may have limited impact on the most important outcome parameter: patient survival. For example, use of three to four procedures or 1 month as the window of technique effectiveness may be of secondary importance if the patient lives for 5 years because of the treatment or if the tumor is completely eradicated over multiple courses of ablation therapy over many years.

**Complete Ablation versus Partial Ablation**

Many reports have surfaced in which different degrees of partial ablation have been reported (22,30,33,73,74). While consensus has been achieved for defining complete and incomplete ablation, there has been a rather arbitrary definition of incomplete ablation. For example, some authors have reported nearly complete ablation as representing greater than 90% induced necrosis, while others have used a threshold of 95% necrosis of the index tumor. Nevertheless, it is the opinion of the majority of the committee that this kind of classification of partial ablation is not warranted in an overwhelming majority of cases, given that adequate data are lacking to support a difference in outcome between different levels of partial ablation. Furthermore, such percentages are often estimates and may be inaccurate. Hence, at this time, such stratification should be avoided. It is important to stress that the elimination of this type of stratification does not negate the utility or imply the lack of benefit of tumor ablation as a palliative method. However, other end points should be chosen (see below) when reporting these cases on the basis of the rationale of palliation.

**Tumor Palliation**

The specified well-defined rationale for palliative therapy, as well as an appropriate method for assessing outcomes, must be provided. For example, if tumor ablation is valid as a vehicle for pain reduction, pre- and postprocedural pain scales should be obtained (75,76). If ablation is performed to reduce symptoms of a syndrome (such as carcinoid or other hormonally active or paraneoplastic tumors [77]), appropriate documentation of laboratory results from blood or urine obtained before and after therapy must be provided, and other symptom end points and grading systems must be specified and used. Needless to say, one cannot “palliate” asymptomatic tumors. Hence, the term debulking should be used to describe a procedure performed with the sole intent of inducing a reduction of tumor burden.

**Failure of Therapy**

*Causes of treatment failure.—The distinction between local incomplete therapy (tumor progression), new foci of disease in the target organ (especially the liver), and distant malignancy should be distinguished whenever possible and reported. Discrimination between “local tumor progression” and new tumor is important for determining the potential utility (ie, local treatment success rate) of*
a given method in the setting of many potentially confounding causes of the death of a given patient. Additionally, for patients with cirrhosis, the causes of mortality should be differentiated between hepatic disease and others.

Local tumor progression.—Many authors have used the term “local recurrence” to describe the appearance over follow-up of foci of untreated disease in tumors that were previously considered to be completely ablated. This is often a misnomer, given the fact that the tumor in essence did not recur but instead was never completely treated. Hence, the process often described is actually “residual unablated tumor.” However, in many cases, it is virtually impossible to determine whether there was incompletely treated viable tumor that continued to grow or if a new tumor (or in the case of hepatocellular carcinoma, “daughter” or “satellite” tumors) grew at the original site. Given this reality, local tumor progression is the preferred term over “local recurrence.”

Patient Mortality

Given that the population of patients that is treated most often are those with cancer, substantial patient mortality that is unrelated to the ablation intervention is anticipated, particularly in clinical studies with long-term follow-up. Therefore, the cause of death should be specified as “tumor related” or due to “other causes.” For tumor-related death, further subclassification (eg, differentiating death due to hepatic or diffuse metastatic burden), if possible, will often be useful because it can potentially shed further light on the effectiveness of therapy.

COMPLICATIONS

The unified standardized SIR grading system should be used as outlined here (78,79). Complications should be reported by using the SIR standard table so that they can be categorized consistently according to severity. The definition of death is self-explanatory and should be reported on a per-patient basis. Any patient death within 30 days of image-guided tumor ablation should be addressed (SIR classification F). The specific cause of death should be reported, with the potential and degree of causality to the ablation procedure clearly specified. Major and minor complications and side effects should be reported on the basis of the number of ablation sessions on a per-session basis. However, ideally, the number of ablations performed should be included because multiple ablations increase the likelihood of complications (80,81).

The definition of major complication is an event that leads to substantial morbidity and disability, increasing the level of care, or results in hospital admission or substantially lengthened hospital stay (SIR classifications C-E). This includes any case in which a blood transfusion or interventional drainage procedure is required. All other complications are considered minor. It is important to stress that several complications, such as pneumothorax or tumor seeding, can be either a major or minor complications, depending on severity. For tumor seeding, this would depend on whether the ec-topic tumor focus can be successfully ablated or otherwise treated.

Differentiation between immediate complications (up to 6–24 hours following the procedure), periprocedural complications (within 30 days), and delayed complications (more than 30 days after ablation) is advised. This stratification will give the reader an idea of when specific complications or side effects are most likely to occur and assist in defining when and how to take adequate precautions. Ablation-related complications should include problems encountered within the periprocedural (30 days) time period that can be related in any way to the procedure, as well as additional complications identified at delayed follow-up imaging that were judged to be highly likely due to the ablation therapy (eg, biliary ductal stricture, tumor seeding along the needle tract). Additionally, it should be specified which complications are being reported on a patient-by-patient basis (such as death) and whether the denominator represents the number of sessions or the number of tumors.

Side Effects

Side effects are expected undesired consequences of the procedure that, although occurring frequently, rarely if ever result in substantial morbidity. These include pain, the postablation syndrome, asymptomatic pleural effusions, and minimal asymptomatic perihilar (or renal) fluid or blood collections seen at imaging. Another such side effect would include asymptomatic imaging evidence of minimal thermal damage to adjacent structures without other evidence of negative sequelae (ie, “collateral damage”). An example of this would include when the zone of ablation extends beyond the liver capsule to include small portions of the diaphragm or kidney. These are not true complications because they do not lead to an unexpected increased level of care.

Pain

Even with appropriate conscious sedation techniques, patients may experience pain during ablation procedures. Additionally, depending on the organ site, many patients may experience grade 1–2 pain for several days, occasionally lasting 1–2 weeks following an ablation procedure. Last, thermal ablation, particularly RF, is being used with increased frequency as a method for treating refractory metastatic and primary bone tumor pain (75,76). We, therefore, propose adoption of the Common Toxicity Criteria of the National Cancer Institute for reporting pain (this document can be downloaded from the Web site step.cancer.gov/reporting/ctc.html) (82): grade 0, no pain; grade 1, mild pain that does not interfere with function; grade 2, moderate pain or pain or analgesics that interfere with function but not interfere with activities of daily living; grade 3, severe pain or pain or analgesics that severely interfere with activities of daily living; and grade 4, disabling pain.

Postablation Syndrome

This syndrome is a transient self-limiting symptom or sign complex of low-grade fever and general malaise (44,83). The duration depends on the volume of necrosis produced and the overall condition of the patient. If small areas are treated, the patient is unlikely to experience postablational syndrome at all. If very large areas of liver tumors are ablated, the syndrome may persist for 2–3 weeks. The majority of patients who have this syndrome will experience some malaise for 2–7 days depending on the volume of tumor and surrounding tissue ablated and the integrity of the patient’s immune system (ie, patients being treated with steroids or those who have small tumors may experience postablational syndrome).

Follow-up and Outcomes

Outcomes of interest may include local response, systemic response (pain, cancer syndromes, etc), quality of life, or survival. For those studies that deal with the quality of life, some form of objective measurement must be used both before and after treatment. Ideally, previously validated scales or metrics should be used and appropriately referenced.
Imaging follow-up.—Currently, despite a reliance on imaging findings to determine the extent of “unablated residual tumor,” there is a lack of consensus on a standard follow-up interval regimen for imaging. The most common approach taken by members of the Working Group includes contrast-enhanced CT or MR imaging within 6 weeks of the initial ablation to determine whether additional ablation therapy is required (many centers perform this examination on the day of the initial procedure) and thereafter every 3–4 months to determine technical effectiveness. Imaging intervals may also vary depending on the type of underlying tumor and the goals of treatment. At a minimum, the intervals at which follow-up imaging was performed should be clearly specified.

Although standard imaging criteria for response assessments have been defined for evaluation of other cancer therapies, these criteria focus almost exclusively on tumor size (84). However, given the heavy reliance on morphologic features other than size in the assessment of results of ablation therapy, exclusive reliance on tumor size does not provide a complete imaging assessment of tumor response and may even lead to erroneous conclusions about the effectiveness of the therapy (85). Therefore, in addition to reporting index tumor and the zone of ablation diameters, assessment of tumor enhancement or lack thereof should also be included in the imaging response assessment following ablation therapy.

Length of follow-up.—Currently, many, if not most, published series do not provide enough technical detail to permit duplication of the investigators’ efforts. This problem is compounded by the fact that there are many different types of ablation equipment on the market and in development and these often change. Hence, the specification of the parameters such as duration of energy applied and manufacturer must be provided. Also, the number of treatment sessions for each tumor should be specified. The procedure approach (ie, whether the procedure was performed percutaneously, laparoscopically, or endoscopically) should also be clearly specified. Additional parameters to be provided for publication should include the following: (a) whether the procedure is performed with general anesthesia or conscious sedation (the specifics of anesthetics and medications administered during the procedure and in the recovery phase should be always reported, including agent, dose, and route), (b) the types of imaging guidance (CT, CT fluoroscopy, US, or MR imaging), (c) whether the patient was hospitalized, (d) the number of sessions required to initially achieve technical success, and (e) the subsequent rates of other tumors requiring additional ablation therapy. Last, any repositioning of the applicator during the ablation and the procedure for applicator removal (ie, use of fiber enclosure or other closure devices) should be noted.

Other Study Population Data to Be Reported

The study population should be rigorously described, including inclusion and exclusion criteria and tumor type and size. The degree of proof of disease required for entry into the study (ie, biopsy, imaging, or serologic criteria) should be clearly specified. Pretreatment evaluation also needs to be reported. In addition to an appropriate focus on anatomy (ie, the organ, tumor size, location, and number), the pretreatment evaluation should also include tumor stage (ie, spread elsewhere), patient comorbidities, age, sex, and overall clinical debility, because outcomes such as mortality will depend on these factors. Obviously, a debilitated cachectic patient with widespread metastases will have a worse outcome following liver RF ablation than will an otherwise well patient.

Findings of a recent study (89) have also suggested the potential complementary effects of chemotherapy and radiation therapy on ablation effectiveness. Hence, the administration of either of these therapies to patients enrolled in clinical trials of ablation should be specified. This should be further classified as having received the conventional oncologic therapies previously, around the time of ablation (within 1 month) or during the follow-up period. The specific therapy protocol and the duration of therapy in relation to the ablation therapy should also be provided.

Accurate and Complete Delineation of Ablation Procedures

Substantial confusion and difficulty in comparing results have arisen regarding the success and complication rates because patients may have had one or more tumors treated over multiple procedure sessions. Ideally, all four parameters (numbers of patients, tumors, treatment sessions, and ablation procedures) should be reported whenever possible. Additionally, results are often reported for heterogeneous populations of patients for which varied rationales for the procedure (cure vs palliation) or outcomes (hepatic metastases vs hepatocellular carcinoma) have been reported. Therefore, stratification of patients into appropriate categories is advised to avoid confusion and best facilitate extraction of clinically meaningful conclusions.

Minimizing Technical Jargon

Although substantial technical jargon and marketing terminology appear in the peer-reviewed medical literature, these should not be used. For example, colloquial phrasing such as “lesioning” and “burning” are to be avoided when describing the application of thermal energy. Another example is the concept of “roll off” to describe the impedance control algorithm of a device of one particular manufacturer; this term should not be used.
Comparison with Other Treatments

Given that most reports of image-guided therapy have been relatively small case series, a major benefit of uniform reporting standards is the ability to perform meta-analyses of outcomes to compare therapies (90). Clinical research studies should be reported in such a manner that the results can be directly compared to various cancer therapies, including other forms of image-guided ablation, surgery, radiation therapy, and chemotherapy. The coin of the realm in oncology is survival, disease-free survival, and quality of life stratified according to disease stage and patient functional status (91,92). Nevertheless, there are limited data addressing these issues for most diseases treated with image-guided ablation (93). Thus, the committee wishes to stress the need for studies on organ-by-organ and disease-by-disease bases. Randomized, controlled, and blinded studies are considered the standard for pivotal studies and should be performed when possible (94–97). By the same token, the committee acknowledges both the very real obstacles to performing such studies, (patient recruitment, long periods of data collection, expense, multicenter organization, etc.) and the benefit of reporting less robust forms of data, including retrospective studies, case series, and case reports (94,98).

Statistical Evaluation

Regardless of the study type, rigorous statistical evaluation appropriate for the data collected should be presented (95,97). The primary and secondary study end points should be clearly stated. By bearing in mind that the data from individual studies may need to be treated differently, in general survival outcomes should be reported by using life-table (Kaplan-Meier) analysis. Patients should be randomized, if possible, and results should be reported on the basis of the intention to treat, whether patients were treated as randomized and whether they were treated per protocol (ie, excluding protocol violations). Outcomes may further need to be stratified according to multiple factors (tumor type, grade, and stage; functional status; comorbidities; etc). Appropriate methods for assessment of quality of life should likewise be selected (99).

CONCLUSIONS

The intent of this proposal for standardization of terminology is to provide an appropriate vehicle for reporting the various aspects of image-guided ablation therapy. Our intent is to provide such a framework to facilitate the clearest communication between investigators and the greatest flexibility in comparison between the many new, exciting, and emerging technologies. Clearly, this is an ongoing process that will require modification as our understanding of these technologies improves, new treatment paradigms emerge, and greater consensus is achieved on standardizing the reporting of currently unresolved issues. Indeed, we welcome constructive feedback from the medical community at large in an attempt to further refine this proposal. Nevertheless, we encourage all of our colleagues to adopt the terminology and reporting strategies outlined in this proposal.

Acknowledgments: The complete list of authors is S. Nahum Goldberg, MD (coordinated the publication of the initial document and served as topic leader during the subsequent revision of the draft), Clement J. Grassi, MD (chair of the SIR Technology Assessment Committee), John F. Cardella, MD, J. William Charboneau, MD, Gerald D. Dodd III, MD, Damjan E. Dupuy, MD, Debra Gervais, MD, Alice R. Gillams, MD, Robert A. Kane, MD, Fred T. Lee, Jr, MD, Tito Livraghi, MD, John McGahan, MD, David A. Phillips, MD, Hyunchul Rhim, MD, Stuart G. Silverman, MD, Luigi Solbiati, MD, Thomas J. Vogel, MD, Bradford J. Wood, MD, Suresh Vedantham, MD, and David Sacks, MD (is SIR Standards Division Councilor). Other members of the Technology Assessment Committee who participated in the development of this clinical practice guideline are John Dean Barr, MD, Gary J. Becker, MD, John J. Borsa, MD, Patricia E. Cole, PhD, MD, Ram VSR Chavali, MD, William B. Crenshaw, MD, Michael D. Dake, MD, Chieh-Min Fan, MD, Ziv Haskal, MD, Randall T. Higashida, MD, Thomas B. Kinney, MD, Sanjoy Kundu, MD, Lindsay Machan, MD, Louis G. Martin, MD, Steve F. Millward, MD, Niles H. Patel, MD, Dheeraj Rajan, MD, James E. Silberzweig, MD, Curtis W. Bakal, MD, Curtis A. Lewis, MD, and Kenneth S. Rholl, MD. The authors also wish to acknowledge other members of the International Working Group on Image-Guided Tumor Ablation who ratified the initial proposal.

Author affiliations: Department of Radiology, Beth Israel Deaconess Medical Center, Boston, Mass (S.N.G., R.A.K.); Department of Radiology, Lahey Clinic Medical Center, Burlington, Mass (C.J.G.); Department of Radiology, University of Colorado Health Sciences Center, Denver, Colo (J.F.C.); Mayo Clinic, Rochester, Minn (J.W.C.); Department of Radiology, University of Texas Health Science Center, San Antonio, Tex (G.D.D.); Division of Ultrasound, Brown University-Rhode Island Hospital, Providence, RI (D.D.); Department of Radiology, GI/CU Division, Massachusetts General Hospital, Boston, Mass (D.A.G.); The Middlesex Hospital, London, United Kingdom (A.R.G.); Clinical Science Center, University of Wisconsin, Madison (F.T.L.); Department of Radiology, Ospedale Civile-Vimercate, Milan, Italy (T.L.); U.C. Davis Medical Center, Sacramento, Calif (J.M.); Department of Radiology, University of Massachusetts Memorial Health Care, Worcester, Mass (D.A.P.); Hanyang University Hospital, Seoul, Korea (H.R.); Department of Radiology, Harvard Medical School, Brigham and Women’s Hospital, Boston, Mass (S.V.); Department of Radiology, General Hospitals of Busto Arsizio, Busto Arsizio, Italy (L.S.); Institute of Diagnostic/Interventional Radiology, Johann Wolfgang Goethe University Hospital, Frankfurt, Germany (T.J.V.); Radiology Department, NIH Clinical Center, National Institutes of Health, Bethesda, Md (B.J.W.); Mallinckrodt Institute of Radiology, Washington University School of Medicine, St Louis, Mo (S.V.); and The Reading Hospital and Medical Center, Reading, Pa (D.S.).

References


PACS and Imaging Informatics: Basic Principles and Applications
H. K. Huang, DSc

The picture archiving and communication system (PACS) originated as an image management system for improving the efficiency of radiologic practices. It has evolved into a hospital-integrated system that stores information media in many forms, including voice, text, medical records, waveform images, and video recordings. The integration of these various types of information requires the technology of multimedia, including hardware platforms, information systems and databases, communication protocols, display technology, and system interfacing and integration. The purpose of this book is to discuss the conceptual and technologic advances in PACS development for all of these previously mentioned technologies. The book also emphasizes the growing field of PACS-based imaging informatics, where existing PACS resources are used for large-scale horizontal and longitudinal clinical service, research, and education applications.

The book is intended for those who are involved in a health care environment that manages the everyday practical realities of PACS planning, operation, and maintenance. The book will also serve as excellent curriculum material for the training of PACS information technology personnel because the author provides comprehensive concepts that tie medical imaging, medical physics, information technology, and clinical workflow with enterprise PACS (multiple, hospital-wide PACS). The author, H. K. Huang, has been involved in PACS development since the 1980s and is a leading scientist in the field.

The book is divided into five parts and 23 chapters. Part 1, titled “Medical Imaging Principles,” summarizes the history of PACS, the fundamental concepts of digital medical imaging, and the different types of modalities integrated into PACS. Part 2, titled “PACS Fundamentals,” describes in detail the PACS subsystems and industry standards (Digital Imaging and Communications in Medicine and Health Level 7) that are used to integrate PACS with the hospital information system and radiology information system. Part 3, titled “PACS Operation,” describes the practical problems encountered by PACS administrators and users and discusses different techniques to solve these problems, with many examples from installed PACS sites around the world. Because the author developed PACS at the University of California, Los Angeles, in 1991 and the University of California, San Francisco, in 1995, he is able to draw from his expertise on the subject. The author also uses his experience from developing many state-of-the-art prototypes (eg, fault tolerant storage system) funded by research grants.

Part 4, titled “PACS-based Imaging Informatics,” presents image-processing techniques that can be used to mine the huge image and clinical data resources stored in the PACS archive. The analysis of lung nodules by using temporal computed tomographic and content-based image indexing are some of the examples of advanced processing applications applied to PACS data mining. Part 5, titled “Enterprise PACS,” describes different models for implementing an enterprise PACS in terms of the architectural design of the system and financial and business models.

This new book builds on two previous volumes on PACS that were written by the author. The text, image collections, and schematic diagrams in this current work achieve the author’s goal of providing up-to-date material concerning the technologic advances in PACS and imaging informatics. The text serves as a guide for those who are planning an enterprise PACS and provides teaching material for training PACS developers in industry and health care. The author has used portions of this book as lecture material for graduate courses at leading universities in the United States and abroad. The author also uses many examples to emphasize that the integration of heterogeneous health care information systems is becoming more complex and that, as more radiologists and clinicians expect higher standards in the reliability of the data, a systemic understanding of Digital Imaging and Communications in Medicine, Health Level 7, and Integrating the Healthcare Enterprise becomes ever more important to optimize enterprise workflow.

The strength of the book lies in the vast experience that the author, who has implemented PACS at numerous institutions in the United States and abroad, has transmitted into this new book. Huang has conducted first-hand analysis of the various systems from PACS manufacturers that are used in clinical operation at large medical institutions and has performed up-to-date research in data mining of the huge image and data resources stored in the PACS repository.

Reviewed by Duk-Woo Ro, PhD
Imaging and Cancer: Research Strategy of the American College of Radiology Imaging Network

The American College of Radiology Imaging Network (ACRIN) is a cooperative group funded by the National Cancer Institute and dedicated to developing and conducting clinical trials of diagnostic imaging and image-guided treatment technologies. ACRIN’s six disease site committees are responsible for developing scientific strategies and resultant trials within the framework of ACRIN’s five key hypotheses: (a) Screening and early detection with imaging can reduce cancer-specific mortality. (b) Less invasive image-guided therapeutic methods can reduce the mortality and morbidity associated with treating cancer. (c) Molecular-based physiologic and functional imaging can improve the diagnosis and staging of cancer, thus improving treatment. (d) Functional imaging can portray the effectiveness of treatment earlier and more accurately, thus reducing mortality and improving the likelihood of a cure. (e) Informatics and other “smart systems” can improve the evaluation of patients with cancer, thus leading to better and more effective treatments. This article details ACRIN’s research strategy according to disease site through the year 2007.

© RSNA, 2005

Denise R. Aberle, MD  
Caroline Chiles, MD  
Constanine Gatsonis, PhD  
Bruce J. Hillman, MD  
C. Daniel Johnson, MD  
Bruce L. McClennan, MD  
Donald G. Mitchell, MD  
Etta D. Pisano, MD  
Mitchell D. Schnall, MD, PhD

A. Gregory Sorensen, MD

Published online  
10.1148/radiol.2353041760  
Radiology 2005; 235:741–751

Abbreviations:  
ACRIN = American College of Radiology Imaging Network  
DCIS = ductal carcinoma in situ  
FLT = 3-deoxy-3-fluorine 18-fluorothymidine  
FDG = 2-deoxy-2-fluorine 18-fluoro-D-glucose  
RF = radiofrequency  
SUV = standardized uptake value

1 From the Departments of Radiology of the University of California, Los Angeles (D.R.A.); Wake Forest University School of Medicine, Winston-Salem, NC (C.C.); the Brown Center for Statistical Sciences, Brown University, Providence, RI (C.G.); University of Virginia Health System, PO Box 800170, Charlottesville, VA 22908 (B.J.H.); Mayo Medical School, Rochester, Minn (C.D.J.); Yale University School of Medicine, New Haven, Conn (B.L.M.); Thomas Jefferson University School of Medicine, Philadelphia, Pa (D.G.M.); University of North Carolina School of Medicine, Chapel Hill, NC (E.D.P.); University of Pennsylvania School of Medicine, Philadelphia, Pa (M.D.S.); and Harvard University School of Medicine, Cambridge, Mass (A.G.S.). Received October 13, 2004; accepted October 27. Address correspondence to B.J.H. (e-mail: bjh8a@virginia.edu).  
© RSNA, 2005
is to survey opportunities and to bring those opportunities to the disease site committees through the membership of modality committee constituents on disease site committees. This cross-fertilization is intended to be an efficient mechanism for the exchange of information concerning innovative technologies that will lead to the development of important clinical trials.

The six ACRIN disease site committees are responsible for the generation of scientific strategy for the organization within the context of ACRIN’s overarching research strategy. Between 2004 and 2007, this overarching strategy will focus on the following five key hypotheses: (a) Screening and early detection with imaging can reduce cancer-specific mortality. (b) Less invasive image-guided therapeutic methods can reduce mortality and the morbidity associated with treating cancer. (c) Molecular-based physiologic and functional imaging can improve the diagnosis and staging of cancer and thus improve treatment. (d) Functional imaging can portray the effectiveness of treatment earlier and more accurately, thus reducing morbidity and improving the likelihood of a cure. (e) Informatics and other “smart systems” can improve the evaluation of cancer, leading to better and more effective treatments.

The individual disease site committee strategies make up the comprehensive ACRIN research agenda, which we believe also circumscribes the agenda for imaging and cancer. The pursuit of clinical trials based on these strategies is expected to provide information that will further ACRIN’s principal goal of improving the length and quality of life in patients with cancer.

BREAST CANCER

The key clinical questions for breast cancer that should be addressed in imaging clinical trials are as follows: What methods can be used to improve the early diagnosis of breast cancer in special populations, particularly in young women and women with radiographically dense breasts? Can breast cancer screening be improved for these and other populations through adaptation of imaging diagnosis and treatment to patient risk and, if so, through the use of which imaging tools? What imaging tools can be used to help distinguish aggressive breast cancers from nonaggressive breast cancers, thus allowing therapies to be adapted to the tumor characteristics?

Can we adapt therapy to tumor aggressiveness and use less aggressive (ie, percutaneous or transcervaneous) therapies for less aggressive malignancies? What imaging tools can be used to better assess the success of therapeutic regimens for breast cancer so that therapies can be adjusted when tumors are not responding?

RATIONAL AND HOW IMAGING MIGHT BE USED TO ADDRESS THESE QUESTIONS

Adaptation of Diagnosis and Treatment to Patient Risk

The mainstay of breast cancer screening is still screen-film mammography. Its value as a test that helps to save lives by allowing early detection of breast cancer has been repeatedly demonstrated through seven screening trials in four countries over the past 40 years (1). Unfortunately, 40 000 women still die of breast cancer in the United States every year (2). The sensitivity of mammography is estimated to range from 83% to 95%, and the specificity is estimated to range from 90% to 98% (3). Sensitivity is lessened in women with dense breasts, with one author estimating it to be as low as 48% (4).

As a result, conditions are not diagnosed with screening in many women with breast cancer; instead, conditions are diagnosed with palpation of a finding at physical examination or self-examination. In addition, there is a high probability of false-positive findings at screening in up to 10% of women (3), and it has been suggested that there is a 50% chance of an abnormal examination in women who undergo screening for 10 years (5).

The current mammographic screening guidelines in the United States are uniform across the entire population of women older than 40 years. Women at high risk for breast cancer might benefit from enhanced screening imaging tools. As an example, there is obvious need for improvement in breast cancer detection, both in increasing sensitivity in women with dense breasts and in increasing specificity. The ACRIN Digital Mammographic Imaging Screening Trial (ACRIN 6652), which is currently concluding patient follow-up and conducting reader studies, will enable researchers to determine whether digital mammography offers such an advance over traditional screen-film mammography. These concepts are also currently being tested in women at high risk for breast cancer in the ACRIN trials of screening breast magnetic resonance (MR) imaging (ACRIN 6667) and breast ultrasonography (US) (ACRIN 6666). Other potentially useful screening tools include genetic tests and serum proteomic markers, which might be combined with imaging screening to enhance accuracy and efficiency.

With the guidance of these trial results, more individualized screening tools and recommendations are potentially available within the next 5 years. Clearly, the development of those tools and guidelines should be a high priority for the ACRIN Breast Committee. Potentially useful technologies on the horizon that might be studied as part of ACRIN trials are digital mammography with tomography and three-dimensional reconstruction and dual-energy or digital subtraction mammography before and after the administration of intravenous contrast agents. These ACRIN trials will generate a vast archive of images that will be made available to academic and commercial entities for the development and testing of computer-aided detection and diagnosis.

Differentiation between More and Less Aggressive Breast Cancers

Ductal carcinoma in situ (DCIS) is not a cause of death per se. Breast cancer can only cause death in a patient after it has invaded and spread to other organs. DCIS is a substantial risk factor for the development of invasive disease, however, and its presence may usually precede the development of invasive tumor. However, the presence of DCIS does not indicate that an invasive tumor will inevitably develop (6). The diagnosis of DCIS has increased approximately tenfold in the United States and other developed countries since the introduction of population-based screening mammography. This suggests that some cancers that would have never injured a patient are being diagnosed and treated, thus causing unneeded morbidity and mortality (7).

DCIS is currently treated with either mastectomy or breast-conserving surgery and accompanied by radiation therapy and hormone therapy with tamoxifen. The recurrence rate is approximately 30% without radiation therapy and approximately 15% with it; there is no difference in mortality between the two therapies when compared with each other and when compared with mastectomy (2%). Recurrences tend to be 50% in patients with DCIS and 50% in patients with invasive breast cancers. As a result, some have argued that DCIS is not overtreated,
that the increasing incidence of DCIS reflects more careful screening of the population, and that eventually this will yield lower breast cancer mortalities in the screened populations. Further evidence to support this viewpoint is found in the form of higher than expected numbers of invasive breast cancers in women with DCIS, increased risk of DCIS and invasive breast cancer in women treated for DCIS, and the results of genetic studies in women with DCIS and invasive breast cancer (8).

Ideally, however, there would be some mechanism for clinicians to determine which DCIS lesions are likely to lead to or become associated with invasive breast cancer and the death of the patient so that the more aggressive tumors can be treated aggressively and those that are not likely to kill the patient can be left alone or treated less aggressively. In general, high-grade DCIS is more likely to be associated with an invasive tumor. While there are some imaging features that loosely correlate with the grade of DCIS that is subsequently diagnosed, these findings are not very sensitive or specific. It is important for breast imagers to evaluate molecular imaging tests that, together with serum markers, might allow the aggressiveness of mammographically apparent lesions to be determined in vivo, perhaps precluding biopsy and even therapy for more indolent cases. One such potential tool is breast MR spectroscopy for tumor choline levels.

In addition, there are multiple tools that might be useful for the less invasive treatment of small early breast cancers, possibly including DCIS, should they be determined to be effective. These currently include radiofrequency (RF) ablation, microwave phased-array ablation, high-intensity focused US, cryoablation, and laser therapy. The role of imaging in administration of the therapies and caring for the patient after therapy has been completed is yet to be determined. ACRIN trials will corroborate the effectiveness of less invasive image-guided therapies with current standard therapies to elucidate which women might not need open surgery.

Evaluation of the Effectiveness of Treatment

In ACRIN’s trial of dynamic contrast agent–enhanced MR imaging as an intermediate marker for therapeutic effectiveness (ACRIN 6657), researchers will evaluate a possible role for breast MR imaging in determining whether novel chemotherapeutic regimens are working well in individual patients, particularly in the treatment of advanced breast cancer (stage 3 or higher). Such patients are at high risk of dying. If dynamic contrast-enhanced MR imaging proves successful in enabling physicians to predict which patients are responding to treatment, the effect on patients will be profound, as this technique will enable the early stoppage of ineffective treatment and the more rapid initiation of alternative therapy.

The use of MR imaging as a tool to monitor the efficacy of chemotherapy is really in its infancy. Patients who respond to chemotherapy may have typical contrast enhancement patterns over time, and the diagnostic criteria for interpreting MR imaging findings in the setting of chemotherapy may differ substantially from that used prior to chemotherapy. Other imaging tools that might be useful in the evaluation of the effects of chemotherapy over time in patients at high risk of death from breast cancer if therapy is unsuccessful—and that might be used in ACRIN trials—are MR spectroscopy, positron emission tomography (PET) with novel agents, US, and digital subtraction mammography.

GASTROINTESTINAL CANCERS

The Gastrointestinal Committee chose the following as its research focuses: (a) methods to screen for colorectal polyps and cancer, (b) development of computer-aided detection tools for colorectal polyp and cancer detection, (c) assessment of patients with chronic liver disease to discriminate dysplastic regenerative nodules from hepatocellular carcinoma, (d) development of imaging tools or techniques to improve discrimination between benign serous pancreatic tumors from the pre-malignant or malignant mucinous tumors, (e) development of better methods to gauge response to therapy in patients with liver metastases, (f) banking of tissue and serum to correlate biomarker findings with imaging studies, and (g) percutaneous ablation of cancer for local control and reduced mortality. The Gastrointestinal Committee has the intent of developing clinical trials around these focuses for liver metastases and colorectal, hepatocellular, and pancreatic carcinomas.

RATIONALS AND PROPOSED IMAGING TRIALS

Colorectal Cancer

Colorectal cancer causes substantial morbidity and mortality, especially in industrialized nations. It is the third most common cancer and second leading cause of malignant death in the United States, with an estimated 134,000 new colorectal cancer cases and 55,000 colorectal cancer deaths each year (2). As the natural history of colorectal cancer permits the recognition and curative treatment of both precursor adenomas and localized cancers, there is an enormous opportunity to save lives with early detection programs broadly applied to a general population. However, the potential efficacy and practicality of such a screening effort are compromised by limitations in the performance, comfort, and expense of available screening tests. Better tools are needed to more effectively screen for colorectal neoplasia.

Computed tomographic (CT) colonography (virtual colonoscopy), a noninvasive technique requiring only bowel preparation, is a structural examination of the entire colorectum that uses volumetric data acquired from a CT scanner combined with advanced computer software for image display. Controversy exists as to its sensitivity and specificity in the detection of large adenomas (especially in a screening population), as some authors have reported performance comparable to that of conventional colonoscopy (9), while others have reported only moderate sensitivity (10,11). A strong need exists to clinically validate widespread use of CT colonography in a screening population for the detection of colorectal neoplasia. In addition, the cost-effectiveness implications of observed performance outcomes need to be addressed. These needs are accommodated by ACRIN 6664—which serves to evaluate CT colonography in a screening population at high risk for colon cancer. Accrual of subjects began in February 2005.

In addition, all data from the trial will be archived and made available for academic and commercial concerns to test computer-aided detection programs aimed at improving observer performance and reducing variability.

Hepatocellular Carcinoma

Hepatocellular carcinoma is the most common primary hepatic tumor and one of the most common malignancies world
wide. This devastating disease accounts for 17,300 new cases of cancer and 14,400 deaths yearly (2). It is believed that regenerative nodules in the liver, which develop as a result of cirrhosis, undergo transformation over the course of many years into dysplastic nodules and, ultimately, carcinoma. During the transformation between a dysplastic nodule and carcinoma, there are changes in the blood supply within these nodules. As the benign nodule changes into a malignancy, the major blood supply also changes from portal venous to arterial (12–14).

The natural history of this tumor leads many investigators to believe that there is high potential to use imaging to screen, detect, and treat many of these premalignant lesions or early cancers before metastases have developed. Imaging techniques that hold promise include MR imaging, CT, and US—especially with the use of novel intravenous contrast agents—for assessment of lesion perfusion. Radionuclide scanning with innovative targeted contrast agents holds the promise of indicating the metabolic or physiologic nature of abnormalities, which may allow better differentiation of benign hyperplastic nodules from malignancy. ACRIN trials that involve obtaining tissue specimens will allow correlation of imaging findings with molecular expression. This may allow identification of hepatocellular carcinoma subtypes, thus opening the door to image-driven customized therapeutic approaches.

There is also a role for ACRIN in the investigation of less invasive treatment for hepatocellular carcinoma. Treatments of hepatocellular carcinoma with demonstrated efficacy include surgical resection and liver transplantation if lesions are small in number and size. For patients who do not fulfill criteria for resection or transplantation, however, treatment options are poor. RF ablation provides a potentially useful treatment option that may be superior to other available nonsurgical treatments (15,16). In addition to the possibility of prolonging the survival of patients who are not surgical candidates, RF ablation can be used in transplantation candidates to treat small hepatocellular carcinomas that arise while a patient is awaiting transplantation.

Cystadenocarcinoma of the Pancreas

Mucous cystadenocarcinomas are not reliably differentiated from benign serous tumors by their morphologic appearance (17,18). As a result, patients with asymptomatic benign neoplasms may undergo unnecessary surgery. However, serous tumors contain abundant low-viscosity glycoprotein-rich fluid, while mucinous tumors contain high-viscosity mucinous fluid. ACRIN trials of imaging methods that could be used to distinguish these tumors according to their fluid differences might prove to be effective diagnostic tools. New imaging methods, such as diffusion MR imaging and MR spectroscopy, hold promise in this regard and are ripe for investigation.

Assessing the Therapeutic Response of Liver Metastases

Liver metastases are the most common malignant neoplasm to affect the liver. Despite great advances in imaging technology that allow physicians to directly visualize liver metastases and increase the ability of physicians to detect these lesions with ever-increasing precision, our ability to measure the effect of treatment on these lesions remains limited.

The Gastrointestinal Committee proposes trials in which researchers will address the need to better assess the volume of residual tumor in the liver (19,20) and the functionality of the remaining tumor after therapy (21,22). New CT and MR imaging technology allows acquisition of three-dimensional volumes of the liver, which may represent an improvement on the traditionally used Response Evaluation Criteria in Solid Tumors. Functional assessment of tumors is also possible with assessment of contrast enhancement within lesions and with the use of PET scanning. Potentially better ways to determine viable tumor within detectable lesions include the use of targeted molecular agents, perhaps with fusion technologies, such as combined PET/CT. In addition, the committee proposes that ACRIN collaborate with therapeutic cooperative groups to test the potential of new molecular imaging technologies to serve as intermediate markers of the effectiveness of treatment, particularly with regard to predicting pathologic response and extended survival.

Surgical resection is the only treatment for colorectal metastases that has been shown to improve 5-year survival. As with hepatocellular carcinoma, treatment options for patients who are not candidates for surgery are poor. RF ablation is a treatment that could perhaps extend survival and potentially improve quality of life by improving local control.

The Gastrointestinal Committee recommends a trial that would serve to test this hypothesis. Such a trial could also serve to evaluate specific characteristics of lesions, such as size, location, and blood supply, which are used to predict the success or failure of RF ablation.

CANCER IN THE THORAX

Although predominantly focusing on lung cancer, the Thoracic Disease Committee also will consider trials of cancer arising from extrapulmonary sites. Its strategy through 2007 is to focus on five major activities. The first of these is continuation of the current National Lung Screening Trial (ACRIN 6664), including analysis of early data.

The second activity is measurement of tumor response to therapy in patients with lung cancer. Specifically, researchers want to answer the following questions: (a) Does measurement of tumor volume on CT scans provide an earlier and more reliable indication of response to antitumor therapy than unidimensional or bidimensional measurement? (b) Is measurement of the standardized uptake value (SUV) or the volume-corrected SUV on PET images a better indicator of tumor response than measurements of tumor size on CT scans?

The third activity is assessing the single pulmonary nodule for malignancy with PET. Specific questions include (a) Is 3-deoxy-3-fluorine-18-fluorothymidine (FLT) PET superior to 2-deoxy-2-fluorine-18-fluoro-D-glucose (FDG) PET in the distinction of benign from malignant pulmonary nodules? (b) Are size-corrected SUVs more accurate than SUVs that are not size corrected in the distinction of benign from malignant pulmonary nodules?

The fourth activity is assessing the role of FDG PET in determining treatment of esophageal cancer—including (a) in patients receiving therapy for localized esophageal cancer, can PET be used to distinguish those who are responding to therapy from those who are not responding? (b) Does incorporation of PET in the follow-up and treatment planning of these patients improve survival?

The fifth activity is RF ablation of lung tumors. Specifically, in patients with medically inoperable early stage lung cancer, can RF ablation, in conjunction with either stereotactic radiation therapy or catheter-delivered brachytherapy, be used to improve survival compared with stereotactic radiation therapy?
RATIONALES AND IMAGING PROPOSALS

National Lung Screening Trial

The National Lung Screening Trial is a collaboration of ACRIN and the Lung Screening Study of the National Cancer Institute Prostate, Lung, Colon, and Ovary trial; it is intended to determine if CT screening for lung cancer reduces lung cancer–specific mortality in high-risk individuals. The National Lung Screening Trial began enrollment in September 2002 and concluded enrollment in February 2004, with a total of 53,418 participants (18,893 patients enrolled in ACRIN, and 34,525 enrolled in the Lung Screening Study). Screening will continue through the middle of 2006, and health information will be obtained through 2009. Early publications will include results of prevalence and incidence screening, results of the recruitment process, development of image quality control standards, and the influence of the screening process on smoking habits and beliefs.

Measurement of Therapeutic Response in Patients with Lung Cancer

Lung cancer is the most common cause of cancer death in both men and women in the United States and worldwide. Evaluating response to the different treatment regimens within clinical trials for lung cancer and comparing these results with other trials requires consistent criteria for determining response. In 1994, the World Health Organization criteria for reporting the results of cancer treatment were reviewed, and revised guidelines known as Response Evaluation Criteria in Solid Tumors were proposed, with treatment response determined by using a single measurement of the largest tumor diameter in the transverse plane. Although interobserver agreement in size determination of small pulmonary tumors is good when spiral CT is used to assess well-defined tumors, considerable variability exists in measurements of lung tumors performed by different readers if the lesion is irregular (23). These interobserver differences in measurements can erroneously affect the data and outcomes of clinical trials. Adaptation of tumor measurements by using volumetric analysis tools may allow an earlier and less variable assessment of tumor response (24).

The development of novel agents, such as gene therapy, as well as treatment protocols that target tumor biology—including tumor cell proliferation and invasion, angiogenesis, and metastasis—further complicate the measurement of tumor response. Antitumor effect in many of these regimes is cytostatic and, unlike anticancer cytotoxic agents, may not cause regression in tumor size. Anatomic measurements consequently may not adequately reflect the efficacy of cytostatic agents in cancer trials. Molecular imaging technologies, such as PET, can enable analysis of the metabolic activity of tumors and will be the subject of ACRIN trials to assess their capabilities to provide early and accurate information on the effectiveness of these new therapeutic agents. This will be performed by embedding ACRIN imaging trials in treatment trials of novel therapeutic agents. Researchers will perform serial measurements of tumor size on CT scans, including unidimensional, bidimensional, and volume measurement, and PET SUV measurements of the primary tumor sites before and after treatment. Researchers will then analyze the results in the context of pathologic response of the tumor and patient survival. ACRIN also will focus on optimal ways of using PET data to assess treatment effect, as PET measurements can be subject to error. The committee proposes to test the possibility that volumetric analysis of SUV values, in a manner similar to anatomic volumetric analysis, may be more reliable in the prediction of tumor response.

PET Assessment of the Solitary Pulmonary Nodule

With the implementation of screening CT for lung cancer and the frequent detection of pulmonary nodules with CT, a noninvasive means of detecting neoplasia is necessary. FDG PET is more sensitive (specificity of 92%–96%) and specific (specificity of 78%–96%) than CT in the detection of malignancy in pulmonary nodules that are larger than 1 cm in diameter. Despite its improved depiction of neoplasia, there are tumors that are not consistently detected with FDG PET. Tumors such as carcinoid bronchioalveolar cell carcinoma and other well-differentiated adenocarcinomas are frequently determined to be false-negative with FDG PET. FDG is also taken up by inflammatory and infectious processes that can mimic neoplasia. Recent studies have shown that FLT is a useful biomarker for tumor cell proliferation (25). Preliminary studies have shown that FLT PET may be better than FDG PET in distinguishing benign nodules.

The committee proposes a multicenter study to evaluate the sensitivity and specificity of FDG PET and FLT PET in the detection of malignancy of pulmonary nodules on the basis of estimation of glucose metabolism with FDG PET and thymidine turnover rate with FLT PET. This study could allow researchers to determine if the uptake of FDG and FLT varies according to tumor characteristics, such as specific tumor types (adenocarcinoma—including bronchioalveolar carcinoma type, large cell, squamous cell, carcinoid), tumor grade (high vs low grade), and differentiation (good vs poor).

PET Evaluation of Treatment for Esophageal Cancer

Esophageal cancer is the sixth leading cause of death from cancer worldwide. Patients with localized esophageal cancer are treated with surgical resection. Trials have also been conducted on preoperative radiation therapy, preoperative chemotherapy, and combined preoperative chemoradiation. There may be a subset of patients with local-regional esophageal cancer who experience a substantial pathologic response to multimodality therapy. If patients who respond to treatment could be distinguished from those who do not respond to treatment, aggressive treatment could be pursued in the first group of patients and discontinued in the second. In preliminary studies, FDG PET was used to predict disease-free survival and overall survival on the basis of change in the SUV at the primary site on pre- and posttreatment scans (26). The relative change in SUV has also been shown to correlate with the percentage of viable cells in the resected surgical specimen (27,28).

In patients with surgically resectable esophageal carcinoma, a prospective randomized trial incorporating pre- and posttreatment FDG PET scanning is proposed. PET data would be correlated with patient survival and morbidity as they relate to chemoradiation, surgery, and palliative care. Tissue specimens would be collected so that imaging phenotype could be correlated with cancer genotype.

RF Ablation of Lung Cancer

Surgical resection is the standard of care in patients with early-stage non-small cell lung cancer. Some patients with potentially curable lung cancers are
not candidates for surgery because of limited pulmonary reserve, coexisting morbidity, or both. In these patients, external beam radiation therapy, with or without chemotherapy, is the most frequently used treatment regimen. The 5-year survival rate with this approach ranges from 10% to 30%. In this patient population, the noninvasive strategies being investigated are based on gaining better local control of the tumor and include RF ablation, brachytherapy, stereotactic radiation therapy, and cryotherapy (29–31).

Because RF ablation targets the center of the tumor, combined therapy with external beam radiation therapy—which has maximal effect on the periphery of the tumor—may prove to be advantageous, particularly in larger tumors. Combined RF ablation and catheter-delivered brachytherapy with either high-activity iridium 192 seeds (temporary implants) or low-activity iodine 125 seeds (permanent implants) has also been reported (32). The appropriate follow-up of patients treated with local therapy has not been established.

The committee proposes a multicenter, randomized controlled trial for the purpose of evaluating outcomes in nonsurgical patients with stage I non-small cell lung cancer. In this study, researchers would compare RF ablation in combination with stereotactic radiation therapy and RF ablation in combination with brachytherapy to stereotactic radiation therapy alone. Follow-up would be performed with a combination of contrast-enhanced CT and FDG PET. This trial could be used to evaluate the efficacy of these approaches for local control of lung cancer and develop standards for follow-up and further treatment.

GYNECOLOGIC MALIGNANCIES

In each woman with a gynecologic malignancy, imaging should facilitate the choice of treatments that will best cure or delay the progression of malignant disease so that adverse effects of unnecessary surgery or other treatment can be avoided. The determined treatments should then be targeted to the tumor, thus sparing adjacent organs and tissues that might be damaged by the treatment. The ACRIN Gynecologic Committee chose the following focuses to guide their strategy for research. First, the committee chose to focus on evaluation of women with cervical cancer to guide surgical and chemoradiation therapy decisions, including evaluation of local extent of disease, discernment of characteristics of the primary tumor, and assessment of lymph node metastases. Second, they chose to determine the need for surgical staging in women with endometrial carcinoma. Third, the committee chose to improve the selection of women with adrenal masses needing oophorectomy. Fourth, they chose to improve determination of which women with recurrent ovarian carcinoma would benefit from optimal secondary cytoreduction surgery and prediction of early response to chemotherapy for recurrence.

RATIONALES AND IMAGING STUDIES

Cervical Cancer

Patients with carcinoma confined to the cervix or upper vagina (stage IIA or lower) can be treated with either chemoradiation therapy or radical hysterectomy with bilateral pelvic lymphadenectomy, whereas those with carcinoma with parametrial or lymphatic involvement (stage IIB or higher) are now treated with chemoradiation therapy (33–35). Treatment failures are generally because of resistance to chemoradiation of the primary tumor, as well as occult lymphatic metastases that are not resected or irradiated. Primary chemoradiation without surgery would generally be preferred if the nodal or parametrial tumor or the potential for involved margins at surgery could be accurately identified before surgery (34). More accurate local staging would also allow refining of the volume of tissue treated with targeted radiation therapy, in contrast to the current practice of irradiating the entire pelvis to ensure treatment of undetected metastases in parametrical tissues and pelvic lymph nodes. Data from a collaborative trial between ACRIN and the Gynecologic Oncology Group, or GOG, (ACRIN 6651/GOG 183) are currently being analyzed and used to compare the staging accuracies of CT, MR imaging, and clinical International Federation of Gynecology and Obstetrics standards in patients with stages IB and IIA cervical carcinoma. Accuracy of CT and MR imaging for anatomic delineation should be reevaluated in a few years, as both of these techniques continue to improve.

Angiogenesis and tumor hypoxia are crucial factors for tumor development and progression. The inhibition of angiogenesis is considered one of the most promising strategies that may lead to novel cancer therapies (36). A noninvasive in vivo measure of angiogenesis could enhance the value of studies involving these new treatment strategies (37). Tumor hypoxia contributes to treatment failure (38). Identification and measurement of tumor hypoxia might help physicians predict outcome and select patients for therapies designed to increase the radiosensitivity and chemosensitivity of hypoxic tumor cells (39).

Doppler US signals can be augmented after intravenous injection of microbubble contrast agents, thus improving the correlation between Doppler US and parameters of angiogenesis. Perfusion imaging with either CT or MR imaging techniques can be used to depict both angiogenesis and hypoxia in tumors (36,40). Perfusion CT or MR imaging can be potentially accomplished as part of a comprehensive examination that is also used to evaluate the local extent of the tumor and detect lymphadenopathy. Perfusion imaging with CT and MR imaging and perhaps with advanced US techniques that use novel contrast agents can be correlated in a multicenter trial with parameters such as pathologic and patient response to chemoradiation therapy.

Since metastases are often present within lymph nodes that are not enlarged, decisions regarding extent of radiation therapy have been dependant on extraperitoneal surgical staging. However, preirradiation surgical staging has been associated with an increase in late radiation-related morbidity.

FDG PET is used to detect lymphatic metastases because of their abnormally high metabolic rate. The accuracy of PET is improved by fusing PET and CT images obtained during the same examination. Most investigators are convinced that PET/CT is more accurate in the diagnosis of lymphatic metastases than are current CT or MR imaging techniques (41–43); however, to our knowledge, its negative predictive value has not yet been determined in a large surgically validated series. MR imaging after intravenous administration of lymphotropic ultrasmall superparamagnetic iron oxide nanoparticles can also be used to distinguish between tumor and benign tissue in normal-sized lymph nodes (44). There has, however, been no published trial on the accuracy of this technique and no comparison with PET/CT in patients with cervical carcinoma. ACRIN and the Gynecologic Oncology Group are preparing a protocol to evaluate PET/CT and ultrasmall superparamagnetic iron oxide–enhanced MR imaging for the diagnosis of
lymphatic metastases in patients with locally advanced cervical carcinoma.

**Endometrial Cancer**

Endometrial cancer is highly treatable; 73% of affected women have disease localized to the uterus, with an estimated 5-year relative survival rate of 96%. Many clinicians omit surgical staging in women thought to be at low risk for extraperitoneal extension because of the potential morbidity of surgical staging and because most patients will not have metastatic spread. Women with unfavorable risk factors that predict extraperitoneal metastasis, such as myometrial invasion, tumor grade, histologic characteristics, tumor size, and lymphovascular space invasion, are selected for surgical staging with pelvic and paraaortic lymph node sampling. In patients with lymphatic metastases, maximal cytoreduction has correlated strongly with duration of survival (45).

The potential benefits of an accurate preoperative estimation of the extent of extraperitoneal disease include a tailored surgical approach and streamlined triage of patients to appropriate clinical personnel for therapy (eg, gynecologic oncologist vs general gynecologist). The most important predictive measure for preoperative staging with imaging is the negative predictive value (so that the staging requirement is not inappropriately deferred). US and MR imaging have been more effective than CT in the diagnosis of deep myometrial invasion, but their negative predictive values have been less than 90%. Evaluation of tumoral blood flow with Doppler US has shown initial promise in the identification of tumor aggressiveness. The committee recommends that perfusion imaging and contrast-enhanced vaginal Doppler US be compared with surgical staging and findings of a pathologic examination. In addition, ACRIN should explore the value of PET/CT and ultrasmall superparamagnetic iron oxide–enhanced MR imaging to determine if either of these tests should be used routinely for preoperative diagnosis.

**Adnexal Masses**

In most large series, about two-thirds of adnexal masses removed for suspected malignancy are in fact benign (46–48). In many of these women, surgery could have been avoided if benign disease could have been diagnosed before surgery. Some reports suggest that complex masses can be characterized by using a combination of gray-scale US appearance and Doppler vascular morphology along with Doppler waveforms, as well as use of microbubble US contrast agents. Other series suggest that MR imaging is currently the most accurate imaging study for diagnosis of adnexal masses (49,50). FDG PET performed after an inconclusive US examination can improve the specificity for diagnosis of ovarian cancer, although it is less sensitive for borderline tumors. In designing trials to test these technologies, the key questions are negative predictive value (so that surgery is not delayed in women with early ovarian carcinoma), overall cost effect, and patient outcomes.

**Ovarian Cancer**

Approximately 80% of patients with epithelial ovarian cancer treated with maximal surgical cytoreduction followed with platinum and taxane-based primary therapy enter complete clinical remission. Unfortunately, 70% of optimally debulked patients experience a relapse at a median duration of approximately 18 months. Treatment options for patients with recurrent disease are dependent on time from prior treatment, previous response to chemotherapy, and extent of disease. Complete or optimal secondary surgical cytoreduction has been associated with prolonged survival in selected patients (51). However, no survival benefit has been shown for suboptimal secondary cytoreduction. Thus, the objective of imaging in the detection of recurrent ovarian carcinoma is to identify patients who are likely to achieve a complete or at least optimal secondary surgical cytoreduction.

Traditionally, we have classified patients with relapse at less than 6 months follow-up as platinum resistant and have selected nonplatinum agents for treatment; those with relapse at greater than 6 months are termed platinum sensitive and are administered a platinum compound or, more recently, a platinum-based combination treatment. Response assessment may require between two and four cycles of therapy, with associated side effects, until response manifests as either a decline in the serum CA-125 level or a radiographic or symptomatic improvement. Patients are often treated empirically with multiple sequential agents, with intervals separated by the time required to assess response.

FDG PET/CT has shown improved accuracy in the identification of bulky recurrent ovarian carcinoma, and it has been used to show early response to chemotherapy before changes in size are visible on conventional images (52). PET/CT and surgical results should be correlated, with overall survival and quality of life as the primary end points in women whose test results are positive for CA-125 but have a negative or equivocal CT scan, as well as in women with a recurrence that appears to be localized on CT scans. In addition, ACRIN should test the hypothesis that a PET/CT response after one cycle of therapy might potentially serve as an in vivo predictor of subsequent clinical response. The magnitude of PET/CT response should be evaluated as a possible predictor of the duration of chemotherapy response. This could be addressed in conjunction with any phase II study of a cytotoxic agent that requires depiction of measurable disease. Changes in PET/CT after one cycle of therapy would be correlated with subsequent clinical response rate and time to treatment failure.

**GENITOURINARY CANCERS**

The Genitourinary Committee will focus on the following issues: (a) evaluation of new imaging methods to enable preoperative staging of cancer as a means of improving the choice of treatment; (b) assessment of new methods of improving surveillance for tumor recurrence, including the use of smart systems to improve tumor recognition and further understanding of the natural history of these cancers; and (c) establishment of percutaneous methods of treating genitourinary malignancies in a less invasive manner.

Specifically, the committee will develop trials related to these considerations for the common genitourinary tract tumors of the prostate, bladder, and kidney.

**RATIONALES AND IMAGING APPROACH**

**Prostate Cancer**

In men, prostate cancer is the second leading cause of death and the most common cancer diagnosed. Screening for prostate cancer with prostate-specific antigen is well established; however, compliance is variable, and the appropriateness of such screening is controversial (53). Radical prostatectomy and radiation therapy remain the established definitive treatment options. Reliable localization of prostate cancer is of increasing therapeutic importance as a result of the
emergence of disease-targeted ablative therapies, such as interstitial brachytherapy, intensity-modulated radiation therapy, high-intensity focused US, and cryosurgery (54). Such focal treatment holds the promise of substantially reducing the morbidity associated with treatment of the entire prostate, whether with surgery or with radiation.

The selection of appropriate treatment options and the use of focal therapy all require accurate noninvasive evaluation of disease extent and aggressiveness. Unfortunately, all current methods of assessment are of limited accuracy. MR imaging and spectroscopy have shown substantial promise as innovative radiologic tools for use in the local evaluation of prostate cancer (55–57). Early data suggest that these modalities can positively affect treatment planning (58,59). A currently open ACRIN trial (ACRIN 6659) will enable researchers to answer key questions about the utility of these technologies and help guide future definitive therapies. This trial will enable researchers to evaluate the contribution of MR spectroscopy (eg, color mapping of MR spectroscopic information) to MR imaging in the localization and determination of the extent of prostate cancer.

**Bladder Cancer**

Transitional cell carcinoma of the bladder treated with radical cystectomy with limited or extended pelvic lymph node dissection has been studied since before 1946 (60). The involvement of local-regional lymph nodes varies and increases with extent of tumor invasion. Despite the use of regional lymphadenectomy at the time of radical cystectomy, there are few data from randomized trials to define an association between the extent of surgical lymph node resection and recurrence or survival.

Recent evidence suggests that lymph node imaging with MR and iron-based contrast agents may improve preoperative diagnosis and help better direct surgical therapy (44). Thus, selective limited lymph node dissections may be more precisely performed. The committee proposes a collaborative trial appended to an American College of Surgeons Oncology Group, or ACOSOC, trial (ACOSOG Z7031), which is a prospective evaluation of the benefit of a limited versus an extended pelvic lymphadenectomy performed at radical cystectomy for bladder cancer. The imaging arm would use a novel contrast agent, ultrasmall superparamagnetic iron oxide (Combidex; Advanced Magnetics, Cambridge, Mass), to image the internal architecture of lymph nodes before surgery and enable researchers to determine whether this approach improves the predictive value for determining the extent of lymphadenectomy. PET and other molecular imaging technologies may also prove useful for tumor staging and lymph node evaluation, and they may be the focus of additional ACRIN trials.

A second focus related to transitional cell carcinoma of the urinary tract will be surveillance of patients who have shown a propensity to develop urothelial cancers. The biology of urothelial cancers is considered a “field theory,” in that once a cancer has been discovered, all the remaining urothelium remains at risk for the development of cancer, and lifelong surveillance becomes a necessity. ACRIN will evaluate screening and surveillance imaging techniques, such as x-ray CT urography for patients with urothelial carcinoma of the bladder and other urothelial neoplasms. ACRIN will use the image archive generated with such trials to test the value of sophisticated three-dimensional software in the manipulation of images as a means of providing enhanced diagnostic information.

**Renal Cell Carcinoma**

Most renal cell carcinomas are discovered today by means of serendipity due to the increased use of CT, US, and MR imaging performed for nonrenal causes. As a result, more renal cell carcinomas are being discovered at a smaller size and, likely, a lower stage. Cross-sectional imaging with CT or contrast-enhanced MR imaging offers a fast and accurate method to closely follow and observe the behavior of small solid renal neoplasms. Newer volumetric software and three-dimensional reconstruction promise improvement over conventional transverse imaging. Functional (molecular) imaging techniques with PET and radiopharmaceuticals such as $^{13}$C-acetate performed to enable lesion detection and characterization may have value in following small renal cell carcinomas, as increased metabolism may occur before changes in the size of a neoplasm. The appropriate use of these imaging technologies to better characterize and follow incidentally discovered renal masses will require well-controlled clinical trials.

Definitive treatment for renal cell carcinoma remains partial or total nephrectomy with open or laparoscopic techniques. However, new therapies, such as RF ablation, have generated much interest in minimally invasive image-guided approaches. RF ablation for small lesions seems promising, but follow-up of patients treated with RF ablation proves challenging, since we have yet to learn the precise posttreatment patterns, as depicted on cross-sectional images. ACRIN trials of RF ablation and other less invasive image-guided treatment technologies could elucidate the answers to such questions.

**Informatics**

In addition to its organ-specific trial activities, the Genitourinary Committee proposes a cross-cutting effort to develop genitourinary radiology ontology, which might build on current activities of the National Library of Medicine Unified Medical Language System project. This project has three components: (a) metathesaurus, which contains information about biomedical concepts and terms from controlled vocabularies and classifications used in patient records, administrative health data, and bibliographic and full text databases; (b) SPECIALIST Natural Language Processing System, which is a lexicon of biomedical terms (available at: www.specialist.nlm.nih.gov); and (c) semantic networks, which categorize concepts represented in the metathesaurus with links between the concepts, thus providing a structure for the representation of important relationships in the biomedical domain. Building such “smart databases” for genitourinary diseases could be an important step into the future of data sharing beyond merely the transfer of clinical images. Such an effort would also lead to collaboration with those who have developed artificial neural networks in genitourinary cancer (61,62).

**MALIGNANCIES OF THE HEAD AND NECK AND CENTRAL NERVOUS SYSTEM**

A number of malignancies affect the head and neck and the central nervous system. The committee’s strategic focus initially is on primary cerebral gliomas and squamous cell carcinoma in the head and neck. These are the most common brain and head and neck malignancies, respectively, and they are associated with substantial morbidity and mortality. The prevalence of these tumors and their dismal prognoses provide substantial motivation for research.

With regard to cerebral gliomas and
squamous cell carcinomas, the committee will develop clinical trials in which researchers will (a) evaluate imaging technologies that may allow for less invasive differentiation of malignant tissue from other abnormalities, such as hemorrhage or fibrosis; (b) use metabolic imaging, physiologic imaging, or both to predict the clinical outcomes of patients with brain tumors; (c) investigate the potential of imaging technologies to improve the staging of head and neck cancers to improve outcomes; and (d) evaluate the potential of imaging technologies to aid in the prediction of the effectiveness of treatment.

RATIONALES FOR INVESTIGATION

Primary Cerebral Gliomas

There are substantial differences between high-grade cerebral gliomas and neoplasms in other body locations, such as the colon, breast, and lung. Cerebral glioma is relatively uncommon, with an estimated 18,400 new primary intracranial gliomas and 12,600 deaths in the United States in 2004 (63). While its low prevalence precludes consideration of screening for cerebral glioma from a cost perspective, detection is rarely difficult. Conventional MR imaging techniques such as T2-weighted or fluid-attenuated inversion-recovery images and contrast-enhanced MR imaging are very sensitive for detection of parenchymal lesions. On the other hand, characterization of a lesion as a neoplasm rather than an infectious or inflammatory process can be problematic. Another key difference from other types of cancer is that patients with high-grade gliomas have 5-year survival rates of less than 15% (64), and high-grade gliomas progress rapidly, which reflects the lack of effective treatments. Thus, cerebral gliomas cause substantial morbidity and mortality.

For these reasons, our proposed imaging research strategy is to focus on imaging technologies that permit rapid and accurate characterization of a brain lesion as a neoplastic process (differential diagnosis) and that accurately reflect tumor viability, activity, or both (treatment monitoring and planning), with a particular eye toward the use of these techniques to aid in the development of effective therapies. Recently developed MR and nuclear imaging tools have become available that have great potential to enable researchers to address these two issues; therefore, there is a high likelihood that patient care may change and disease-associated morbidity and mortality may be potentially reduced.

A common difficult clinical scenario is the task of distinguishing neoplasm from infection, inflammation, or infarction or vascular causes with conventional imaging (65). It is imperative not to delay therapy by awaiting follow-up data; conversely, one does not want to perform unnecessary biopsy or surgery. Because of the rapid progression of high-grade gliomas, early accurate characterization is crucial for instituting appropriate therapy quickly, which may increase survival time. Furthermore, the therapeutic approach for a low-grade lesion may be different from that for a high-grade lesion. Hence, accurate diagnosis, including assigning a tumor grade, would be of great clinical value.

Our second focus in the evaluation of early therapeutic response relies on accurate and early assessment of tumor viability, activity, or both. Such assessments are imperative for quick adaptation of treatment regimens in patients for whom life expectancy is currently very limited and changes to alternative therapies on a rapid basis is crucial.

In addition, the noninvasive assessment of tumor viability, activity, or both could greatly aid the development and testing of new therapies. Imaging techniques that provide “biomarkers” of tumor pathobiology, such as angiogenesis, vascular permeability, and metabolic activity, might permit rapid noninvasive evaluation of the efficacy of many experimental therapies before more conventional morphologic (eg, tumor size, vascularity, edema) and clinical (eg, neurologic scores, death) parameters would be expected to change. For example, high-grade gliomas appear to require an increased blood supply compared with normal brain tissue, and there is a strong correlation between increased tumor neovascularity (eg, irregular enlarged microvessels recruited via angiogenesis) and malignancy (66). Drugs that prevent angiogenesis and consequently inhibit tumor growth are actively being pursued by researchers because of the lack of success of standard cancer treatments. To evaluate the clinical potential of these new drugs, a standardized widespread method for early noninvasive monitoring of their efficacy would be tremendously beneficial. The recovery of tissue for genomic and proteomic assessment, so as to best develop molecular imaging tools in the future, will be an important feature of the committee’s trial portfolio because of the potential of molecular diagnoses in the future.

Squamous Cell Carcinoma in the Neck

There were an estimated 28,260 new cancers and 7230 deaths from lesions of the oral cavity and pharynx in the United States in 2004, with almost all of these being squamous cell carcinoma (63). Curative resection can be disfiguring and cause severe functional limitation; for this reason, the minimally required surgical extent is desired, and it is crucial that staging be accurate to remove nodal metastases but spare as much function as possible.

The presence or absence of nodal disease is one of the most important factors in predicting disease recurrence and survival in patients with squamous cell carcinoma; for example, the presence of nodal metastases can reduce 5-year survival by as much as 50% (67,68). Unfortunately, the surgical morbidity associated with extensive neck dissection is substantial, and some reports suggest that two-thirds of patients who do not have clinically palpable lymph spread and yet undergo surgical exploration to rule out microscopic disease end up not having nodal disease. They experience the morbidity of neck dissection without benefit. Similar concerns are present in monitoring patients after initial surgical therapy. Attempts to develop scintigraphic, dye-based, or other staging techniques are still controversial and largely confined to single-center studies. The lack of success of these approaches has been in part because the radiotracers or dyes used to date have not been cancerspecific markers. More specific markers, such as radiolabeled thymidine—which is a PET tracer—might provide more specific information about the status of lymph nodes and eventually reduce the need for surgical exploration while providing accurate nodal status (67).

Potential Trials

There are some commonalities in the approach with which the committee intends to address the two neoplasms. In addition, problems that can be solved with current technology should naturally come before those that require new technical solutions. Thus, our short-term priorities include designing multicenter trials to (a) determine the prognostic value of MR spectroscopy in newly identified intracranial mass lesions (differential diag-
(b) determine the prognostic value of perfusion MR imaging in newly identified intrathelial mass lesions and the role of perfusion MR indexes as a surrogate for degree of tumor angiogenesis and a marker for response to therapy, (c) determine the utility of FLT in staging squamous cell carcinoma, (d) determine the role of diffusion-weighted or diffusion-tensor imaging as a marker of acute cytotoxicity, and (e) investigate tissue hypoxia markers for cerebral glioma and squamous cell carcinoma.

As new technologies develop, new opportunities will follow. Later investigatory priorities will focus on the evaluation of novel targeted molecular agents by using PET, MR imaging, and other image receptor technologies to identify intrathelial mass lesions or squamous cell carcinoma; diffusion tensor MR imaging as a potential marker for radiation therapy toxicity or as a tool for treatment planning; and evaluation of tools for differentiating radiation necrosis from tumor recurrence.

References
39. Chiarotto JA, Hill RP. A quantitative analysis of the reduction in oxygen levels required to induce up-regulation of vascular endothelial growth factor (VEGF)
41. Lin W, Hung YC, Yeh LS, Kao CH, Yen RF, Shen YY. Usefulness of (18)F-fluorodeoxyglucose positron emission tomography to detect para-aortic lymph nodal metastasis in advanced cervical cancer with negative computed tomography findings. Gynecol Oncol 2003; 89:73–76.
Contrast Medium–induced Nephrotoxicity: Which Pathway?  

Acute renal failure (ARF), the sudden and rapid deterioration in renal function that results in the failure of the kidney to excrete nitrogenous waste products and to maintain fluid and electrolyte homeostasis, can be the direct result of parenteral contrast medium administration, which is the third most common cause of hospital-acquired ARF (1,2). Contrast medium–induced ARF may range in severity from asymptomatic, nonoliguric transient renal dysfunction to oliguric (urine volume of less than 400 mL/d), severe ARF that necessitates dialysis.

Findings in recent reports in the cardiology literature indicate that the development of contrast medium–induced ARF after diagnostic coronary angiography and percutaneous interventional association is prolonged hospitalization, marked increases in morbidity, and early and late mortality. In 1826 consecutive patients who were undergoing percutaneous intervention, McCullough et al in 1997 (3) reported an incidence of contrast-induced nephropathy in 14.5% who did not require dialysis and in 0.7% who did require it. For patients who did not require dialysis, there was a sixfold increase in inhospital mortality; there was a 33-fold increase in mortality in those patients who required dialysis. In 2082 percutaneous interventions for acute myocardial infarction in patients without shock, Sadeghi et al in 2003 (4) reported a more than 13-fold (1.2% in patients without vs 16.2% in patients with contrast-induced ARF) increase in 30-day mortality and a more than sevenfold (3.2% vs 23.3%) increase in 1-year mortality.

Other than hydration, findings in most prior studies of measures to prevent contrast-induced ARF have been neutral, or researchers have found deleterious effects or, in the case of N-acetylcysteine, have reported mixed results (5–8).

It is difficult to prevent contrast-induced nephropathy when the cause is not understood. There are three relatively distinct mechanisms or pathways proposed for the pathophysiology in contrast-induced ARF: hemodynamic effects, direct contrast medium molecule tubular cell toxicity, and endogenous biochemical disturbances such as an increase in oxygen-free radicals and/or a decrease in antioxidant enzyme activity (Figure). On the other hand, an interaction of any two, all, or some other effects are also presumed as additional possibilities. The noteworthy investigation by Heinrich et al in the current issue of Radiology (9), and important results in recent clinical reports about potential antidotes, are the stimuli for this commentary.

Hemodynamic Effects

Of the three major pathways depicted in the Figure, the greatest amount of recent clinical attention has focused on the hemodynamic effects of contrast media. This is somewhat curious in view of a previous insight, almost 25 years ago, by Gilbert Mudge, MD, a leading expert in the field of contrast-induced ARF, who stated in a comprehensive review of this subject that “Despite repeated statements to the contrary, there is little evidence that ischemia is the mechanism by which urographic agents produce renal failure” (10). My academic interest has been to study the hemodynamic effects of and mechanisms of action of contrast media on renal function, and I favor this assessment.

Interest in the hemodynamic pathway was stimulated by three seminal animal investigations that included an assessment of the effects of hypertonic contrast media on renal blood flow. A decrease in renal blood flow following the direct intraarterial injection of contrast material into the renal vascular bed, which is unique in comparison with other vascular beds, was initially noted by Talner and Davidson in 1968 (11), Caldicott et al in 1970 (12), and Sherwood and Lavender in 1969 (13). The hemodynamic response with the direct intraarterial injection of contrast material is actually biphasic, with an initial increase followed by a decrease. Although the decrease in blood flow is only about 30% in comparison with baseline and is transient, lasting for only several minutes, this observation initiated the hypothesis that ischemia could be a likely candidate for the cause of contrast-induced renal toxicity. Subsequently, there have been numerous pharmacologic agents proposed to eliminate or minimize the contrast-induced decrease in blood flow: dopamine, endothelin antagonists, adenosine antagonists, calcium-channel blockers, and fenoldopam mesylate, none of which have proved to be clinically efficacious (14).

The biphasic changes in renal blood flow seen during direct injections of contrast material into the renal artery are far too transient and minimal to produce ischemic damage. Animal models of ischemic ARF require severe prolonged in-
Sion can lead to necrosis of the medullary crease in blood flow and/or oxygen tension (22). The hypothesis is that either a decrease in renal vascular resistance in the cortex, observations are valid, a decrease in the ARF in the hospital setting and includes any condition that induces hypovolemia, such as hemorrhage, dehydration, low cardiac output, and lowered systemic vascular resistance that can occur with general anesthesia (1,2,6). Most intravascular contrast media are osmotic diuretics and can, thus, induce or exacerbate hypovolemia secondary to their dehydrating effects. Interestingly, Fang et al in 1980 (29) found low fractional excretion of sodium during the oliguric phase of contrast-induced ARF, a more common characteristic of hypovolemic ARF than of direct renal injury. In addition, contrast media cause vasodilatation, usually mild and transient, but uncommonly lead to prolonged systemic hypotension (6,30). The indirect effect, therefore, of contrast media on renal hemodynamics is a reasonable hypothesis in regard to the pathophysiology involved and is rapidly reversible when restoration of renal blood flow and normal glomerular filtration rate occurs with hydration and/or administration of pharmacologic agents that correct hypotension.

Direct Contrast Medium Molecule Tubular Cell Toxicity

The study by Heinrich et al (9) that appears in the current issue of Radiology was rigorously performed, is timely, and has substantial results. Prior evidence for direct contrast medium toxicity to the renal tubular cell has been suggested by findings in studies in which the renal extraction of p-aminohippurate was assessed. A decline in p-aminohippurate indicates a reduction in tubular cell transport (secretory) of the proximal tubules of the cortical nephrons and can be an indicator of a toxic effect independent of the hemodynamic changes. The first such indication was suggested by Talner...
and Davidson in 1968 (31). In their study in animals, they determined that a hypertonic contrast agent is capable of inducing a decrease in p-aminohippurate, exclusive of the effect of hypertonicity. Findings in that study were later corroborated by Dibona in 1978 (32). Fabico et al in 1989 (33) performed a study in dehydrated dogs in which dehydration was sufficient to reduce the glomerular filtration rate and effective renal plasma flow. They found that the parenteral administration of a hypertonic contrast agent further diminished the already compromised renal hemodynamic function, caused by dehydration, and impaired the renal tubular cell transport mechanism, as evidenced by a marked reduction in p-aminohippurate. The depressive effect on p-aminohippurate could not be attributed to the osmotic properties of the contrast medium itself.

Humes et al in 1987 (34) reported a possible direct deleterious effect of sodium diatrizoate on renal tubular cells; they used suspensions enriched in rabbit proximal tubular cell segments incubated with sodium diatrizoate. A variety of well-established metabolic parameters used to quantitate the extent of cell injury were measured. Sodium diatrizoate produced marked declines in tubular cell potassium (K⁺), adenosine triphosphate, and total adenine nucleotide contents; marked decreases in tubular cell basal and uncoupled respiratory rates; and a marked increase in tubular cell calcium content, an increase that is indicative of cell injury. A period of 22 minutes 30 seconds of hypoxia also caused deleterious changes in each of these quantitative indexes of cell viability, and diatrizoate potentiated the degree of hypoxia-induced cell injury. The effects could not be explained on the basis of the toxicity of the contrast agent, since equimolar concentrations of mannitol had no detrimental effects on the cell viability parameters. It was concluded that a possible direct toxic effect of the contrast material on plasma membrane permeability and transport processes was a likely cause.

The current investigation in Radiology reinforces and expands on the importance of direct tubular cell toxicity on the pathophysiology of contrast-induced ARF (9). The authors have shown convincing evidence of a direct cellular toxicity of contrast agents independent of either hemodynamic mechanisms or osmolarity. This refocuses attention on the contrast medium molecule itself and on direct cellular mechanisms for elucidation of the pathophysiology of contrast-induced ARF and, thus, on the potential for a solution. The model also suggests development of an assay for preclinical testing of newly developed contrast media molecules and potential strategies for engineering safer agents.

The finding of differential effects of contrast media molecules with incubation for 24 hours supports our speculation that situations that lead to a prolongation of contrast media dwell time in the tubular lumen can be a major factor in renal toxicity. Specific circumstances where there is a delay in the excretion of contrast media molecules from the tubular lumen would occur in prerenal clinical conditions, such as hypovolemia secondary to dehydration (that can be exacerbated by the osmotic diuretic effects of contrast media) and with the osmotic diuresis that can occur with diabetics mellitus (1). The effects also may be exacerbated by an increased concentration of the contrast medium molecules in the renal tubular lumen, as would occur in dehydration.

Oxygen-Free Radicals

The third pathway that is proposed is an increase in oxygen-free radicals or a decrease in antioxidant enzyme activity triggered by contrast medium administration (35–37). In some ways, this could be a sequela of the direct tubular cell toxicity pathway if the endogenous biochemical disturbances are simply the product of tubular cell damage rather than the primary cause of the resultant tubular cell damage.

Free radicals are atoms or molecules that contain one or more unpaired electrons. In vivo, oxygen molecules are changed into water molecules after successive reduction reactions. Intermediate species are called “reactive oxygen species.” The catabolism of adenosine by xanthine oxidase leads to the formation of one of the reactive oxygen species that has been demonstrated by Katholi et al (38) in vivo and in vitro with contrast material administration. At high concentrations, free radicals have highly deleterious effects on all cellular constituents and cause oxidative stress and protein damage. It has been suggested that reactive oxygen species are important in the renal damage caused by contrast agents (35,36,39). In laboratory animals, contrast agents increase lipid peroxidation (40), and superoxide dismutase, a scavenger of reactive oxygen species, preserves renal function (35).

For many years, N-acetylcysteine has been known to promote detoxification and act as a reactive oxygen species scavenger. On the basis of the assumption that the reactive oxygen species could be involved in the pathophysiology of contrast-induced ARF, researchers in clinical studies have assessed the potential prophylactic effects of the oral or intravenous administration of N-acetylcysteine. The initial positive results by Tepel et al (37) precipitated numerous subsequent assessments of the potential value of N-acetylcysteine in the prevention of contrast-induced nephrotoxicity. The subsequent clinical investigations have been mixed, however, and all lack a sufficiently powered prospective randomized placebo-controlled end point. In the most recent meta-analysis of 16 prospective controlled clinical trials with a total of 1538 patients, Kshirsagar et al (8) in 2004 showed neither a confirmation of validity nor a recommendation for the routine use of N-acetylcysteine. Thus, there is yet no clinical substantiation that the liberation of oxygen-free radicals is the mechanism for contrast-induced ARF.

The current article by Heinrich et al in Radiology provides a renewed focus on the direct tubular cell toxicity pathway of contrast medium–induced ARF and, in my opinion, a promising pathway to solve a growing clinical crisis. In the meantime, we know that close attention to patient hydration before and after the administration of contrast material, attention to the risk factors of preexisting renal insufficiency and diabetes, and active communication with referring clinicians are effective strategies for prevention.

Acknowledgments: I express my appreciation to Patty Gimlen, Laurie Finch, and Vonnie Simpkins for their outstanding secretarial assistance.

References

Multi–Detector Row CT Systems and Image-Reconstruction Techniques

The introduction in 1998 of multi–detector row computed tomography (CT) by the major CT vendors was a milestone with regard to increased scan speed, improved z-axis spatial resolution, and better utilization of the available x-ray power. In this review, the general technical principles of multi–detector row CT are reviewed as they apply to the established four- and eight-section systems, the most recent 16-section scanners, and future generations of multi–detector row CT systems. Clinical examples are used to demonstrate both the potential and the limitations of the different scanner types. When necessary, standard single-section CT is referred to as a common basis and starting point for further developments. Another focus is the increasingly important topic of patient radiation exposure, successful dose management, and strategies for dose reduction. Finally, the evolutionary steps from traditional single-section spiral image-reconstruction algorithms to the most recent approaches toward multisection spiral reconstruction are traced.

Computed tomography (CT) was introduced in the early 1970s and has revolutionized the practice not only of diagnostic radiology but also of the whole field of medicine. CT was the first technology to marry a computer to a medical imaging machine, the first to display x-ray images as cross sections, and the first modality to herald a new era of digital imaging. A glossary of terms used in this review is available online in Appendix E1.

EVOLUTION OF SPIRAL CT: FROM ONE SECTION TO 16

The introduction of spiral CT in the early 1990s constituted a fundamental evolutionary step in the development and ongoing refinement of CT imaging techniques (1,2). For the first time, volume data could be acquired without misregistration of anatomic detail. Volume data became the basis for applications such as CT angiography (3), which has revolutionized the noninvasive assessment of vascular disease. The ability to acquire volume data also paved the way for the development of three-dimensional (3D) image-processing techniques such as multiplanar reformation (MPR), maximum intensity projection, surface-shaded display, and volume-rendering techniques (4), which have become a vital component of medical imaging today.

Ideally, volume data are of high spatial resolution and are isotropic in nature: Each image data element (voxel) is of equal dimensions in all three spatial axes, and this forms the basis for image display in arbitrarily oriented imaging planes. For most clinical scenarios, however, single-section spiral CT with a 1-second gantry rotation is unable to fulfill these requirements. To prevent motion artifacts and optimally utilize the contrast agent bolus, body spiral CT examinations need to be completed within a certain time frame of, ordinarily, one breath hold (25–30 seconds). If a large scan range such as the entire thorax or abdomen (30 cm) has to be covered within a single breath hold, a thick collimation of 5–8 mm must be used. While the in-plane resolution of a CT image depends on the system geometry and on the reconstruction kernel selected by the user, the
Multi–detector row CT allows substantial reduction in examination time for standard protocols, coverage of extended anatomic volumes, and, most important, substantially increased longitudinal resolution by means of reduced section width.

Near-isotropic spatial resolution in routine examinations, which has been achieved with 16-section CT systems, enables 3D renderings of diagnostic quality and oblique MPRs and maximum intensity projections with resolution similar to that of the transverse images.

Scanning at narrow collimation does not markedly increase the radiation dose to the patient, as long as the effective milliampere-seconds level is kept constant.

A key challenge for image reconstruction with multi–detector row CT is the cone angle of the measurement rays; this requires novel reconstruction techniques such as 3D back projection, AMPR, or weighted hyperplane reconstruction.

Z filtering makes it possible to reconstruct images retrospectively with different section widths from the same raw CT data set, trading off, in this way, z-axis resolution and image noise.

ESSENTIALS

- Multi–detector row CT allows substantial reduction in examination time for standard protocols, coverage of extended anatomic volumes, and, most important, substantially increased longitudinal resolution by means of reduced section width.
- Near-isotropic spatial resolution in routine examinations, which has been achieved with 16-section CT systems, enables 3D renderings of diagnostic quality and oblique MPRs and maximum intensity projections with resolution similar to that of the transverse images.
- Scanning at narrow collimation does not markedly increase the radiation dose to the patient, as long as the effective milliampere-seconds level is kept constant.
- A key challenge for image reconstruction with multi–detector row CT is the cone angle of the measurement rays; this requires novel reconstruction techniques such as 3D back projection, AMPR, or weighted hyperplane reconstruction.
- Z filtering makes it possible to reconstruct images retrospectively with different section widths from the same raw CT data set, trading off, in this way, z-axis resolution and image noise.
became feasible. This 1.0–1.25-mm longitudinal resolution combined with the improved contrast resolution of modern CT systems enabled noninvasive depiction of the coronary arteries (19–22). Initial clinical studies demonstrated the potential of multi–detector row CT to not only demonstrate but to some degree also characterize noncalcified and calcified plaques in the coronary arteries on the basis of plaque CT attenuation (22,23).

The limitations of four– and eight–detector row CT systems, however, have so far prevented the successful integration of CT coronary angiography into routine clinical algorithms: Stents or severely calcified arteries constitute a diagnostic dilemma, mainly because of partial volume artifacts as a consequence of insufficient longitudinal resolution (22). For patients with a higher heart rate, careful selection of separate reconstruction intervals for different coronary arteries has been mandatory (25). It is almost impossible for patients with manifest heart disease to comply with the breath-hold time of about 40 seconds required to cover the entire heart volume (approximately 12 cm) with four-section CT. The ongoing technical refinement of multi–detector row CT, however, holds the promise of gradually overcoming some of these limitations. The most important steps toward this goal are gantry rotation times faster than 0.5 second (26,27) for improved temporal resolution and robustness of use, 16-section submillimeter acquisition for increased longitudinal resolution and shorter breath-hold times, and novel sophisticated approaches for image acquisition and reconstruction.

In this review, ECG-synchronized examinations of the heart and of the cardiothoracic anatomy will be very succinctly discussed, since this topic has been extensively reviewed elsewhere (28). Similarly, advanced 3D postprocessing techniques are omitted. In this article, we will review the general technical principles of multi–detector row CT as they apply to the established four– and eight–detector row systems, the more recent 16–detector row scanners, and generations of CT systems yet to come. On the basis of the technologic description of different scanner types and image-reconstruction approaches, we provide practical “take-home points” to enable better translation into daily clinical practice of the technology and science reviewed here. Useful up-to-date information regarding multi–detector row CT is also readily available on the Internet at, for example, the UK Medicines and Healthcare products Regulatory Agency CT Web site (www.medical-devices.gov.uk) or the Advanced Medical Imaging Laboratory site (www.ctisus.org).

CURRENT TECHNIQUES

System Design

Detector design.—For clinical purposes, different section widths must be available to adjust the optimum scan speed, longitudinal resolution, and image noise for each application. With a single–detector row CT scanner, different collimated section widths are obtained by means of prepatient collimation of the x-ray beam to obtain different collimated section widths with a single–detector row CT detector. FOV = field of view.

The limitations of four– and eight–detector row CT systems, however, have so far prevented the successful integration of CT coronary angiography into routine clinical algorithms: Stents or severely calcified arteries constitute a diagnostic dilemma, mainly because of partial volume artifacts as a consequence of insufficient longitudinal resolution (22). For patients with a higher heart rate, careful selection of separate reconstruction intervals for different coronary arteries has been mandatory (25). It is almost impossible for patients with manifest heart disease to comply with the breath-hold time of about 40 seconds required to cover the entire heart volume (approximately 12 cm) with four-section CT. The ongoing technical refinement of multi–detector row CT, however, holds the promise of gradually overcoming some of these limitations. The most important steps toward this goal are gantry rotation times faster than 0.5 second (26,27) for improved temporal resolution and robustness of use, 16-section submillimeter acquisition for increased longitudinal resolution and shorter breath-hold times, and novel sophisticated approaches for image acquisition and reconstruction.

In this review, ECG-synchronized examinations of the heart and of the cardiothoracic anatomy will be very succinctly discussed, since this topic has been extensively reviewed elsewhere (28). Similarly, advanced 3D postprocessing techniques are omitted. In this article, we will review the general technical principles of multi–detector row CT as they apply to the established four– and eight–detector row systems, the more recent 16–detector row scanners, and generations of CT systems yet to come. On the basis of the technologic description of different scanner types and image-reconstruction approaches, we provide practical “take-home points” to enable better translation into daily clinical practice of the technology and science reviewed here. Useful up-to-date information regarding multi–detector row CT is also readily available on the Internet at, for example, the UK Medicines and Healthcare products Regulatory Agency CT Web site (www.medical-devices.gov.uk) or the Advanced Medical Imaging Laboratory site (www.ctisus.org).
tions according to the selected beam collimation and the desired section width.

For established four-section CT systems, two detector types are commonly used. The fixed-array detector consists of detector elements with equal sizes in the longitudinal direction. A representative example of this scanner type, the Lightspeed scanner (GE Medical Systems, Milwaukee, Wis), has 16 detector rows, each of them defining a 1.25-mm collimated section width in the center of rotation (8,10,29). The total coverage in the longitudinal direction is 20 mm at the isocenter; owing to geometric magnification, the actual detector is about twice as wide. By means of prepatient collimation and combination of the signals of the individual detector rows, the following section widths (measured at the isocenter) can be realized: four sections at 1.25 mm, 2.5 mm, 3.75 mm, and 5.0 mm (Fig 3a). The same detector design is used for the eight-section version of this system and provides eight sections at 1.25- and 2.5-mm collimated section widths.

A different approach uses an adaptive-array detector design, which comprises detector rows with different sizes in the longitudinal direction. Scanners of this type, the Mx8000 four-section scanner (Philips Medical Systems, Best, the Netherlands) and the Somatom Sensation 4 scanner (Siemens), have eight detector rows (7,9). Their widths in the longitudinal direction range from 1 to 5 mm at the isocenter and allow the following collimated section widths: two sections at 0.5 mm, four at 1.0 mm, four at 2.5 mm, four at 5.0 mm, two at 8.0 mm, and two at 10.0 mm (Fig 3b).

The selection of the collimated section width determines the intrinsic longitudinal resolution of a scan. In a “step-and-shoot” sequential mode, any multiple of the collimated width of one detector section can be obtained by adding the detector signals during image reconstruction. In a spiral mode, the effective section width can be obtained by adding the detector signals during image reconstruction. In a spiral mode, the effective section width can be obtained by adding the detector signals during image reconstruction. In a spiral mode, the effective section width is determined as the FWHM of the detector signal, and the collimated section width is determined as the FWHM of the detector signal. The effective section width is determined as the FWHM of the detector signal, and the collimated section width is determined as the FWHM of the detector signal.

Radiation Dose

Radiation dose and dose efficiency.—Radiation exposure to the patient at CT and the resulting potential radiation hazard have recently gained considerable attention in both the public and the scientific literature (30,31). Typical values for the effective patient dose for selected CT protocols are 1–2 mSv for a head CT, 5–7 mSv for a chest CT, and 8–11 mSv for abdominal and pelvic CT (32,33). This radiation exposure must be appreciated in the context of the average annual background radiation, which is 2–5 mSv (3.6 mSv in the United States). Despite the undisputed clinical benefits, multissection CT scanning is often considered to require increased patient dose compared with the dose from single-section CT. Indeed, a certain increase in radiation dose is unavoidable owing to the underlying physical principles.

In the x-ray tube of a CT scanner, a small area on the anode plate, the focal spot, emits x-rays that penetrate the patient and are registered by the detector. A collimator between the x-ray tube and the patient, the prepatient collimator, is used to shape the beam and to establish the dose profile. In general, the collimated dose profile is a trapezoid in the longitudinal direction. In the umbral region (ie, plateau region of the trapezoid), x-rays emitted from the entire area of the focal spot illuminate the detector. In the penumbral regions, only a part of the focal spot illuminates the detector, while the prepatient collimator blocks off other parts.

With single-section CT, the entire trapezoidal dose profile can contribute to the detector signal, and the collimated section width is determined as the FWHM of this trapezoid. The relative dose utilization of a single-section CT system can therefore be close to 100%. In most cases with multi–detector row CT, only the plateau region of the dose profile is used to ensure an equal signal level for all detector elements. The penumbral region is then discarded, either by a postpatient collimator or by the intrinsic self-collimation of the multissection detector, and represents “wasted” dose. The relative contribution of the penumbral region in-

Figure 3. Illustrations show examples of (a) fixed-array and (b, c) adaptive-array detectors used in commercially available four- and 16-section CT systems.
Radiology

Figure 4. Dose profiles for four- and 16-section CT systems with identical collimated width of one detector (Det.) section. The relative contribution of the penumbral region, which represents wasted dose, decreases with increasing number of simultaneously acquired sections.

Figure 5. Transverse (top) and coronal maximum intensity projection (bottom; 5-mm slab thickness) thoracic CT images in a patient with pulmonary embolism. Scans were acquired with 16-section scanner at 100 kV and 120 mAs. Effective patient dose was 2.3 mSv, 25% less than that for the standard 120 kV protocol. (Images courtesy of Peter Herzog, MD, Klinikum Grosshadern, Munich, Germany.)

creases with decreasing section width, and it decreases with increasing number of simultaneously acquired images. This is demonstrated in Figure 4, which compares the “minimum width” dose profiles for a four-section CT system and a corresponding 16-section CT system with equal collimated width of one detector section. Correspondingly, the relative dose utilization with four-section 1-mm collimation CT is 70% or less (10), depending on the scanner type. With 16-section CT systems and submillimeter collimation, dose utilization can be improved to 84%, again depending on scanner type (25). Some multi–detector row CT systems offer special implementations of even more dose-efficient modes that use a portion of the penumbral region.

A clinically appropriate measure for dose is the weighted CT dose index, or CTDIw (34), which uses the absorbed dose in a polymerized methyl methacrylate (acrylic plastic) phantom as an approximation of the dose delivered to a cross section of the patient’s anatomy (see Appendix E2, radiology.rsna.jnl.org/cgi/content/full/2353040037/DC1). Figure E1 (radiology.rsna.jnl.org/cgi/content/full/2353040037/DC1) shows CTDIw at 120 kV for the 32 cm body phantom as a function of the total collimated width of the detector for a four-section CT system and a 16-section CT system with a similar system geometry. The CTDIw for 16-section CT at 0.75-mm collimation is 7.8 mGy/100 mAs, whereas the CTDIw for four-section CT at 1.0-mm collimation is 9 mGy/100 mAs. Thus, different from the transition from single-section CT to 4-section CT systems, a further increase in radiation exposure with the more widespread availability of 16-section CT systems is not to be expected.

Concepts for radiation dose reduction.—The most important factor for reducing radiation exposure is an adaptation of the dose to the patient’s size and weight (35–37).

As a general rule for the practicing radiologist, the dose necessary to maintain constant image noise has to be doubled if the patient diameter is increased by 4 cm. Correspondingly, for a patient diameter that is 4 cm smaller than average, half the standard dose is sufficient to maintain adequate image quality. This is of particular importance in pediatric imaging. Dose reduction can be achieved by reductions in the milliampere-seconds and voltage settings. Most CT manufacturers provide dedicated pediatric protocols with, for example, milliampere-seconds and voltage settings adjusted according to the weight of the child.

Another means to reduce radiation dose is to adapt the x-ray tube voltage to the intended application. In contrast agent–enhanced studies such as CT angiography, the contrast-to-noise ratio for fixed patient dose increases with decreasing x-ray tube voltage. As a consequence, to obtain the desired contrast-to-noise ratio, the patient dose can be reduced by choosing a lower voltage setting. The potential for dose saving is more substantial for patients with a smaller diameter. This can be demonstrated, for example, by means of phantom measurements of small tubes filled with diluted contrast agent embedded in acrylic plastic phantoms with different diameters (38). The iodine contrast-to-noise ratio at constant radiation dose for various voltage settings is shown in Figure E2 (radiology.rsna.jnl.org/cgi/content/full/2353040037/DC1) as a function of the phantom diameter. Compared with a standard scan at 120 kV in a 32-cm-diameter phantom (corresponding to that for an average adult), the same contrast-to-noise ratio is obtained with 0.49 times the dose (1.3 times the milliampere-seconds setting) for 80 kV and 0.69 times the dose (1.1 times the milliampere-seconds) for 100 kV. Thus, ideally, 80 kV should be used for CT angiography in order to reduce patient dose.

Clinical studies (38) have confirmed these findings and demonstrated a potential for dose reduction of about 50% when 80 kV is used for CT angiography instead of 120 kV. In reality, however, the maximum x-ray tube current available at 80 kV is generally not sufficient to scan bigger patients, which limits the routine application of this approach. Therefore, use of 100 kV appears to be a suitable compromise and the method of choice for CT angiography. Figure 5 shows pulmonary CT angiographic images of a patient with pulmonary embolism; the scan was performed on a 16-section scanner at 100 kV and 120 mAs, and the effective patient dose for this scan was 2.3 mSv, 25% less than that for the standard 120-kV protocol. Authors of recent study (39) recommended 100 kV as the standard mode for thoracic and abdominal CT angiography and report dose savings of 30% without loss of diagnostic information.
An approach that is finding increased implementation in clinical practice is anatomic tube current modulation. With this technique, tube output is adapted to the patient geometry during each rotation of the scanner to compensate for strongly varying x-ray attenuation in asymmetric body regions such as the shoulders and pelvis. The variation of the tube output is either predefined by means of an analysis of a localizer scan (topogram, scout view) or is determined online by evaluating the signal from a detector row. With this technique, dose can be reduced by 15%–35% without degrading image quality, depending on the body region (40,41).

In more sophisticated approaches, tube output is modified according to the patient geometry not only during each rotation but also in the longitudinal direction (automatic exposure control), to maintain adequate dose when moving to different body regions (eg, from thorax to abdomen). In one implementation, the attenuation for each body region of a “standard-sized” patient is stored in the control computer. This attenuation corresponds to the milliampere-seconds setting of the standard protocol. If the actual attenuation of the patient deviates from the “standard” attenuation, the tube output is adapted correspondingly. Figure 6 shows the variation of the milliampere-seconds output for a CT scan of the chest and abdomen in a 6-year-old child. Although the standard protocol with 165 mAs was used—which would have resulted in substantially higher radiation dose than necessary in a standard mode of operation—the average milliampere-seconds value throughout the scan was adjusted to 38 mAs by means of automatic exposure control. Automatic adaptation of tube current to patient size prevents both over- and underirradiation, considerably simplifies the clinical workflow for the technician, and eliminates the need for look-up tables of patient weight and size for adjusting the milliampere-seconds settings.

Radiation dose for ECG-synchronized CT for cardiac applications has been a topic of considerable controversy. Recent studies (32,33) based on four-section CT systems find an effective patient dose of roughly 1 mSv for ECG-triggered calcium scoring with 3-mm section width and roughly 10 mSv for ECG-gated CT angiography of the coronary arteries with 1.0- or 1.25-mm section width. Radiation dose in ECG-gated spiral CT can be reduced by 30%–50% with use of ECG-controlled dose modulation (42,43). During the spiral scan, the output of the x-ray tube is modulated according to the patient’s ECG trace. It is kept at its nominal value during a user-defined phase of the cardiac cycle—in general, the mid- to end-diastolic phase. During the rest of the cardiac cycle, the tube output is typically reduced to 20% of the nominal values, although it is not switched off entirely, to allow image reconstruction throughout the entire cardiac cycle. Thus, although the signal-to-noise ratio is decreased at certain phases of the cardiac applications, the low-dose images are still sufficient for evaluation of functional parameters such as ejection fraction, should this kind of information be desired.

SEQUSTURAL SCANS AND IMAGE-RECONSTRUCTION TECHNIQUES

With the advent of multi–detector row CT, sequential “step-and-shoot” scanning has remained in use for only a few clinical applications, such as head CT, high-spatial-resolution lung CT, perfusion CT, and interventional applications. A detailed theoretical description to predict the performance of multi–detector row CT in sequential mode can be found in reference 44.

The number of images acquired during a sequential scan corresponds to the number of active detector sections. By adding the detector signals of the individual sections during image reconstruction, the number of images per scan can be reduced, and the image section width can be increased. As an example, a scan with four sections at 1.0-mm collimation provides either four images with 1.0-mm section width, two images with 2.0-mm section width, or one image with 4.0-mm section width.

The option to realize a wider section by summing several thin sections is beneficial for examinations that require narrow collimation to prevent partial volume artifacts and low image noise to allow detection of low-contrast details (eg, neurologic examinations of posterior fossa or cervical spine). In the head, partial volume artifacts typically manifest as dark streaks or areas of hypoattenuation and are due to a nonlinear effect that has been described in reference 45. Figure 7 shows an example of a patient who underwent follow-up CT after surgical removal of a pituitary tumor. From the same scan data—four sections at 1.0-mm collimation—both 4.0-mm-thick images with a standard head kernel for soft-tissue evaluation and 1.0-mm-thick images with a bone kernel were reconstructed. For best image quality, the posterior fossa should be scanned with a collimated section width not larger than 1.25 mm, whereas wider collimation can be used in the supratentorial region (46).

SPIRAL SCANS AND IMAGE-RECONSTRUCTION TECHNIQUES

Spiral scanning is the method of choice for the majority of all multi–detector row CT examinations and requires more at-
tion holds for both single-section and multi-detector row CT. It shows whether data acquisition occurs with gaps \((p > 1)\) or with overlap \((p < 1)\) in the longitudinal direction. With 16 sections at 0.75-mm collimation and a table-feed of 18 mm per rotation, the pitch is \(p = 18/(16 \times 0.75) = 18/12 = 1.5\). With four sections at 1.0-mm collimation and a table-feed of 6 mm per rotation, the pitch again is \(p = 6/(4 \times 1) = 6/4 = 1.5\). In the early days of four-section CT, the term detector pitch had been additionally introduced, which accounts for the width of a single section in the denominator. For the sake of clarity and uniformity, the detector pitch should no longer be used.

**Definition of Spiral Pitch**

An important parameter for characterizing a spiral CT scan is the pitch. According to International Electrotechnical Commission specifications (34), the pitch \(p\) is given by \(p = TF/W\), where \(TF\) is the table feed per rotation, and \(W\) is the total width of the collimated beam. This definition holds for both single-section and multi-detector row CT because it is conceptually more demanding.

**Short Review of Single-Section Spiral CT Reconstruction**

Spiral CT requires an interpolation of the acquired measurement data in the longitudinal (through-plane) direction to estimate a complete CT data set at the desired plane of reconstruction. The most commonly used single-section spiral interpolation schemes are the 360° and 180° linear interpolation methods.

The 360° linear interpolation method exploits the 360° periodicity of the projection data \((1,2)\). For each projection angle, a linear interpolation is performed between those two projections on either side of the image plane that are positioned closest to the image plane and are 360° apart (ie, are measured in subsequent rotations). The 180° linear interpolation technique makes use of the fact that for each measurement ray, an interpolation partner is already available after approximately half a rotation \((47)\), when the x-ray tube and detector have exchanged positions. This is the so-called complementary ray. In spiral CT, z-axis resolution is determined not only by the collimated beam width (as in sequential scanning) but also by the effective section width, which is established in the spiral interpolation process. Usually, the effective section width is defined as the FWHM of the SSP. Effective section width increases with increasing pitch for both 360° and 180° linear interpolation, and longitudinal resolution degrades (Fig E3, radiology.rsna.org/cgi/content/full/2353040037/DC1). This is a consequence of the increasing longitudinal distance of the projections used for spiral interpolation. With 180° linear interpolation, the effective sections width equals the collimated section width at a pitch of 1, but effective section width equals 1.27 times the collimated width at a pitch of 2, so that a collimated 5-mm-thickness section is an actual 6.4-mm-thickness section at a pitch of 2. The image noise in single-section spiral CT is independent of the pitch if the tube current (in milliamperes) is left unchanged, and patient dose decreases with increasing pitch (see Appendix E2, radiology.rsna.org/cgi/content/full/2353040037/DC1).

Single-section spiral CT is based almost exclusively on 180° linear interpolation, owing to the narrower SSP of this algorithm, despite its increased susceptibility to artifacts and increased image noise. For the same milliampere-seconds setting, image noise is about 15% higher than that in sequential CT mode. Spiral artifacts gradually increase as pitch is increased. Spiral artifacts typically manifest as hyper- or hypointenuating “windmill” structures surrounding z-axis inhomogeneous high-contrast objects (eg, bones), which rotate when scrolling through a stack of images. Spiral artifacts are caused by the spiral interpolation process and can also be observed on multi-detector row CT images (see Fig 8). With single-section CT, scanning at a higher pitch is often used to reduce patient dose at the expense of section broadening—if the collimation is kept constant—and increased spiral artifacts. For CT angiographic applications in particular, it is more favorable to scan a given volume in a given time by using narrow collimation at a high pitch rather than wider collimation at a low pitch. The motivation for increasing pitch and reducing collimation is to improve longitudinal resolution by narrowing the SSP \((48)\).
The Cone-Angle Problem in Multi–Detector Row CT

Two-dimensional image-reconstruction approaches used in commercially available single-section CT scanners require all measurement rays that contribute to an image to run in a plane perpendicular to the patient's longitudinal axis. In multi–detector row CT systems, this requirement is violated. Figure 9 shows the geometry of a four-section scanner: The measurement rays are tilted by the so-called cone angle with respect to the center plane. The cone angle is largest for the sections at the outer edges of the detector, and it increases as the number of detector rows increases, if their width is kept constant. As a first approximation, the cone angle is neglected in multi–detector row CT reconstruction approaches. The measurement rays are treated as if they traveled perpendicular to the z-axis, and modified two-dimensional image-reconstruction algorithms are used. The data are then inconsistent, however, and produce cone-beam artifacts at high-contrast objects such as bones. It has been demonstrated that cone-beam artifacts can be tolerated if the maximum number of simultaneously acquired sections does not markedly exceed four (49). As a consequence, the image-reconstruction approaches of all commercially available four-section CT systems and of some systems with even more sections neglect the cone angle of the measurement rays.

MULTI-DETECTOR ROW SPIRAL CT RECONSTRUCTION APPROACHES THAT NEGLECT CONE-BEAM GEOMETRY

Multi–Detector Row 180° and 360° Linear Interpolation

The 360° and 180° linear interpolation single-section spiral reconstruction approaches can be extended to multi–detector row spiral scanning in a straightforward way (29,50,51). Both 360° and 180° multidetector linear interpolation methods are characterized by a projection-wise linear interpolation between two rays on either side of the image plane. The cone angle of the measurement rays is not taken into account. In the 360° linear interpolation spiral reconstruction approach, rays measured either at the same projection angle by different detector rows or in consecutive rotations of the scanner (ie, 360° apart) are used for spiral interpolation. In the 180° spiral reconstruction approach, both direct and complementary rays are considered. At the isocenter, direct and complementary rays interleave in the z-axis direction for selected pitch values. This way, the distance between measured samples is substantially reduced and equals half the collimated section width, which results in the desired narrow SSPs. Appropriate pitch values are 0.75 for four-section scanning (29) and 0.5625 or 0.9375 for 16-section scanning (50). The 180° and 360° multidetector linear interpolation approaches are schematically illustrated in Figure E4 (radiology.rsnaajnl.org/cgi/content/full/235/3/640#DC1) for the example of a four-section CT scanner.

In general, scanners that rely on 180° or 360° multidetector linear interpolation techniques and extensions thereof provide selected discrete pitch values to the user, such as 0.75 and 1.5 for four-section scanning (29) or 0.5625, 0.9375, 1.375, and 1.75 for 16-section scanning (50). These pitch values are intended to provide optimized sampling schemes in the longitudinal direction and, hence, optimized image quality.

The user has to be aware of pitch-dependent effective section widths. For low-pitch scanning (pitch of 0.75 for four sections and 0.5625 or 0.9375 for 16 sections), the effective section width approximates the collimated section width; for a 1.25-mm collimated section width, the resulting effective section width remains 1.25 mm. The narrow SSP, however, is achieved by using 180° multidetector linear interpolation reconstruction with conjugate interpolation at the price of increased image noise (29,50). For high-pitch scanning (pitch of 1.5 for four sections and 1.375 or 1.75 for 16 sections), the effective section width is approximately 1.27 times the collimated section width, and a 1.25-mm collimated section width results in a 1.5–1.6-mm effective section width.

When comparing dose and image noise for different pitch values, the widening of the SSP has to be taken into account. To obtain the same image noise as in a sequential scan with the same collimated section width, 0.73–1.68 times the dose (depending on spiral pitch) is required, with the lowest dose at the highest pitch (see reference 50). Some manufacturers provide a semiautomatic adaptation of the milliampere value to keep the image noise constant if the pitch is changed. In clinical practice, therefore, it is permissible to assume that scanners offering discrete optimized pitch values based on 180° and 360° multidetector linear interpolation techniques are comparable to single-section CT systems in some core aspects: At high pitch, the section widens and the longitudinal resolution degrades; at low pitch, the narrowest possible SSP (comparable to that of 180° single-section linear interpolation at pitch 1) can be obtained, but a higher dose is necessary to maintain the signal-to-noise ratio. Thus, as a take-home point, when one selects the scan protocol for a particular application, scanning at low pitch optimizes image quality and longitudinal resolution at a given collimation but at the expense of increased patient dose. To reduce patient dose, either milliampere settings should be reduced at low pitch values or high pitch values should be chosen.

Z-Filter Approaches

In a z-filter multi–detector row spiral reconstruction (51,52), the spiral interpo-
the image plane. Instead, all direct and complementary rays within a selectable distance from the image plane contribute to the image. The weighting function for the rays is selectable, which allows one to adjust both the functional form and the FWHM of the spiral SSP. Still, the cone angle is neglected. A representative example of a z-filter approach is the adaptive axial interpolation algorithm (51) implemented in Siemens CT scanners, which is illustrated in Figure E5 (radiology.rsna.org/cgi/content/full/2353040037/DC1). Another example is the “multislice cone-beam tomography,” or MUSCOT, algorithm (52) used by Toshiba. Z filtering allows the system to trade off z-axis resolution (the SSP) with image noise (which directly correlates with required dose).

With adaptive axial interpolation, the spiral pitch is freely selectable in the range 0.5–2.0, and the same effective section width, which is defined as the FWHM of the spiral SSP, is generated at all pitch values (7,51,53). Therefore, longitudinal resolution is independent of pitch, unlike single-section spiral CT and multi–detector row CT that relies on 180° and 360° linear interpolation (51,54). Figure E6 (radiology.rsna.org/cgi/content/full/2353040037/DC1) shows the SSPs of a 2-mm section (for four-section CT at 1-mm collimation) and MPRs of a spiral z-axis resolution phantom for selected pitch values. As a consequence of the pitch-independent spiral section width, the image noise for a fixed tube current (in milliamperes) would decrease as pitch is decreased, owing to the increasingly overlapping spiral acquisition. Instead, the user selects an “effective” milliamperes-seconds value, and the tube current is then automatically adapted to the pitch of the spiral scan to compensate for dose accumulation. The dose for fixed effective milliamperes-seconds is independent of the spiral pitch and equals the dose of a transverse scan with the same milliamperes-seconds setting (see Appendix E2, radiology.rsna.org/cgi/content/full/2353040037/DC1).

Thus, as a take-home point, unlike in single-section spiral CT a change in pitch does not result in a change in dose to the patient. Accordingly, the use of a higher pitch does not result in a dose saving, which is an important practical consideration with CT systems that rely on adaptive axial interpolation.

The intrinsic resolution of a multi–detector row spiral scan is determined by the choice of collimation (eg, four sections at 1.0 or 2.5 mm). Z filtering makes it possible to reconstruct images retrospectively with different section widths from the same raw CT data set. Only section widths equal to or larger than the section width of one active detector row can be obtained. In many cases, both thick sections for initial viewing and recording and thin sections for detailed diagnosis or as an input for advanced 3D postprocessing are routinely reconstructed.

The thinnest available section width is the collimated section width (1.0 mm for four sections at 1.0-mm collimation), which is created by using nonlinear spiral weighting functions at the expense of increased image noise and increased susceptibility to artifacts. Thus, as a take-home point, the thinnest available section should only be used for high-contrast applications such as high-spatial-resolution lung imaging. For general purpose scanning, a 1.25-mm section width for four-section CT at 1.0-mm collimation (and 3.0-mm section width for four sections at 2.5-mm collimation) is recommended as the most suitable trade-off between longitudinal resolution, image noise, and artifacts, in particular when thin sections are reconstructed as an input for 3D postprocessing such as for MPR, maximum intensity projection, or volume-rendering techniques. For a 1.25-mm spiral section width reconstructed from four-section CT at 1.0-mm collimation, 0.61–0.69 times the dose (depending only slightly on spiral pitch) is required to maintain the image noise of a sequential scan at the same collimation (see references 54,55). Unlike 180° and 360° multidetector linear interpolation, image noise is therefore practically independent of pitch at constant dose.

For a given collimation, such as four sections at 2.5 mm, image quality can be optimized with regard to spiral artifacts by lowering the pitch (56). Another means to reduce spiral artifacts is to use narrow collimation: A given section width (eg, 3.0 mm) can be obtained with different collimations, in this case four sections at 1.0 mm and at 2.5 mm. For optimum image quality, collimation that is narrow relative to the desired section width is preferable (51). Furthermore, a more rectangular SSP can be established. Figure 10a shows the SSPs of a 3.0-mm section for four-section CT at both 1.0- and 2.5-mm collimation. Figure 10b shows 3.0-mm transverse sections of a thorax phantom scanned with four-section CT at 2.5- and 1.0-mm collimation. Despite the higher pitch, the 3.0-mm image obtained at 1.0-mm collimation shows fewer artifacts. Similar to single-section spiral CT, narrow collimation at
high pitch is preferable to wide collimation at low pitch for artifact reduction.

Except for a minor dose increase due to the different relative contributions of the penumbral zones of the dose profile, scanning at narrow collimation does not result in higher radiation dose to the patient as long as the effective milliampereseconds level is kept constant. Narrow-collimation scanning should, therefore, be the protocol of choice for all applications that require 3D postprocessing as part of the clinical evaluation. In the clinical treatment of uncooperative or trauma patients or for protocols such as routine oncologic staging, the use of wider collimation can be considered. The best suppression of spiral artifacts is achieved by using both narrow collimation (relative to the desired section width) and reduced spiral pitch.

In general, more challenging clinical protocols, such as CT of the spine and of the skull base, are reliant on a combination of narrow collimation and low pitch. When multi-detector row spiral CT of the head is performed with narrow collimation, low pitch, and z-filter reconstruction of wider sections, the results are equivalent to those of traditional sequential CT. Figure 11 shows an example of a head scan performed with a four-section CT system in which a sequential image (two-section CT at 8 mm) and a spiral image (6-mm section width from four-section CT at 1-mm collimation) are compared in the same patient.

Some manufacturers who use a z-filter approach do not provide completely free selection of the spiral pitch but recommend a selection of fixed pitch values (eg, pitch of 0.625, 0.75, 0.875, 1.125, 1.25, 1.375 and 1.5 for four-section CT with the MUSCOT algorithm [52]) that are aimed at optimizing the z-axis sampling scheme and reducing spiral artifacts.

**MULTI-DETECTOR ROW SPIRAL RECONSTRUCTION APPROACHES THAT ACCOUNT FOR CONE-BEAM GEOMETRY**

**Overview of Cone-Beam Reconstruction Algorithms**

For CT scanners with 16 or more detector rows, modified reconstruction approaches that account for the cone-beam geometry of the measurement rays have to be considered. Some manufacturers (Toshiba, Philips) have extended the Feldkamp algorithm (57,58), an approximate 3D convolution back-projection reconstruction that was originally introduced for sequential scanning, to multisection spiral scanning (59,60). With this approach, the measurement rays are back projected into a 3D volume along the lines of measurement, accounting in this way for their cone-beam geometry. Three-dimensional back projection is computationally demanding and requires dedicated hardware to achieve acceptable image-reconstruction times. Other manufacturers use variations and extensions of nutating-section algorithms (61–66) for image reconstruction. These algorithms split the 3D reconstruction task into a series of conventional two-dimensional reconstructions on tilted intermediate image planes, in this way benefiting from established and very fast two-dimensional reconstruction techniques. Representative examples are AMPR (Siemens) (67,68) and the weighted hyperplane reconstruction (proposed by GE Medical Systems) (69,70) techniques.

**AMPR Method**

The AMPR approach (67,68) is an extension and generalization of the “advanced single-slice rebinning” (63,64) method. AMPR allows free selection of the spiral pitch with optimized dose utilization, which is beneficial for medical applications. With advanced single-slice rebinning, a partial scan interval (about 240° of scan data) is used for image reconstruction. The image planes are no longer perpendicular to the patient axis; instead, they are tilted to match the spiral path of the focal spot; see Figure 12 for a 16-section scanner at a pitch of 1.5. For every view angle in this partial scan interval, the focal spot is positioned in or near the image plane—that is, measurement rays running in or very close to the image plane are available. These conditions need to be fulfilled for a standard two-dimensional reconstruction. In a final z-axis reformation step, the traditional transverse images are calculated by interpolating between the tilted original image planes.

Advanced single-slice rebinning encounters its limitations when the spiral pitch is reduced to make use of the overlapping spiral acquisition and the resulting dose accumulation. The AMPR algorithm (67,68) addresses this problem: Instead of all available data being used for a single image, the data are distributed to several partial images on double-oblique image planes, which are individually adapted to the spiral path and fan out like the pages of a book (Fig 13, left). To ensure full dose utilization the number of partial images (“pages” in the book), as well as the length of the data interval per image, depend on the spiral pitch. The final transverse (or arbitrarily oriented) images are calculated by means of z-axis interpolation between the tilted partial image planes (Fig 13, right). The shape and the width of the z-axis interpolation functions are selectable. Different SSPs and different section widths can therefore be adjusted, so that z-axis resolution (SSP) can be traded off with image noise. The spiral pitch is freely selectable and the section width—and consequently the z-axis resolution—are independent of the pitch. The concept of effective milliampereseconds and automatic adaptation of the tube current to the pitch also apply to
AMPR (see Appendix E2, radiology.rsna.org/cgi/content/full/2353040037/DC1).

With the AMPR approach, sufficient image quality is obtained for all pitch values between 0.5 and 1.5 (68). Figure 14 shows transverse sections and MPRs of an anthropomorphized thorax phantom. Scan data for 16 sections at 0.75-mm collimation and pitch of 1 were reconstructed with 1-mm section width with z-folding, the AMPR algorithm, and 3D back projection. Neglecting the cone angle leads to artifacts at high-contrast objects and geometric distortions, particularly in MPRs (Fig 14, top). Both AMPR and 3D back projection restore the spatial integrity of the high-contrast objects, reduce cone-beam artifacts, and are fully equivalent for 16-section scanning. Recent studies have demonstrated the adequacy of extended versions of AMPR for medical CT systems with up to 64 detector rows (71).

The remaining artifacts in Figure 14 are spiral interpolation artifacts (windmill artifacts), not cone-beam artifacts. Windmill artifacts are not related to the cone-beam geometry and result from the finite width of the detector rows, which require interpolation between the rows for image reconstruction. Hence, windmill artifacts occur independent of the reconstruction approach. They are exaggerated in the mathematic phantom shown (Fig 14) and can be reduced by decreasing the pitch and/or increasing the reconstruction section width relative to the collimation (Fig 8). Figure 15 shows MPRs of a z-axis resolution phantom scanned with 16-section CT at 0.75-mm collimation and pitches of 0.75, 1.0, 1.25, and 1.5. Independent of the pitch, all cylinders down to 0.6 mm in diameter can be resolved, the MPRs are relatively free of geometric distortions, and the spatial integrity of the 3D image is maintained.

Multi-detector row spiral CT with AMPR is characterized by the same key properties as adaptive axial interpolation, which can be directly derived from information in the section on z-filter reconstruction presented earlier in this review. Thus, all recommendations regarding selection of collimation and pitch that were discussed there also apply for AMPR. In particular, a change in pitch does not result in a change in radiation exposure to the patient, and the use of higher pitch does not result in dose saving. Narrow collimation should be used whenever possible. With 16-section 0.75-mm-collimation CT, the thinnest available reconstruction section width of 0.75 mm is created by using nonlinear weighting functions at the z-axis image-reformation step, at the expense of increased image noise and increased susceptibility to artifacts. As a take-home point, this approach again should only be used for high-contrast applications such as high-spatial-resolution lung imaging. When thin sections are reconstructed as input for 3D postprocessing such as MPR, maximum intensity projection, or volume-rendering techniques, a 1.0-mm section width is recommended as the most suitable trade-off between longitudinal resolution, image noise, and artifacts.

**Weighted Hyperplane Reconstruction**

The weighted hyperplane reconstruction method, which has been described elsewhere (69,70), uses concepts related to AMPR but is derived differently. Similar to AMPR, 3D reconstruction is split into a series of two-dimensional reconstructions. Instead of reconstruction of traditional transverse sections, convex hyperplanes are proposed as the region of reconstruction. The increasing spiral overlap with decreasing pitch is handled by introducing subsets of detector rows, which are sufficient to reconstruct an image at a given pitch value. At pitch of 0.5625 with a 16-section scanner, the data collected by detector rows one to nine form a complete projection data set. Similarly, projections from detector rows two to 10 can be used to reconstruct another image at the same z-axis position. Projections from detector rows three to 11 yield a third image and so on. In a way, these “subimages” are related to the “book pages” of AMPR. The final image is

**Figure 12.** Left: Schematic 3D illustration of “advanced single-slice rebinning” approach for 16-section CT system at pitch of 1.5. Left: Curved line represents spiral path of the focal spot. Intermediate image plane is indicated by gradient-shaded rectangle and is no longer perpendicular to patient axis; instead, it is tilted to match spiral path of the focal spot. Right: Projection onto a plane containing the z-axis, where the spiral path is represented as a sinusoidal line. A partial scan interval (about 240°) is used for image reconstruction. For all view angles, focal spot is close to the image plane.

**Figure 13.** Illustration of AMPR approach. Left: First, multissection spiral CT data are used to reconstruct several partial images on double-oblique image planes, which are individually adapted to the spiral path. Partial images fan out like pages of a book. Right: Second, final images with full dose utilization are calculated with z-axis interpolation between tilted partial image planes.
based on a weighted average of the subimages. In the article by Hsieh et al (70), good image quality was demonstrated for a 16-section CT system (Lightspeed 16; GE Medical Systems) with which the weighted hyperplane reconstruction approach was used. By performing parameter optimizations, an optimal balance among various system performance parameters, such as noise, artifacts, and SSFs, can be achieved (72).

**ECG-SYNCHRONIZED SCAN AND IMAGE-RECONSTRUCTION TECHNIQUES**

One of the most exciting new applications of multi-detector row CT is the ability to image the heart and the cardiothoracic anatomy without motion artifacts. For ECG-synchronized scanning of the cardiothoracic anatomy, either ECG-triggered sequential scanning or ECG-gated spiral scanning can be used. In ECG-triggered sequential scanning, the heart volume is covered by subsequent transverse scans with a step-and-shoot technique. For each transverse scan, the number of images corresponds to the number of active detector sections. A partial scan data interval is acquired with a predefined temporal offset relative to the R waves of the patient’s ECG trace, which can be either relative (as a certain percentage of the R-R interval) or absolute (in milliseconds) and either forward or reverse (17). Some 16-section CT systems offer gantry rotation times shorter than 0.5 second (eg, 0.42, 0.40, or 0.37 second). In this case, temporal resolution can be as good as 0.21, 0.20, or 0.185 second (26,27).

With retrospective ECG gating, the heart volume is covered continuously by a spiral scan. The basic concepts for ECG-gated sequential scanning, the heart volume is covered by subsequent transverse scans with a step-and-shoot technique. For each transverse scan, the number of images corresponds to the number of active detector sections. A partial scan data interval is acquired with a predefined temporal offset relative to the R waves of the patient’s ECG trace, which can be either relative (as a certain percentage of the R-R interval) or absolute (in milliseconds) and either forward or reverse (17). Some 16-section CT systems offer gantry rotation times shorter than 0.5 second (eg, 0.42, 0.40, or 0.37 second). In this case, temporal resolution can be as good as 0.21, 0.20, or 0.185 second (26,27).

With retrospective ECG gating, the heart volume is covered continuously by a spiral scan. The basic concepts for ECG-gated sequential scanning, the heart volume is covered by subsequent transverse scans with a step-and-shoot technique. For each transverse scan, the number of images corresponds to the number of active detector sections. A partial scan data interval is acquired with a predefined temporal offset relative to the R waves of the patient’s ECG trace, which can be either relative (as a certain percentage of the R-R interval) or absolute (in milliseconds) and either forward or reverse (17). Some 16-section CT systems offer gantry rotation times shorter than 0.5 second (eg, 0.42, 0.40, or 0.37 second). In this case, temporal resolution can be as good as 0.21, 0.20, or 0.185 second (26,27).

With retrospective ECG gating, the heart volume is covered continuously by a spiral scan. The basic concepts for ECG-gated sequential scanning, the heart volume is covered by subsequent transverse scans with a step-and-shoot technique. For each transverse scan, the number of images corresponds to the number of active detector sections. A partial scan data interval is acquired with a predefined temporal offset relative to the R waves of the patient’s ECG trace, which can be either relative (as a certain percentage of the R-R interval) or absolute (in milliseconds) and either forward or reverse (17). Some 16-section CT systems offer gantry rotation times shorter than 0.5 second (eg, 0.42, 0.40, or 0.37 second). In this case, temporal resolution can be as good as 0.21, 0.20, or 0.185 second (26,27).

With retrospective ECG gating, the heart volume is covered continuously by a spiral scan. The basic concepts for ECG-gated sequential scanning, the heart volume is covered by subsequent transverse scans with a step-and-shoot technique. For each transverse scan, the number of images corresponds to the number of active detector sections. A partial scan data interval is acquired with a predefined temporal offset relative to the R waves of the patient’s ECG trace, which can be either relative (as a certain percentage of the R-R interval) or absolute (in milliseconds) and either forward or reverse (17). Some 16-section CT systems offer gantry rotation times shorter than 0.5 second (eg, 0.42, 0.40, or 0.37 second). In this case, temporal resolution can be as good as 0.21, 0.20, or 0.185 second (26,27).

With retrospective ECG gating, the heart volume is covered continuously by a spiral scan. The basic concepts for ECG-gated sequential scanning, the heart volume is covered by subsequent transverse scans with a step-and-shoot technique. For each transverse scan, the number of images corresponds to the number of active detector sections. A partial scan data interval is acquired with a predefined temporal offset relative to the R waves of the patient’s ECG trace, which can be either relative (as a certain percentage of the R-R interval) or absolute (in milliseconds) and either forward or reverse (17). Some 16-section CT systems offer gantry rotation times shorter than 0.5 second (eg, 0.42, 0.40, or 0.37 second). In this case, temporal resolution can be as good as 0.21, 0.20, or 0.185 second (26,27).

Figure 14. Transverse sections (left) and sagittal MPRs (right) of anthropomorphic thorax phantom. Scan data for 16-section CT at 0.75-mm collimation and pitch of 1 were reconstructed with 1.0-mm section width and z filtering that neglected the cone angle of measurement rays (top), with AMPR algorithm (middle), and with 3D back projection (bottom). Neglecting cone angle leads to artifacts at high-contrast objects, particularly in MPRs (top). Both AMPR (middle) and 3D back-projection (bottom) images reduce cone-beam artifacts and are fully equivalent for 16-section CT.

Figure 15. MPRs of z-axis resolution phantom at isocenter, scanned with 16-section CT at 0.75-mm collimation and pitches of 1.5, 1.25, 1.0, and 0.75 (MPR section width, 0.75 mm; increment, 0.4 mm). Phantom consists of polymerized methyl methacrylate plate with rows of cylindrical holes (diameters of 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.2, 1.5, 2.0, and 3.0 mm) aligned in the longitudinal direction. Independent of pitch, all cylinders down to 0.6 mm in diameter can be resolved.
struction mode (16,73–77), where $t_{rot}$ is the gantry rotation time of the CT scanner. With increased $N$, better temporal resolution is achieved but at the expense of slower volume coverage: Increased $N$ and slower patient heart rate require a reduction in spiral pitch.

Multisegment approaches rely on a complete periodicity of the heart motion, and these approaches encounter their limitations in patients with arrhythmia or a heart rate that changes during scan acquisition. Multisegment reconstruction may improve image quality in selected cases, but the reliability of good-quality image acquisitions with $N$-segment reconstruction is compromised with increases in $N$.

In general, clinical practice suggests the use of one segment at lower heart rates and two or more ($N \geq 2$) segments at higher heart rates. Use of single-segment versus multisegment reconstruction is integrated in the data acquisition process in a variety of ways, depending on the scanner type. One approach consists of automatic division of the partial-scan data segment into one or two subsegments, depending on the patient’s heart rate during acquisition (“adaptive cardio volume” algorithm [74]). With a different approach, single-segment partial-scan images are prospectively reconstructed as baseline images, followed by retrospective two-segment reconstruction for improved temporal resolution in patients with a higher heart rate. Yet another approach is prospective adjustment of the gantry rotation time to the heart rate of the patient to obtain an optimized temporal resolution for a multisegment reconstruction. Again, this approach requires a stable and predictable heart rate during scan acquisition.

Prospective ECG triggering combined with sequential step-and-shoot acquisition of transverse sections has the benefit of smaller patient dose than that of ECG-gated spiral scanning, because scan data are acquired only during the desired heart phases. However, this technique does not provide continuous volume coverage with overlapping sections, and misregistration of anatomic details cannot be avoided. Furthermore, reconstruction of images in different phases of the cardiac cycle for functional evaluation is not possible. Since ECG-triggered sequential scanning depends on a reliable prediction of the patient’s next R-R interval by using the mean of the preceding R-R intervals, the method encounters its limitations in patients with arrhythmia. To maintain the benefits of ECG-gated spiral CT but reduce patient dose, ECG-controlled dose modulation has been developed (42,43) (see earlier discussion).

The major improvements of 16-section CT, compared with established four-section scanners, include improved temporal resolution due to shorter gantry rotation time, better spatial resolution owing to submillimeter collimation, and considerably reduced scan acquisition times (26,27). The time to cover the entire heart volume (about 12 cm) with four-section CT at 1.0-mm collimation is about 40 seconds, which is at the limit for a scan requiring patient breath holding. ECG-gated CT of the entire thorax or the aorta is not possible within reasonable scan durations. For a 16-section CT system, the time to cover the entire heart volume with submillimeter collimation is about 15 seconds. With 16-section CT, coverage of the entire thorax (30 cm) can be completed in about 38 seconds at 0.75-mm collimation and in about 19 seconds at 1.5-mm collimation. ECG-gated examinations of extended cardiothoracic anatomy became feasible with 16-section CT, which lends itself to a spectrum of applications where suppression of cardiac pulsation is desired. Typical diagnostic pitfalls caused by transmitted cardiac pulsation can be avoided, such as an artificial intimal flap resembling dissection in the ascending aorta (79). Suppression of cardiac pulsation improves the assessment of paracardiac lung segments and allows confident exclusion of small peripheral pulmonary emboli in segmental and subsegmental arteries (80). In routine thoracic studies, which are not synchronized to the patient’s ECG signal, cardiac motion usually precludes the assessment of coronary bypass grafts. Figure 16 shows an example of an ECG-gated scan of the entire thorax for a patient with bypass grafts; this scan was acquired with 16-section CT at 0.75-mm collimation and 0.42-second gantry rotation.

**APPLICATIONS**

Clinical applications benefit from multidetector row CT technology in several ways: (a) shorter scan time (important for trauma patients and pediatric patients, CT angiography), (b) extended scan range (important for CT angiography, combined chest-abdomen scans such as in oncologic staging), and (c) improved longitudinal resolution (beneficial for all reconstructions,
particularly when 3D postprocessing is part of the clinical protocol).

Most protocols even benefit from a combination of all of these advantages. The near isotropic spatial resolution in routine examinations enables 3D renderings of diagnostic quality and oblique MPRs and maximum intensity projections of a resolution similar to that of the transverse images. The availability of multi–detector row CT technology has already begun to change the traditional perception of CT imaging. In CT, a distinction is traditionally made between longitudinal and in-plane resolution. With spiral CT, collimation is no longer the only factor used to determine longitudinal resolution; the spiral interpolation function also comes into play. This has been a first step toward decoupling the image section width from the beam width as determined by the collimation. Multi–detector row CT now allows reconstruction of arbitrary section widths from a given collimation by using z-filter techniques, as long as the desired section width is not smaller than the collimation. The potential to trade off z-axis resolution and image noise for the same data set is the most important benefit of z-filter reconstruction. In many applications, data acquisition with narrow collimation is recommended independently of the section width desired for primary viewing.

The distinction between longitudinal and in-plane resolution will gradually become a historical curiosity, and the traditional transverse section will lose its clinical importance. In its place, interactive viewing and manipulation of isotropic volume images will become commonplace, with only the key sections or views in arbitrary directions recorded and stored.

Spiral scanning with 16 submillimeter sections, in particular, represents a breakthrough on the way to true isotropic resolution for routine clinical applications. Improved longitudinal resolution is combined with considerably reduced scan times, which facilitate examinations in uncooperative patients and reduce the amount of contrast material needed (although optimized contrast material protocols are also required).

Furthermore, new clinical applications are evolving as a result of the increased speed of volume scanning. CT angiography of the carotid arteries and the circle of Willis with 16 sections at 0.75-mm collimation, 0.5-second rotation time, and pitch of 1.5 requires only 9 seconds for a scan range of about 300 mm (with table feed of 36 mm/sec). For the first time, true arterial phase imaging of the entire carotid artery with high spatial resolution can be performed. Clinical practice indicates the potential of 16-section CT angiography to replace conventional interventional angiography in the evaluation of carotid artery stenosis (81). Evaluation of the supraaortic vessels with 16-section CT is particularly useful in emergency situations, since CT allows a quick diagnosis with optimized patient access.

For patients suspected of having ischemic stroke, both the status of the vessels supplying the brain and the location of the intracranial occlusion can be assessed during the same examination (82). Brain perfusion CT can be performed by using the same modality, with the goal of differentiating irreversibly damaged brain tissue from reversibly impaired tissue at risk. The combined use of nonenhanced CT, perfusion CT, and CT angiography may rapidly provide comprehensive information regarding the extent of ischemic damage in patients with acute stroke (46).

Scan acquisition of the entire thorax (350 mm) with submillimeter collimation can now be performed in approximately 11 seconds. Owing to the short breath-hold time, central and peripheral pulmonary embolism can be reliably and accurately diagnosed even in severely dyspneic patients with limited ability to cooperate (11,83). Meanwhile the use of multi–detector row CT for a combined diagnosis of pulmonary embolism and deep venous thrombosis has been clinically established (83). Both a native and a contrast-enhanced scan of the thorax can be obtained within the same breath hold for matching of both image volumes as a basis for investigational applications such as lung perfusion imaging.

Sixteen-section CT enables whole body angiographic studies with submillimeter resolution in a single breath hold. Also, 16-section CT yields the same morphologic information as invasive angiography (84,85). CT angiography of the chest and abdomen with submillimeter collimation can be completed in about 17 seconds for a scan range of 600 mm (Fig 17). When true isotropic resolution is not required, the use of 16-section CT at 1.25- or 1.5-mm collimation enables even shorter examination times or extended scan ranges (eg, for onologic screening, trauma cases, or CT angiography). Whole-body 16-section CT angiography with 1500-mm scan range, 1.5-mm collimation, 0.5-second rotation time, and pitch of 1.25 (table feed, 60 mm/sec) can be completed in only 26 seconds.

ECG-gated cardiac scanning benefits from both improved temporal resolution and improved spatial resolution. Detection and characterization of coronary
plaque, even in the presence of severe calcifications, greatly benefits from the increased robustness of the technology. Sixteen-section CT allows assessment of small, peripheral coronary segments that, until now, could not be evaluated. In a recent study (86) in which coronary CT angiography with a 16-section system was investigated in 59 patients, 86% specificity and 95% sensitivity were demonstrated for identification of significant coronary artery stenosis. None of the patients had to be excluded, unlike in previous studies that were based on less-advanced scanner technology. Other investigators have reported similar results (87). Early clinical experience with 0.37-second gantry rotation indicates improved image quality due to reduced cardiac motion and increased clinical robustness at higher heart rates, which thereby potentially reduce the number of patients who require heart rate control (Fig 18).

FUTURE DIRECTION OF MULTI-DETECTOR ROW CT

Sixteen-section CT, which has become widely available, enables truly isotropic submillimeter imaging for virtually any application. In the case of cardiac imaging, 16-section CT sets today’s benchmark in spatial resolution for noninvasive coronary artery imaging. Motion artifacts in patients with a higher heart rate remain the most important challenge for multi–detector row coronary CT angiography, although diagnostic image quality can be achieved in most cases by administering β-blockers to such patients. Improved temporal resolution is desirable in the future to prevent the need for heart rate control. Increased gantry rotation speed, rather than multi-segment reconstruction, appears to be preferable for robust clinical performance. Obviously, substantial development efforts are needed to account for the notable increase in mechanical forces (about 17 g for 0.42-second rotation, >33 g for 0.3-second rotation) and increased data transmission rates. A rotation time of less than 0.2 second (mechanical force > 75 g), which is required to provide a temporal resolution of less than 100 msec independent of heart rate, appears to be beyond today’s mechanical limits. An alternative to further increases in rotation speed is to reconsider the scanner concept with multiple tubes and multiple detectors that had already been described in the early years of CT (88,89).

Owing to its ease of use and its widespread availability, general-purpose CT continues to evolve into the most widely used diagnostic modality for routine examinations, especially in emergency situations or for oncologic staging. CT primarily provides morphologic information; in combination with other modalities, however, functional and metabolic information can also be obtained (90). Therefore, combined systems for obtaining comprehensive structural and functional diagnoses will gain increasing importance in the near future.

The combination of state-of-the-art multi–detector row CT with positron emission tomographic (PET) scanners, for instance, opens a wide spectrum of applications ranging from oncologic staging to comprehensive cardiac examinations. The clinical potential of these scanners is currently being evaluated (91). Reconstruction of the CT images in a sufficient field of view without truncation of anatomic structures (eg, arms) is a prerequisite for adequate attenuation correction of the PET images. An enlarged field of view of up to 70 cm can be realized by extrapolating from the measured CT data. Pertinent algorithms can be found in, for example, reference 92. Figure 19 shows MPRs from CT images in a 46-year-old man with renal cancer who had undergone nephrectomy and chemotherapy, with PET images superimposed. Areas with increased metabolism are enhanced, and a metastatic mediastinal lymph node can be identified, which supports the notion of PET as adding a “new contrast agent” to CT.

Systems that combine CT and single-photon emission computed tomography are another promising modality. Potential applications are currently being investigated and range from the localization of parathyroid lesions (93) and heterotopic splenic tissue (94) to detection of recurrent nasopharyngeal carcinomas (95) to imaging of aortic prosthesis infection (96).

CT virtual simulation is gaining increasing importance with a more widespread adoption in 3D conformal and intensity-modulated radiation therapy. With general-purpose CT systems that have a gantry opening with a typical diameter of 70 cm, some patients (eg, women with breast cancer) cannot always be scanned in the treatment position. Such applications, along with interventional procedures and trauma protocols, will be facilitated by CT systems with a larger bore (97). Recently, concepts have been introduced for four- and 16-section CT scanners with a bore diameter of up to 85 cm and a reconstruction field of up to 82 cm, owing to image reconstruction based on extrapolation. These systems will probably gain considerable importance in the near future, in particular with regard to the dramatically increasing number of severely obese patients in the Western countries.

For general purpose CT, we will witness a moderate increase in the number of simultaneously acquired sections in the near future. A new generation of CT systems with 32, 40 and—in combination with refined z-axis sampling techniques—64 simultaneously acquired sections are currently being introduced. However in contrast to
the transition from single-section to four- and 16-section CT, clinical performance will improve only incrementally with further increases in the number of detector rows. The achievable clinical benefit will have to be carefully considered in the light of the necessary technical efforts and the cost. Clinical progress can more likely be expected from further improvements in spatial resolution rather than from an increase in the volume-coverage speed. In clinical reality, the latter has only rarely been a limiting factor since the introduction of 16-section CT. As soon as all relevant examinations can be performed in a comfortable breath hold of not more than 10 seconds, a further increase in the number of sections will not provide a substantial clinical benefit.

At this point, a qualitative enhancement of CT that allows new clinical applications may again bring substantial clinical benefit. A scanner with 256 0.5-mm detector elements has been proposed by one manufacturer and appears to be conceptually promising, but this system is still in the prototype stage. Prototype systems by other vendors use cesium iodide–amorphous silicon flat-panel detector technology that was originally used for conventional angiography, which is limited in terms of low contrast resolution and imaging speed. Owing to the intrinsic slow signal decay of flat-panel detectors, rotation times of at least 20 seconds are needed to acquire a sufficient number of projections (600 projections). The spatial resolution of such systems is excellent, though, because of the small detector pixel size. Excessive dose requirements to date, however, preclude the examination of larger objects. Initial experimental results are thus limited to small high-contrast objects such as joints, the inner ear, or contrast material-filled vessel specimens.

Figure 20 shows a prototype set-up, where a flat-panel detector was incorporated into a standard CT gantry (Somatom Sensation 16; Siemens). The detector covers a 250 × 250 mm scan field of view, and the pixel size is 0.25 × 0.25 mm, both measured at the center of rotation. Figure 21 shows volume renderings of a heart specimen (80 kV, 20 mA, 20-second gantry rotation) that demonstrate excellent spatial resolution, which enables visualization of even very small side branches of the coronary artery tree. The combination of area detectors that provide sufficient image quality with fast gantry rotation speed will be a promising technical concept for medical CT systems. The vast spectrum of potential applications may bring about another...
quantum leap in the evolution of medical CT imaging; however such systems will probably not be available in the near future.

References


75. Toth TL, Simoni P. A quantitative measure of distortion caused by the artifact (abstr). Radiology 2001; 221(P):217.


Performance Benchmarks for Diagnostic Mammography

PURPOSE: To evaluate a range of performance parameters pertinent to the comprehensive auditing of diagnostic mammography examinations, and to derive performance benchmarks therefrom, by pooling data collected from large numbers of patients and radiologists that are likely to be representative of mammography practice in the United States.

MATERIALS AND METHODS: Institutional review board approval was met, informed consent was not required, and this study was Health Insurance Portability and Accountability Act compliant. Six mammography registries contributed data to the Breast Cancer Surveillance Consortium (BCSC), providing patient demographic and clinical information, mammogram interpretation data, and biopsy results from defined population-based catchment areas. The study involved 151 mammography facilities and 646 interpreting radiologists. The study population included women 18 years of age or older who underwent at least one diagnostic mammography examination between 1996 and 2001. Collected data were used to derive mean performance parameter values, including abnormal interpretation rate, positive predictive value (for abnormal interpretation, biopsy recommended, and biopsy performed), cancer diagnosis rate, invasive cancer size, and the percentages of minimal cancers, axillary node-negative invasive cancers, and stage 0 and I cancers. Additional benchmarks were derived for these performance parameters, including 10th, 25th, 50th (median), 75th, and 90th percentile values.

RESULTS: The study involved 332,926 diagnostic mammography examinations. Mean performance parameter values were abnormal interpretation rate, 8.0%; positive predictive value for abnormal interpretation, 31.4%; positive predictive value for biopsy recommended, 31.5%; positive predictive value for biopsy performed, 39.5%; cancer diagnosis rate, 25.3 per 1000 examinations; invasive cancer size, 20.2 mm; percentage of minimal cancers, 42.0%; percentage of axillary node-negative invasive cancers, 73.6%; and percentage of stage 0 and I cancers, 62.4%.

CONCLUSION: The presented BCSC outcomes data and performance benchmarks may be used by mammography facilities and individual radiologists to evaluate their own performance for diagnostic mammography as determined by means of periodic comprehensive audits.

© RSNA, 2005

Within the United States, Food and Drug Administration regulation requires limited auditing of clinical outcomes for all screening and diagnostic mammograms assessed as either suspicious for malignancy or highly suggestive of malignancy (1). More comprehensive auditing is performed by many mammography facilities in both the United States and other countries. It is generally accepted that auditing is a useful quality assurance procedure that provides performance parameter feedback both to mammography facilities and to individual interpreting radiologists (2–4). Outcomes have been extensively reported for screening mammography, leading to the publication of several performance benchmarks that are currently in widespread use (5–8).

Recent reports indicate significantly different clinical outcomes for diagnostic compared with screening mammography, the diagnostic examinations being defined as those performed for indications other than the periodic screening of asymptomatic women (9,10).
However, these reports involve only a moderate number (approximately 10,000) of examinations, performed at a single institution, which may limit generalization of the observed findings. There also is evidence of considerable variability in performance parameters among interpreting radiologists; this is probably related to a complex interaction of experience and expertise (7,11–17). For diagnostic mammography, the published reports on performance variability are based on data from only 10 interpreting radiologists (9) and are described by investigators as likely being at the ends of the spectrum of performance rather than representing average performance (18,19). Clearly, there is need for more robust data on the clinical outcomes of diagnostic mammography examinations.

The Breast Cancer Surveillance Consortium (BCSC) is a group of mammography registries from geographically diverse areas in the United States, funded by the National Cancer Institute, that collects patient demographic and clinical information, mammogram interpretation data, and biopsy results in the defined catchment areas of its participating facilities (20). The primary purpose of the BCSC is to collect data from diverse population-based settings to examine the practice and performance of mammography throughout the United States. Six BCSC registries collect data on the full range of clinical outcomes pertinent to the comprehensive auditing of mammography performance parameters. Pooling of the data from these registries provides by far the largest reported experience involving diagnostic mammography practice, from which reasonable and realistic performance benchmarks may be derived. Thus, the purpose of our study was to evaluate a range of performance parameters pertinent to the comprehensive auditing of diagnostic mammography examinations, and to derive performance benchmarks therefrom, by pooling data collected from large numbers of patients and radiologists that are likely to be representative of mammography practice in the United States.

MATERIALS AND METHODS

Data Sources

Data were collected from six BCSC registries: Carolina Mammography Registry (Chapel Hill, NC), Group Health Cooperative (Seattle, Wash), New Hampshire Mammography Network (Lebanon, NH), New Mexico Mammography Project (Albuquerque, NM), San Francisco Mammography Registry (San Francisco, Calif), and Vermont Breast Cancer Surveillance System (Burlington, Vt). To determine cancer outcomes, each registry links its data to a state tumor registry or to the Surveillance Epidemiology and End Results program. The North American Association of Cancer Registries maintains statistics for each of the cancer registries. All cancer registries were found to be at least 94.3% complete, except for the Vermont registry, which did not have statistics available. To supplement cancer registry information, each registry is also linked to pathology databases. Each registry obtains annual approval from its institutional review board to collect and maintain registry data. Individual informed consent has not been required by the institutional review boards because of the strict maintenance of anonymity and the observational nature of the study. Our study was compliant with the Health Insurance Portability and Accountability Act. Linkage procedures follow protocols specifically designed to preserve patient confidentiality (21).

Each registry and the BCSC Statistical Coordinating Center, or SCC, have developed data management and quality control procedures that result in high-quality data collection that is comparable across registries. Prior to sending data to the SCC, data quality checks are conducted at each registry by using their own procedures, such as manual validation of a random sample of records, double data entry, monitoring of facility volume over time, and comparing different sources (eg, cancer registry and pathology databases) for consistency. After each annual data submission from the individual registries, the SCC performs additional quality checks of the pooled data by flagging coding errors and by comparing information across registries and over time for consistency and outlying values. The SCC also conducts biennial site visits to each registry and annual meetings involving data managers from the registries to review data management and quality control procedures, as well as to check data quality.

Across the six BCSC registries, 151 mammography facilities contributed to the pooled data. This represents 1.5% of the approximately 10,000 Food and Drug Administration–certified mammography facilities in the United States in 2000. The pool of data contains diagnostic mammogram interpretations made by 646 radiologists. We have been unable to find reliable estimates of how many radiologists met Food and Drug Administration requirements to read mammograms in 2000.

Two authors (E.A.S. and D.L.M., by consensus) compared the demographic makeup (rural-urban mix, race, ethnicity, education level, and socioeconomic status) of the population living in the catchment areas of the six BCSC registries included in our study with that of the entire U.S. population by using 2000 census data. To describe the BCSC population, we included census data from all counties in which there was a participating mammography facility.

Subjects

The study population included women 18 years of age or older who had undergone at least one diagnostic mammography examination during the years 1996–2001. Mammography examinations performed after December 2001 were excluded to ensure that there was a period of at least 12 months following examination during which cancer could be diagnosed and a period of an additional 24 months for reporting cancer data to tumor registries. Cancer reporting was at least 95% complete.

Diagnostic examinations are designed to solve specific problems and almost always include as many mammograms as are necessary to make a Breast Imaging Reporting and Data System (BI-RADS) final assessment, as well as case management recommendations. However, under certain circumstances a diagnostic examination is occasionally assessed as “incomplete—needs additional imaging evaluation” (BI-RADS assessment category 0). In this study, 15,971 (4.6%) of 348,897 examinations were given a category 0 assessment. For this study, when one or more diagnostic examinations followed an initial diagnostic examination that was assessed as category 0, all examinations up to and including the first examination with a non-zero assessment (within 180 days) were treated as a single observation. The date of and indication for examination were considered to be those from the initial examination (the first one with a category 0 assessment). However, we used the assessment and management recommendations from the first non-zero assessment and attributed the observed clinical outcomes to the radiologist who made that first non-zero assessment. If there was no non-zero assessment within 180 days, all of the examinations were excluded (10,662 of 348,897 examinations, 3.1%).
Data Collection Procedures

Across all BCSC registries, mammography patients complete a questionnaire that requests medical history and demographic data (including date of most recent mammography examination, family history of breast cancer, previous percutaneous or surgical biopsies, personal history of breast cancer, and description of breast symptoms experienced within the past 3 months). Women were considered to have a family history of breast cancer if they reported having at least one female first-degree relative (mother, sister, or daughter) with breast cancer. Women were considered to have a personal history of breast cancer if they self-reported previous breast cancer or if there was evidence of previous breast cancer in the cancer registry or pathology database. Each woman was considered to have undergone a previous mammography examination if she self-reported a history of prior mammography or there was data from a prior mammography examination in the BCSC database.

Diagnostic mammography is performed for a variety of problem-solving indications, including work-up of abnormalities detected at screening mammography, evaluation of abnormalities found at clinical examination, and short-interval follow-up examinations both for probably benign lesions and for cancer patients recently treated by means of breast preservation surgery. Other special breast problems, such as the presence of implants or the evaluation of extent of disease for a known malignancy may also represent indications for diagnostic mammography. Across all BCSC registries, the interpreting radiologist prospectively classifies each diagnostic mammography examination into one of three categories: additional work-up of an abnormality detected at screening examination, short-interval follow-up, or evaluation of a breast problem. We further subdivided the “evaluation of a breast problem” category according to whether the patient indicated the presence of a palpable lump on the medical history questionnaire that she completed at the time of her mammography examination, because results in previous published reports have shown substantially different clinical outcomes based on this approach (9,10). If the self-reported response concerning palpable lump was missing for a given examination, we used the first nonmissing response, if any, within the previous 90 days.

The mammography registry also collects data on image interpretation, including management recommendations and the BI-RADS assessment category assigned by the interpreting radiologist. An example of such discordance is a finding assessed as suspicious, accompanied by the recommendation for anything other than biopsy or surgical consultation. In this study, we have chosen to analyze mammography interpretation data by using both BI-RADS assessments and management recommendations to parallel the BI-RADS auditing approaches that will be discussed in the paragraph concerning positive predictive value (PPV) calculations.

Mammography patients were considered to have breast cancer if a state tumor registry, Surveillance Epidemiology and End Results program registry, or pathology database indicated the diagnosis of invasive carcinoma or ductal carcinoma in situ (DCIS) within 12 months after a diagnostic mammography examination. Additional data collected for breast cancer cases included tumor size (for invasive cancers), axillary lymph node status (for invasive cancers), and American Joint Committee on Cancer stage (26).

Outcome Measures

A positive (abnormal) assessment at diagnostic mammography was defined as an overall assessment of suspicious for or highly suggestive of malignancy. Cancer diagnosis rate was defined as the number of cancer cases identified at mammography (mammographically true-positive) divided by the total number of diagnostic mammography examinations. A true-positive case is one that is followed by the diagnosis of invasive breast cancer or DCIS within 12 months of a positive assessment at diagnostic mammography. Conversely, a case was considered to be false-positive if results at diagnostic mammography were negative for malignancy but the patient developed breast cancer within 12 months of mammography.

**TABLE 1**

Demographics for the Study Population Compared with Those for the Entire U.S. Population

<table>
<thead>
<tr>
<th>Demographic</th>
<th>Study Population*</th>
<th>U.S. Population†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total population in selected counties</td>
<td>11 874 535</td>
<td>281 421 906</td>
</tr>
<tr>
<td>Rural-urban mix (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>23.0</td>
<td>21.0</td>
</tr>
<tr>
<td>Urban</td>
<td>77.0</td>
<td>79.0</td>
</tr>
<tr>
<td>Race (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>82.7</td>
<td>84.9</td>
</tr>
<tr>
<td>African American</td>
<td>9.7</td>
<td>10.8</td>
</tr>
<tr>
<td>Other</td>
<td>7.5</td>
<td>4.3</td>
</tr>
<tr>
<td>Hispanic ethnicity (%)</td>
<td>6.3</td>
<td>7.3</td>
</tr>
<tr>
<td>No high school degree (%)</td>
<td>6.3</td>
<td>7.3</td>
</tr>
<tr>
<td>Economic status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Living in poverty (%)</td>
<td>11.2</td>
<td>12.4</td>
</tr>
<tr>
<td>Unemployed (%)</td>
<td>3.7</td>
<td>4.0</td>
</tr>
<tr>
<td>Median family income ($)</td>
<td>53 933</td>
<td>51 197</td>
</tr>
</tbody>
</table>

* Based on 2000 census data for all counties in which there was a mammography facility that contributed data to this study.
† Based on 2000 census data for the entire U.S. population.
‡ For women age 40 years and older.
§ For women age 25 years and older.
mammography were interpreted as positive and no breast cancer was diagnosed within the next 12 months.

In this article, we do not report on measures of sensitivity or specificity because such measures require the enumeration of false-negative and true-negative cases, respectively, involving tumor registry linkage data that are not generally available to mammography facilities or individual practicing radiologists. These measures, as well as other data beyond the scope of this article, are available to interested readers on the BCSC Web site (breastscreening.cancer.gov/benchmarks/diagnostic).

Statistical Analysis
Calculations of PPV were made by dividing the number of true-positive cases by the sum of true-positive and false-positive cases. Three separate PPV calculations were performed by using BI-RADS methods: PPV1, probability of cancer after positive mammography interpretation; PPV2, probability of cancer after recommendation for biopsy or surgical consultation, following positive mammography interpretation; and PPV3, probability of cancer after biopsy, following positive mammography interpretation and a recommendation for biopsy or surgical consultation. “Biopsy” included the performance of any type of biopsy (fine-needle aspiration, core, or surgical biopsy), whether or not imaging guidance was used to perform the biopsy.

Because the principal aim of this study was to provide outcomes data to be used for the derivation of clinically relevant performance benchmarks, we have chosen to provide only descriptive statistics such as those enumerated previously. Because benchmarks are more meaningful if they indicate ranges of performance as well as arithmetic means, we also have calculated percentile values for selected outcomes. For example, the combination of 25th and 75th percentile values defines the range within which the middle 50% of performance is found, and the combination of 10th and 90th percentile values defines the range within which the middle 80% of performance is found. To reduce the number of radiologists involved data from 332 926 diagnostic mammography examinations, and 150 494 (45.2%) were performed to evaluate a breast problem.

Demographic Factors
The demographic makeup of the population living in the catchment areas of the six BCSC sites in our study is compared with that for the entire U.S. population in Table 1. There were only slight differences, none greater than five per-
percentage points, between our study population and the U.S. population. Our study population was slightly more rural, contained slightly fewer African American and Hispanic women, was slightly more educated, and had a slightly higher median family income than the entire U.S. population.

Previous reports have shown that clinical outcomes for screening mammography are affected by several other demographic factors, specifically age, family

---

**TABLE 3**

<table>
<thead>
<tr>
<th>Abnormal Interpretation</th>
<th>Abnormality Detected at Screening Mammography</th>
<th>Short-Interval Follow-up</th>
<th>Evaluation of Breast Problem</th>
<th>No Lump or Lump Unknown</th>
<th>Palpable Lump</th>
<th>All Diagnostic Examinations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abnormal interpretation rate (%)</td>
<td>12.3</td>
<td>3.4</td>
<td>5.7</td>
<td>10.5</td>
<td>8.0</td>
<td></td>
</tr>
<tr>
<td>Abnormal interpretations</td>
<td>12 431</td>
<td>2804</td>
<td>5120</td>
<td>6421</td>
<td>26 776</td>
<td></td>
</tr>
<tr>
<td>All examinations</td>
<td>101 147</td>
<td>81 285</td>
<td>89 593</td>
<td>60 901</td>
<td>332 926</td>
<td></td>
</tr>
<tr>
<td>PPV, (abnormal interpretation) (%)</td>
<td>25.1</td>
<td>24.4</td>
<td>31.7</td>
<td>46.5</td>
<td>31.4</td>
<td></td>
</tr>
<tr>
<td>Cancer</td>
<td>3120</td>
<td>685</td>
<td>1622</td>
<td>2984</td>
<td>8411</td>
<td></td>
</tr>
<tr>
<td>Abnormal interpretation</td>
<td>12 431</td>
<td>2804</td>
<td>5120</td>
<td>6421</td>
<td>26 776</td>
<td></td>
</tr>
<tr>
<td>PPV, (biopsy recommended) (%)</td>
<td>24.6</td>
<td>24.6</td>
<td>32.9</td>
<td>48.0</td>
<td>31.5</td>
<td></td>
</tr>
<tr>
<td>Cancer</td>
<td>2732</td>
<td>577</td>
<td>1357</td>
<td>2507</td>
<td>7173</td>
<td></td>
</tr>
<tr>
<td>Abnormal interpretation</td>
<td>11 099</td>
<td>2347</td>
<td>4130</td>
<td>5223</td>
<td>22 799</td>
<td></td>
</tr>
<tr>
<td>PPV, (biopsy performed) (%)</td>
<td>30.3</td>
<td>32.3</td>
<td>43.2</td>
<td>59.4</td>
<td>39.5</td>
<td></td>
</tr>
<tr>
<td>Cancer</td>
<td>2724</td>
<td>576</td>
<td>1348</td>
<td>2495</td>
<td>7143</td>
<td></td>
</tr>
<tr>
<td>Abnormal interpretation</td>
<td>8976</td>
<td>1782</td>
<td>3120</td>
<td>4198</td>
<td>18 076</td>
<td></td>
</tr>
</tbody>
</table>

**Note.**—Except where indicated, data are numbers of examinations.
* At assessment, either suspicious or highly suggestive of malignancy (BI-RADS category 4 or 5).
† Abnormal interpretation and recommendation for either biopsy or surgical consultation.
‡ Abnormal interpretation, biopsy recommended, and biopsy results available.

---

**TABLE 4**

<table>
<thead>
<tr>
<th>Cancer Data</th>
<th>Abnormality Detected at Screening Mammography (n = 101 147)</th>
<th>Short-Interval Follow-up (n = 81 285)</th>
<th>Evaluation of Breast Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cancers</td>
<td>3120</td>
<td>685</td>
<td>1622</td>
</tr>
<tr>
<td>Cancer diagnosis rate (per 1000)</td>
<td>30.8</td>
<td>8.4</td>
<td>18.1</td>
</tr>
<tr>
<td>Histologic finding*</td>
<td>DCIS</td>
<td>840 (26.9)</td>
<td>210 (30.7)</td>
</tr>
<tr>
<td>Invasive carcinoma</td>
<td>2280 (73.1)</td>
<td>475 (69.3)</td>
<td>1362 (84.0)</td>
</tr>
<tr>
<td>Invasive cancer size (mm)†</td>
<td>1–5</td>
<td>275 (13.9)</td>
<td>62 (14.9)</td>
</tr>
<tr>
<td></td>
<td>6–10</td>
<td>635 (32.0)</td>
<td>133 (32.0)</td>
</tr>
<tr>
<td></td>
<td>11–15</td>
<td>532 (26.8)</td>
<td>98 (23.6)</td>
</tr>
<tr>
<td></td>
<td>16–20</td>
<td>246 (12.4)</td>
<td>57 (13.7)</td>
</tr>
<tr>
<td></td>
<td>&gt;20</td>
<td>294 (14.9)</td>
<td>66 (15.9)</td>
</tr>
<tr>
<td>Unknown</td>
<td>298</td>
<td>59</td>
<td>201</td>
</tr>
<tr>
<td>Mean</td>
<td>14.3</td>
<td>14.4</td>
<td>20.9</td>
</tr>
<tr>
<td>Median</td>
<td>11</td>
<td>11</td>
<td>17</td>
</tr>
<tr>
<td>Minimal cancer‡</td>
<td>1750 (62.0)</td>
<td>405 (64.7)</td>
<td>529 (37.2)</td>
</tr>
<tr>
<td>Axillary lymph node status (invasive cancers)§</td>
<td>Negative</td>
<td>1745 (84.2)</td>
<td>360 (86.7)</td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>327 (15.8)</td>
<td>55 (13.3)</td>
</tr>
<tr>
<td>Unknown</td>
<td>208</td>
<td>60</td>
<td>154</td>
</tr>
<tr>
<td>Cancer stage‖</td>
<td>0</td>
<td>840 (30.0)</td>
<td>210 (34.8)</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>1448 (51.7)</td>
<td>285 (47.2)</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>461 (16.5)</td>
<td>92 (15.2)</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>42 (1.5)</td>
<td>10 (1.7)</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>10 (0.4)</td>
<td>7 (1.2)</td>
</tr>
<tr>
<td>Unknown</td>
<td>319</td>
<td>81</td>
<td>217</td>
</tr>
</tbody>
</table>

* Numbers in parentheses are percentages of all cancers.
† Numbers in parentheses are percentages of invasive cancers of known size.
‡ Defined as DCIS or invasive cancer 10 mm or smaller; numbers in parentheses are percentages of DCIS and invasive cancers of known size.
§ Numbers in parentheses are percentages of invasive cancers of known nodal status.
‖ Numbers in parentheses are percentages of DCIS and invasive cancers of known stage.
history of breast cancer, personal history of breast cancer, and mammography performed previously (3–5, 27–30). Because it is likely that these factors will also affect the outcomes for diagnostic mammography, appropriate data are presented for our study population (Table 2). Those who use the benchmarks derived from observed outcomes in this study are advised to compare the clinical demographic factors of their own patient population with those reported here. Those who are able to break down their own audit data as a function of one or more of these factors should consult the BCSC Web site, where such data breakdowns are provided for selected observed outcomes in the study.

Abnormal Interpretations

For our entire study population, results from 26 776 (8.0%) of the 332 926 diagnostic mammography examinations were interpreted as abnormal (positive). Among these examinations with abnormal results, biopsy was recommended in 22 799 (6.8%) cases, and biopsy was actually performed in 18 076 (5.4%) cases. For our study, PPV₁ (abnormal interpretation) was 31.4%, PPV₂ (biopsy recommended) was 31.5%, and PPV₃ (biopsy performed) was 39.5%. Table 3 shows these data stratified by indication for diagnostic mammography. All three PPVs are higher for examinations performed to evaluate a breast problem than for examinations performed as work-up of screening-detected abnormalities or for short-interval follow-up, with the highest PPVs observed for the subset of “breast problem” examinations for which the patient reported a palpable lump.

Breast Cancers

For our entire study population, breast cancer was found at 8411 of the 332 926 diagnostic mammography examinations with findings interpreted as abnormal, which is a cancer diagnosis rate of 25.3 per 1000 examinations. Cancer diagnosis rates varied considerably according to indication for examination, ranging from a low of 8.4 per 1000 for short-interval fol-
low-up examinations to a high of 49.0 per 1000 for palpable lump cases (Table 4).

Patient-reported data on the presence or absence of a palpable lump were available for 7653 (91.0%) examinations that led to a diagnosis of breast cancer. Among these, a palpable lump was reported for 3181 (41.6%) examinations, and all but 197 of these were performed for evaluation of that symptom.

Some breast lesions are found to be palpable only in retrospect, after diagnostic mammography is performed (ie, once the presence of a lesion is verified and its three-dimensional location is precisely determined). During the study period (1996–2001), the performance of imaging guidance for tissue diagnosis was limited primarily to those lesions that were nonpalpable even in retrospect. Data on the use of imaging guidance for tissue diagnosis were unavailable for 4773 (56.7%) examinations that led to a breast cancer diagnosis. This very high percentage of missing data precludes reliable determination of the frequency with which breast cancer may be palpable in retrospect, after having been identified at diagnostic mammography.

In our overall study population, among the diagnostic mammography examinations with findings interpreted as abnormal, there were 1473 (17.5%) cases of DCIS and 6938 (82.5%) cases of invasive carcinoma. The highest percentages of DCIS were found for abnormal screening work-up and short-interval follow-up cases (26.9% and 30.7%, respectively); the lowest percentage of DCIS (5.5%) was found for cases of breast problem with a palpable lump reported (Table 4).

Data on tumor size were available for 5998 (86.5%) of the invasive cancers in this study. The mean and median sizes for these cancers were 20.2 mm and 15 mm, respectively. When stratified by indication for examination, as shown in Table 4, invasive cancer size was smallest (therefore prognosis was most favorable) for abnormal screening work-up and short-interval follow-up cases (mean, 14.3 mm; median, 11 mm) and short-interval follow-up cases (mean, 14.4 mm; median, 11 mm). Invasive cancer size was largest for palpable lump...
evaluation cases (mean, 25.6 mm; median, 21 mm).

Another widely used outcome measure indicating favorable prognosis is the frequency of minimal cancer, which is defined as either DCIS or invasive carcinoma 10 mm or smaller. For the entire study population, there were 3140 minimal cancers, representing 42.0% of the study population if DCIS and invasive cancers only of known size are considered. The highest percentages of minimal cancer were found for abnormal screening work-up and short-interval follow-up cases (62.0% and 64.7%, respectively). The lowest percentage of minimal cancer (17.5%) was found for palpable lump cases (Table 4).

Conversely, a measure of poor prognosis is the frequency of invasive carcinoma larger than 20 mm in size. For the entire study population, there were 2040 such cases, representing 34.0% of invasive cancers of known size. The lowest percentages of these large cancers were found for cases of abnormal screening work-up and short-interval follow-up (14.8% and 15.9%, respectively), whereas the highest percentage (50.8%) was found for cases of palpable lump (Table 4).

Data on axillary lymph node status were available for 6324 (91.2%) of the invasive cancers. For the entire study population, the percentage of these cancers that were node-negative (favorable prognosis) was 73.6%. The highest percentages were found for abnormal screening work-up and short-interval follow-up cases (84.2% and 86.7%, respectively), whereas the lowest percentage (65.6%) was found for palpable lump cases (Table 4).

Data on cancer stage were available for 7381 (87.8%) of the cancers. For the entire study population, the percentage of these cancers that were stage 0 or stage I (favorable prognosis) was 62.4%. The highest percentages were found for abnormal screening work-up and short-interval follow-up cases (81.7% and 82.0%, respectively), whereas the lowest percentage (40.0%) was found for palpable lump cases (Table 4).
Performance Benchmarks

The data presented in Tables 3 and 4 represent arithmetic mean values of clinical outcomes for all diagnostic mammography examinations in our study. However, because it is unlikely that outcomes for a given radiologist will closely approximate these average values, we also present ranges of performance, displayed in graphical format as smoothed plots of frequency distributions overlaid with vertical lines indicating the 10th, 25th, 50th (median), 75th, and 90th percentile values for those participating radiologists who contributed sufficient numbers of cases to provide useful data. The breadth of these ranges, shown in Figures 1–9, indicates the wide variability in individual performance among radiologists. For example, in Figure 5, E (cancer diagnosis rate, all diagnostic examinations), only 10% of eligible radiologists had a cancer detection rate lower than or equal to 13.0 cancers per 1000 examinations, whereas 90% of radiologists had a rate lower than or equal to 38.0 cancers per 1000 examinations.

DISCUSSION

The geographic diversity of the patient population served by the six BCSC mammography registries that contributed data to this study is evidenced by the fact that major demographic features (rural-urban mix, ethnicity, education level, socioeconomic status) of people in the studied catchment areas are very similar to those features found for the entire U.S. population. This, combined with the large number of examinations studied, suggests that the outcomes we report for diagnostic mammography are reasonably representative of what occurs throughout the United States. The BCSC does not collect sufficient data to reliably characterize the experience or skill of its participating radiologists. However, the patient population-based nature of BCSC data, as well as the large number of radiologists who contribute cases, makes it very likely that our population of participating radiologists is representative of the population of U.S. radiologists.
ologists is as representative as is our patient population. Therefore, we believe that realistic performance benchmarks for the practice of diagnostic mammography may be derived from our data.

In general, the outcomes we observe for diagnostic mammography are considerably different from published performance benchmarks for screening mammography (5–8) as were reported recently from the University of California at San Francisco, or UCSF (9,10). The cancer diagnosis rate is substantially greater at diagnostic mammography, and the cancers identified at diagnostic mammography are larger, more frequently node-positive, and are found at a more advanced stage than are those detected at screening mammography. These similarities between BCSC and UCSF data are due partially to inclusion of some UCSF cases in the BCSC data set. However, cancers reported from the UCSF represent only 606 (7.2%) of the 8411 BCSC cancers. Furthermore, these same general observations are valid for both the UCSF and non-UCSF cases in our study.

The overall BCSC data also confirm the previously reported UCSF observation that diagnostic mammography outcomes vary substantially by indication for examination. All three PPVs are lower for examinations performed as work-up of screening-detected abnormalities and short-interval follow-up than for those performed to evaluate a breast problem and especially those performed to evaluate palpable lumps. Similar observations apply concerning the prognostic factors of cancers identified at diagnostic mammography. Cancers identified among examinations performed as work-up of screening-detected abnormalities and short-interval follow-up are smaller, are more frequently node-negative, and are earlier in stage than are those identified among examinations performed to evaluate a breast problem and especially among those examinations performed to evaluate palpable lumps. These observations have been reported previously (9,31) and are to be expected because the populations of patients undergoing diagnostic mammography for work-up of ab-

Figure 5. Smoothed plots of frequency distributions of cancer diagnosis rate for 274,946 diagnostic mammography examinations (among radiologists with 100 or more examinations), as a function of indication for examination. An overlaid solid line indicates the 50th percentile (median), paired dashed lines indicate the 25th and 75th percentiles, and paired dotted lines indicate the 10th and 90th percentiles. A, Work-up of abnormal results detected at screening examination; B, short-interval follow-up; C, evaluation of breast problem, no lump or lump unknown; D, evaluation of breast problem, palpable lump; E, all diagnostic examinations. Corresponding arithmetic mean values for all 332,926 examinations are listed in Table 4.
normal results detected at screening examinations and for short-interval follow-up involve asymptomatic women similar to the general population of healthy women undergoing routine screening mammography (women among whom advanced cancer outcomes are less likely). The subset of patients undergoing diagnostic mammography for work-up of screening-detected abnormalities differs from the general screening population only in that mammographic abnormalities are present in all cases, thereby accounting for increased abnormal interpretation (BI-RADS category 4 and 5) and cancer diagnosis rates. Our results also reinforce previously published observations that cancer is identified very infrequently (in less than 1% of cases) among patients undergoing diagnostic mammography for short-interval follow-up (32–34).

Traditionally, performance benchmarks are derived by panels of expert practitioners from critical analysis of scientific data published in the peer-reviewed literature. This approach has been used in the development of screening mammography benchmarks. The screening benchmarks currently most widely used in the United States are stated to represent “desirable goals” achieved by “highly skilled experts” in mammography (6).

The authors of this article collectively have the appropriate expertise in breast imaging practice, epidemiology, and biostatistics to evaluate the existing scientific data on clinical outcomes for diagnostic mammography, but we find a paucity of previously published scientific data on the subject. The BCSC data reported here involve by far the most extensive published experience with diagnostic mammography and are likely to be representative of results in general practice throughout the United States rather than results achieved by highly skilled specialists. We have chosen to use only these BCSC data in deriving performance benchmarks.

To achieve the goal of presenting representative and reliable performance benchmarks in a format that is easy to understand by practicing radiologists, we have chosen to present our data not only as arithmetic means but also in the form of smoothed plots of frequency distributions of invasive cancer size for 4733 invasive cancers of known size that were identified at diagnostic mammography (among radiologists finding five or more invasive cancers of known size), as a function of indication for examination. An overlaid solid line indicates the 50th percentile (median), paired dashed lines indicate the 25th and 75th percentiles, and paired dotted lines indicate the 10th and 90th percentiles. A, Work-up of abnormal results detected at screening examination; B, short-interval follow-up; C, evaluation of breast problem, no lump or lump unknown; D, evaluation of breast problem, palpable lump; E, all diagnostic examinations. Corresponding arithmetic mean values for all 5998 invasive cancers of known size are listed in Table 4.
of frequency distribution graphs overlaid with selected calculated percentiles. Note that we have chosen to depart from the previous practice of reporting performance benchmarks as "desirable goals" based on outcomes achieved by "highly skilled experts." It is unclear at what level specialists really perform in the context of BCSC data, although the little scientific evidence already published on the subject suggests that their performance would be at the high end of the numeric scale for all performance parameters except for mean invasive cancer size, for which this would be at the low end of the numeric scale (16,18,19). Rather, the data we report are meant to indicate the range of current clinical outcomes in general practice, and percentile calculations serve as indicators of average and not-so-average performance. These data should not be used to define either standards of care or prescriptive regulatory thresholds for the clinical practice of diagnostic mammography; these issues are beyond the scope of this article. Instead, these data should be used by practicing radiologists to place into perspective the clinical outcomes observed from their own facility-wide and individual audits, for the purpose of continuing quality improvement.

How to Use Benchmark Data

How then should a mammography facility or individual radiologist use the benchmark data presented in this article? First, it will be important to collect data on most if not all of the outcomes reported in this article. One will gain very little insight into either mammography facility or individual radiologist performance if auditing is limited to the cancer-versus-no-cancer tracking of biopsy-recommended cases that is mandated in the United States by Food and Drug Administration regulation (35,36). This approach provides only PPV₂ data, which are essentially meaningless unless analyzed in combination with data on cancer detection rate, size, nodal status, and stage. Furthermore, data (mammography outcomes) collection procedures should be either fairly complete or realistically
judged to be representative in order to reduce the extent to which case selection bias confounds observed results.

Next, it will be necessary to perform a mammography audit that segregates diagnostic from screening examinations to analyze diagnostic outcomes separately. The methods used in this article parallel the BI-RADS auditing approaches developed by the American College of Radiology (22,23), so these should be followed as closely as possible. If feasible, audit data should be analyzed collectively and also separately by indication for diagnostic examination. Next, selected demographic factors of the diagnostic mammography patient population (age, family history of breast cancer, personal history of breast cancer, mammography performed previously) should be compared with those factors reported in Table 2 to determine whether and to what degree patient-related differences might confound the comparison of one’s data with those of the BCSC. For example, if one interprets mammograms from a patient population at very high or very low risk for breast cancer, the interpretations, management recommendations, and clinical outcomes will be different than those reported for the BCSC (35).

Finally, appropriate outcomes should be compared with the benchmarks reported for the BCSC, by using both arithmetic mean data from Tables 3 and 4 and the graphical data shown in Figures 1–9. For each clinical outcome, then, one will be able to judge the level of performance in terms of being above or below mean and also in terms of an estimated percentile. In so doing, it is important to recognize that larger amounts of data will be collected at the mammography facility level, which will provide more statistical precision (and therefore be less subject to random statistical variation) than data collected at the level of the individual radiologist. For relatively low-volume facilities, and especially for individual radiologists who interpret relatively few diagnostic mammograms, it may be necessary to analyze audit data collected from a period longer than the past year. Despite this limitation, it is
very important for radiologist-specific data to be analyzed because this is the only approach that will enable one to identify whether there are individual radiologists within a group practice who need to improve performance.

For those mammography facilities that are able to link their audit data with those in a regional tumor registry, thereby permitting reliable compilation of data on true-negative and false-negative results, calculations of sensitivity and specificity also should be obtained and those calculations should then be compared with parallel BCSC data posted on the BCSC Web site. For either the mammography facility or the individual radiologist who prefers to conduct an online self-versus-BCSC comparison of data, the BCSC is developing a Web site that has a secure user-driven module that employs computer prompts for data entry and validation, followed by interactive displays of performance data for diagnostic mammography for entered-versus-BCSC data. Finally, as the BCSC continues to collect mammography outcomes data over the subsequent years, we also plan to update the performance benchmarks posted on the BCSC Web site, perhaps once a year, so that repeat users will be able to compare their annual audit data with even more robust BCSC data obtained during similar periods of time.

**Study Limitations**

There are five principal limitations to the use of data from our study. First, insofar as clinical outcomes are expected to vary with changes in the demographic factors of a given patient population (3–5,26–29), those who anticipate such problems, particularly those from countries other than the United States, should...
make appropriate comparisons of their own demographic data with those of the BCSC before considering BCSC performance benchmarks to be representative of their practice.

The second limitation concerns the subset of patients undergoing diagnostic mammography for evaluation of a breast problem with no self-reported palpable lump or unknown lump status. This group of cases covers a wide variety of indications for diagnostic mammography ranging from indications similar to those for screening (patients with breast implants or breast pain) to evaluations actually ordered for a palpable lump in cases in which the patient did not self-report the presence of a lump. In the BCSC series, these cases are grouped together because no query for these specific indications was prospectively made. It is likely that the diversity of miscellaneous indications for diagnostic mammography (breast problem, no lump/lump status unknown) will vary somewhat, perhaps even widely, among different mammography facilities. Therefore, one should be cautious in comparing results from this specific subset of diagnostic examinations with results from the BCSC.

The third limitation concerns the concurrent interpretation of mammograms and US images as part of an integrated diagnostic breast imaging evaluation. Because some BCSC registries do not collect US-specific interpretation data, we cannot determine the extent to which US may have affected diagnostic mammography assessments or management recommendations. However, some mammography facilities and some radiologists probably did report integrated mammography-US assessments whereas others did not. Note that the November 2003 publication of a new edition of BI-RADS guidelines (23), in which the use of integrated mammography-US assessments is actively recommended for the first time, may confound comparison of clinical outcomes data collected in the future with the 1996–2001 data that we report in this article.

The fourth limitation concerns our calculation of benchmark percentiles based on outcomes only from those radiologists who contributed at least a designated minimum number of cases for each outcome. Although this approach reduces the number of radiologists who contribute no useful or informative data, it necessarily excludes outcomes from examinations interpreted by low-volume radiologists, ranging from exclusion of 15% of radiologists for abnormal interpretation rate benchmarks to 21% of radiologists for invasive cancer size benchmarks. Therefore, our reported data on performance benchmarks apply principally to those individual radiologists with moderate to high amounts of diagnostic mammography experience.

The fifth limitation is that many (perhaps most) mammography facilities and individual radiologists in the United States do not now conduct the type of comprehensive auditing required to properly utilize the performance benchmark data presented in this article. There simply may not be anyone available to set up, conduct, or analyze comprehensive audits. For practices that use auditing software programs, the program in use may not be able to generate data in a format that permits appropriate comparison with our data. In still other practices, it may be difficult to justify the added cost and effort to conduct comprehensive audits, especially in view of the limited reimbursement now received for breast imaging examinations. However, publication of our performance benchmark data may encourage more mammography facilities and radiologists to conduct comprehensive audits now that clinically relevant comparison data are available.

We have presented a very extensive set of data on diagnostic mammography outcomes and performance benchmarks, among a patient population judged to be representative of the population examined in general radiology practice in the United States, with data designed to be used by mammography facilities and individual radiologists to evaluate their own performance for diagnostic mammography as determined by periodic comprehensive audits. A parallel effort with similar methodology is underway to utilize BCSC data to provide clinically realistic performance benchmarks for screening mammography. Results of this effort will be reported separately.

References
PURPOSE: To retrospectively compare three different doses of gadobenate dimeglumine with a standard dose of gadopentetate dimeglumine for magnetic resonance (MR) imaging evaluation of breast vessels and to evaluate the accuracy of one-sided increased vascularity seen on gadobenate dimeglumine–enhanced MR images as an indicator of ipsilateral breast cancer.

MATERIALS AND METHODS: The original study had local ethics committee approval; informed consent was obtained from all enrolled patients. Ninety-five patients known to have or suspected of having breast cancer were randomly assigned to four groups to receive gadobenate dimeglumine at a dose of 0.05, 0.10, or 0.20 mmol per kilogram of body weight or gadopentetate dimeglumine at a dose of 0.10 mmol/kg. T1-weighted gradient-echo MR images were acquired before and 2 minutes after intravenous contrast material injection. Subtracted images were used to obtain maximum intensity projections (MIPs). Two readers blinded to the type and dose of contrast agent administered scored the MIPs obtained in the dose groups for vessel number, length, and conspicuity from 0, which indicated absent or low breast vascularity, to 3, which indicated high breast vascularity. The sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) of one-sided increased vascularity in association with ipsilateral malignancy for 69 histopathologically confirmed lesions (reference standard) were determined after gadobenate dimeglumine–enhanced MR imaging.

RESULTS: The mean MIP scores assigned to the gadobenate dimeglumine groups were significantly higher than those assigned to the gadopentetate dimeglumine group ($P = .044$). Histopathologic analysis revealed malignant lesions in 52 of 69 patients examined with gadobenate dimeglumine MR imaging: invasive ductal carcinoma in 45, invasive lobular carcinoma in four, and invasive mixed ductal-lobular carcinoma in three patients. Seventeen patients had benign lesions. Two cases of bilateral invasive cancer with symmetric breast vascular maps were excluded. Thus, the overall sensitivity, specificity, accuracy, PPV, and NPV of one-sided increased vascularity as a finding associated with ipsilateral malignancy were 88% (44 of 50 patients), 82% (14 of 17 patients), 87% (58 of 67 patients), 94% (44 of 47 patients), and 70% (14 of 20 patients), respectively.

CONCLUSION: Gadobenate dimeglumine is effective for MR imaging evaluation of breast vessels at doses as low as 0.05 mmol/kg. One-sided increased vascularity is an MR imaging finding frequently associated with ipsilateral invasive breast cancer.

© RSNA, 2005

The role of contrast material–enhanced magnetic resonance (MR) imaging of the breast for breast cancer diagnosis and management is increasing. Defined indications include pre-
surgical local tumor staging in dense breasts; surgically treated breasts in which a residual or recurrent tumor is suspected; evaluation of the effects of neoadjuvant chemotherapy; the search for occult breast cancer with known metastases; and screening of women who are at high genetic-familial risk of having breast cancer (1). Currently, breast MR imaging is considered to have very high (94%–99%) sensitivity for detection of invasive cancers but lower (50%–80%) sensitivity for detection of in situ cancers (2–8). Moreover, the specificity of MR imaging for detection of breast cancer is, at best, only moderate (65%–79%), even in interpretation models in which morphologic and dynamic criteria are integrated (7,9).

Since the first use of gadopentetate dimeglumine (Magnevist; Schering, Berlin, Germany) for contrast-enhanced MR imaging of the breast in 1986 (10), clinical breast MR imaging examinations have been performed exclusively with either this agent or similar nonspecific paramagnetic agents. All agents of this type are chelates of the paramagnetic ion gadolinium, and all are characterized by two-compartment pharmacokinetics that reflect their vascular-interstitial biodistribution. All of these agents have similar T1 relaxivities—ranging between 4.3 and 4.8 (mmol/L)-1·sec-1—so they yield comparable contrast enhancement on T1-weighted breast MR images when injected at equivalent dose (11).

Gadobenate dimeglumine (Multihance; Bracco Imaging, Milan, Italy) also is a gadolinium-based contrast agent characterized by two-compartment pharmacokinetics (12–14). However, unlike an injected dose of gadopentetate dimeglumine, a small fraction of the injected dose of gadobenate dimeglumine is taken up by functioning hepatocytes, so this agent has a partially hepatobiliary biodistribution (13,14). Furthermore, owing to its capacity for a weak and transient interaction with serum albumin (12), gadobenate dimeglumine has a twofold higher T1 relaxivity in vivo (9.7 [mmol/L]-1·sec-1) compared with gadopentetate dimeglumine and other conventional gadolinium-based agents. This property has been shown to be advantageous not only for MR imaging of intraaxial lesions of the central nervous system (15) but also for MR angiography (16–19).

The presence of increased blood flow in the breast tumor itself has been demonstrated previously at positron emission tomography (20), at temporally resolved contrast-enhanced MR imaging (with images acquired at intervals of ≤15 seconds) (21–24), and in a comparison between color Doppler ultrasonography and contrast-enhanced MR imaging (25). Furthermore, increased blood flow in the skin of the breast harboring cancer has been demonstrated by using laser Doppler perfusion (26). More recently, an ipsilateral association between cancer and increased breast vascularity has been demonstrated at MR imaging enhanced with conventional two-compartment gadolinium chelates (27,28).

The aims of this study were to compare three different doses of gadobenate dimeglumine with a standard dose of gadopentetate dimeglumine for the MR imaging evaluation of breast vessels and to evaluate the accuracy of one-sided increased vascularity seen on gadobenate dimeglumine–enhanced MR images as an indicator of ipsilateral breast cancer.

MATERIALS AND METHODS

Bracco Imaging allowed us free access to its archives of pathology data on the subjects included in this study. However, three authors (F.S., A.I., and A.F.) who are not employees of Bracco Imaging had complete control over the inclusion of any data and/or information submitted for publication that might have represented a conflict of interest for the author who is an employee of Bracco Imaging (M.A.K.).

Study Subjects

The present study was a retrospective evaluation of the MR imaging results and pathology records of patients who were enrolled at seven centers in Europe as part of a double-blind, multicenter, randomized phase II dose-finding parallel-group study to compare three different doses of gadobenate dimeglumine with a single dose of gadopentetate dimeglumine for contrast-enhanced MR imaging of the breast (29). Each center in the original study had local ethics committee approval for the study, and written informed consent was obtained from each enrolled patient. Institutional review board approval and written informed consent included approval and consent for the use and publication of any data from the study, with patient confidentiality maintained.

From July 1998 to April 1999, 95 women (mean age, 54.3 years ± 12.0 [standard deviation]; age range, 27–77 years) known to have or suspected of having breast cancer were enrolled, and a total of 95 breast MR imaging examinations were performed (one examination per patient). These patients randomly received either gadobenate dimeglumine at a dose of 0.05 (n = 24), 0.10 (n = 24), or 0.20 (n = 24) mmol per kilogram of body weight or gadopentetate dimeglumine at a dose of 0.10 mmol/kg (n = 23).

MR Imaging

MR imaging was performed by using a 1.5-T unit (Magnetom Vision or Magnetom Symphony; Siemens Medical Solutions, Erlangen, Germany) and a dedicated double-breast coil in all patients while they were in the prone position.

Contrast-enhanced MR images were obtained by using commercially available formulations of 0.5 mol/L gadobenate dimeglumine and 0.5 mol/L gadopentetate dimeglumine. All doses were administered intravenously in a bolus at a rate of 2 mL/sec and were followed by a standardized 20-mL saline flush administered at the same rate. The contrast medium and the saline flush were administered by using a power injector. Contrast-enhanced MR image acquisition started immediately after contrast agent injection, at the end of the saline flush.

MR images were acquired in the coronal plane before (unenhanced) and after (contrast enhanced) contrast agent injection by using a three-dimensional T1-weighted spoiled gradient-echo sequence with fat and water in-phase (echo time between 4.0 and 4.8 msec in the present study); a repetition time of 13 msec or less; one signal acquired; a rectangular field of view of 36 cm or smaller; a matrix of 128 × 256; a flip angle of 20°–30°; an in-plane resolution of 2 mm2 or less, which covered both whole breasts; with thin (≤3-mm) partitions but no intersection gap.

A series of five contrast-enhanced MR images were acquired, beginning at 0, 2, 4, 6, and 8 minutes after contrast material injection. Only the first dynamic phase images (acquired at 2 minutes after injection in this study) were considered, so that we had the best “angiographic effect” for both arteries and veins; in fact, in the subsequent acquisitions, a more pronounced distribution of contrast material in the interstitial space reduces the vascular enhancement (11,30). Unenhanced images were subtracted from the contrast-enhanced images on a pixel-by-pixel basis.

Image Analysis

All images were evaluated on the same workstation (MagicView 1000; Siemens
Medical Solutions) by using identical software (Numaris; Siemens Medical Solutions). Coronal and transverse maximum intensity projections (MIPs) were prepared from the subtracted MR images for assessment. For the coronal views, the posterior part of the volume was cut from the image to avoid superimposition of the contrast material–filled heart on the left breast. Two off-site readers who had 12 (F.S.) and 4 (A.I.) years of experience in contrast-enhanced breast MR imaging and were blinded to the contrast agent administered, the dose injected, and the results of histopathologic examination performed the image evaluations in consensus and in a randomized manner. They evaluated the coronal and transverse MIPs by using a one-on-one format, free windowing, and electronic magnification and distance measurement.

A score ranging from 0, indicating absent or very low breast vascularity, to 3, indicating high breast vascularity, was assigned to each pair (coronal and transverse) of MIP images on the basis of the number of vessels seen and the length and conspicuity of the vessels. The score was assigned according to the number of vessels per breast that were 3 cm or greater in length and 2 mm or greater in maximal transverse diameter: A score of 0 indicated absent or very low vascularity—that is, the complete absence of vessels or the presence of vessels less than 3 cm in length and less than 2 mm in maximal transverse diameter. A score of 1 indicated low vascularity—that is, only one vessel was 3 cm or greater in length and 2 mm or greater in maximal transverse diameter. A score of 2 indicated moderate vascularity—that is, two to four vessels were 3 cm or greater in length and 2 mm or greater in maximal transverse diameter. A score of 3 indicated high vascularity—that is, five or more vessels were 3 cm or greater in length and 2 mm or greater in maximal transverse diameter. The mean vascularity grade for the two breasts was calculated.

Two months after this evaluation, a second evaluation of the images obtained in the 72 patients who underwent gadobenate dimeglumine–enhanced MR imaging was performed. This evaluation also was performed in consensus and in a randomized manner at the satellite MR imaging workstation by the two readers mentioned earlier (F.S. and A.I.). They were blinded to the contrast agent dose injected and to the histopathologic examination results. Again, the number of vessels that were 3 cm or greater in length and 2 mm or greater in maximal transverse diameter was considered. When the difference in the number of these vessels between the two breasts was two or higher, the vascularity of the breast with the most vessels was considered to be increased compared with the vascularity of the other breast. In such cases, vascular asymmetry and the side of its occurrence were recorded, and the patient was considered to have one-sided increased vascularity. The presence and morphologic features of enhancing nodules were not considered during these evaluations.

**Histopathologic Reference-Standard Findings**

Of the 72 patients examined at gadobenate dimeglumine–enhanced MR imaging, 69 underwent histopathologic examination of their breast lesion—as the reference standard—after surgical excision (n = 63) or core-needle biopsy (n = 6). Excision or biopsy and subsequent histopathologic examination were performed at each center by local pathologists according to the 1981 World Health Organization breast cancer classification system (31). The remaining three patients opted not to be a part of this protocol and did not undergo surgery or biopsy during the original study. These three patients were therefore excluded from our analyses. Surgery and biopsy were permitted 48 hours to 1 month after MR imaging. When a malignant lesion was found at histopathologic examination of a surgical or core-needle biopsy specimen from the breast with one-sided increased vascularity, according to the findings at blinded evaluation of the MIP maps, vascular asymmetry was considered to be a true-positive associated finding.

The lesion size at histopathologic examination, when reported, was classified in terms of the maximum lesion diameter into one of the following categories: 5 mm or less, greater than 5 mm but less than or equal to 10 mm, greater than 10 mm but less than or equal to 20 mm, greater than 20 mm but less than or equal to 50 mm, and greater than 50 mm. When more than one malignant focus per breast was found, the one that was the largest in diameter in each breast was considered for the current analysis.

**Statistical Analyses**

Results were evaluated statistically by using nonparametric Kruskal-Wallis and two-tailed Mann-Whitney U tests. Kruskal-Wallis testing was used for global variability analysis of the angiographic scores assigned to the MIPs obtained in patients in the four dose groups (0.05, 0.10, and 0.20 mmol/kg gadobenate dimeglumine and 0.10 mmol/kg gadopentetate dimeglumine), and two-tailed Mann-Whitney U testing was used for post hoc intergroup comparisons of angiographic scores (SPSS for Windows, release 6.0). Kruskal-Wallis testing was also used for global variability analysis of the age and size of the malignant lesions in the four dose groups. The middle value in the lesion size classification system was considered to represent the size of all lesions with a diameter of greater than 5 mm but less than or equal to 50 mm, whereas 2.5 mm was considered to represent the size of lesions with diameters of less than or equal to 5 mm and 75 mm was considered to represent the size of lesions with diameters greater than 50 mm. Thus, the lesion sizes considered were 2.5, 7.5, 15.0, 35.0, and 75.0 mm. P < .05 was considered to indicate a significant difference.

The sensitivity, specificity, positive predictive value, and negative predictive value of one-sided increased breast vascularity as an indicator of ipsilateral invasive breast cancer were calculated. Corresponding 95% confidence intervals were calculated by using the Fisher exact method (32).
Vascular Maps and Contrast Agents

The mean scores assigned to the vascular maps obtained with gadobenate dimeglumine-enhanced MR imaging at 0.05, 0.10, and 0.20 mmol/kg doses were 1.90 ± 1.07, 1.94 ± 0.95, and 2.00 ± 0.75, respectively. The corresponding mean score assigned to maps obtained with 0.10 mmol/kg gadopentetate dimeglumine–enhanced MR imaging was 1.24 ± 0.84 (Fig 1). Global variability analysis performed with the Kruskal-Wallis test revealed significant differences (P = .018) in vascular map scores. Post hoc analysis performed by using the Mann-Whitney U test revealed significantly higher scores for each of the gadobenate dimeglumine dose groups compared with the score for the gadopentetate dimeglumine dose group (P = .044, P = .009, and P = .002, respectively, for 0.05, 0.10, and 0.20 mmol/kg gadobenate dimeglumine groups vs 0.10 mmol/kg gadopentetate dimeglumine group).

Vascular Maps and Histopathologic Lesion Types

Two of the 69 patients examined with gadobenate dimeglumine–enhanced MR imaging and for whom results of breast lesion histopathologic analysis were available had bilateral cancers with symmetric breast vascular maps. These two cases were excluded from further analysis. The lesions in the remaining 67 patients comprised 50 invasive cancers (43 invasive ductal carcinomas, four invasive lobular carcinomas, and three invasive mixed ductal-lobular carcinomas) and 17 benign lesions (four papillomas or cases of papillomatosis, three cases of hyperplasia, three cases of adenosin, three fibroadenomas, and four other benign lesions).

Asymmetric breast vascularity due to the presence of one-sided increased vascularity was observed in 47 of the 67 patients examined with gadobenate dimeglumine–enhanced MR imaging. In 44 (94%) of these 47 patients, the increased vascularity was associated with histopathologically proved ipsilateral breast cancer. These were considered to be true-positive cases. The sizes of only 36 lesions in these 44 true-positive cases were reported at histopathologic analysis: One lesion was 5 mm or smaller, seven were larger than 5 mm but smaller than or equal to 10 mm, 12 were larger than 10 mm but smaller than or equal to 20 mm, 14 were larger than 20 mm but smaller than or equal to 50 mm, and two were larger than 50 mm. The sizes of the lesions in the remaining eight patients with true-positive lesions were not reported by the investigating pathologists.

The three patients with one-sided increased vascularity who were shown not to have cancer at histopathologic examination of the suspicious lesion in the ipsilateral breast were considered to be false-positive cases. Their lesions comprised a 7-mm area of papillomatosis in the 0.10 mmol/kg gadobenate dimeglumine group and an 8-mm papilloma and a 15-mm area of hyperplasia in the 0.20 mmol/kg gadobenate dimeglumine group.

Considering one-sided increased vascularity as a finding associated with ipsilateral breast cancer led to the recording of 1.24 mm ± 3.97 for the 0.10 mmol/kg gadobenate dimeglumine group, 0.75, respectively, for 0.05, 0.10, and 0.20 mmol/kg gadobenate dimeglumine groups vs 0.10 mmol/kg gadopentetate dimeglumine group.

Note.—Data are given for 67 of the 72 original study patients who were known to have or suspected of having breast cancer, with the exclusion of three patients who had indeterminate histopathologic reference-standard results and two patients who had bilateral invasive breast cancers. No patients with pure in situ cancer were included in this series.

* Data are percentages. Numbers used to calculate given percentages are in parentheses.
of six false-negative cases among the 67 patients examined with gadobenate dimeglumine–enhanced MR imaging, with the two patients with bilateral cancers excluded. All six of these false-negative lesions were invasive ductal carcinomas and were identified in each of the three dose groups: One, three, and two lesions were identified in the 0.05, 0.10, and 0.20 mmol/kg gadobenate dimeglumine groups, respectively, and one of these lesions had a large intraductal component. Only two of these six false-negative lesions were 10 mm or smaller. The remaining false-negative lesions comprised two foci larger than 10 mm but smaller than or equal to 20 mm and two foci larger than 20 mm but smaller than or equal to 50 mm. Of 36 invasive cancers for which the size was reported at histopathologic examination, eight were minimal (≤10 mm in diameter) and 28 were not minimal (>10 mm in diameter). Ipsilateral increased vascularity was detected for six (75%) of eight minimal cancers and 24 (86%) of 28 nonminimal cancers. Consequently, two (25%) of the eight minimal cancers and four (14%) of the 28 nonminimal cancers were false-negative cases.

A cross tabulation of the results of the analysis of one-sided increased breast vascularity as a predictor of ipsilateral invasive breast cancer versus the histopathologic results is presented in Table 1. The sensitivity, specificity, positive predictive value, and negative predictive value of one-sided increased vascularity as an indicator of ipsilateral invasive breast cancer, with corresponding 95% confidence intervals, are reported in Table 2.

Examples of true-positive one-sided increased vascularity are shown in Figure 2 in association with a large (48-mm) invasive ductal carcinoma studied by using 0.10 mmol/kg gadobenate dimeglumine; in Figure 3 in association with a 14-mm mixed (lobular-ductal) invasive carcinoma studied by using 0.05 mmol/kg gadobenate dimeglumine; and in Figure 4 in association with bifocal 8- and 15-mm invasive ductal carcinomas (arrows) in the lower quadrants.

An example of false-positive one-sided increased vascularity in association with papillomatosis studied by using 0.10 mmol/kg gadobenate dimeglumine is shown in Figure 5.
MR angiography of the breast is intrinsically integrated into the standard breast examination when a dynamic three-dimensional contrast-enhanced imaging technique is used. Typical MIP images obtained by postprocessing the subtracted images reveal not only the presence of enhancing lesions but also the angiographic vascular map of vessels within the breast. This detection permits breast arteries and veins, including the internal mammary vessels, to be evaluated for conspicuity and symmetry on a routine basis. Because the acquisition time of each sequence in the dynamic MR imaging study of the breast is usually between 60 and 120 seconds, the first or second postinjection phase is generally the most suitable for angiographic evaluation. Subsequent acquisitions are more dependent on whole-body interstitial diffusion followed by renal excretion of the contrast agent during later phases (11,30).

Our experience revealed that MR imaging with gadobenate dimeglumine administered at doses as low as 0.05 mmol/kg enables high-quality vascular maps of the breast to be obtained and that the angiographic effect at this and higher doses is significantly greater than the effect following the administration of a standard 0.10 mmol/kg dose of gadopentetate dimeglumine. These findings are in agreement with those in previous MR angiography studies of other vascular territories examined with gadobenate dimeglumine enhancement (16–19). Our findings also suggest that vascular map asymmetry may be a finding that is frequently associated with ipsilateral invasive breast cancer: Sensitivity and specificity values of 88% and 82%, respectively, suggest that vascular asymmetry could be considered a corollary MR imaging sign of invasive breast cancer.

The small number of false-negative cases (n = 6) in this study precluded an accurate comparison between the dimensions of cancers associated with ipsilateral vascular prevalence and the dimensions of cancers not associated with this feature. However, it is noteworthy that at histopathologic examination, at least eight of the 44 true-positive lesions had a diameter of 10 mm or smaller, while four of the six false-negative lesions were larger than 10 mm and only two were smaller than or equal to 10 mm. These findings suggest that the dimension of the cancer probably is not the key factor in the ipsilateral prevalence of increased breast vascularity. Despite the small number of false-negative lesions, the roughly even distribution of these foci among the three dose groups is not indicative of a dose-related trend.

The presence of ipsilateral breast vascular prevalence in association with cancer may be due to reduced flow resistance in the tumor vessels, the tumor’s higher metabolism, angiogenic stimulation of the whole breast, or a combination of these factors. In our opinion, the first two possibilities may have a role in determining whole-breast increased vascularity when the cancer is relatively large in comparison with the dimension of the breast. Conversely, angiogenic stimulation of the whole breast by the tumor is more likely when the cancer is small. The role of neoangiogenic peptides in the prognosis of breast cancer remains an area of active research (33,34).

Previously, ipsilateral vascular prevalence in association with cancer was studied exclusively with gadopentetate dimeglumine. Mahfouz et al (27) administered 0.10 mmol/kg gadopentetate dimeglumine in 106 randomly selected patients—85 had unilateral malignant lesions, and 21 had unilateral benign lesions—and obtained sensitivity, specificity, accuracy, positive predictive, and negative predictive values of 77% (65 of 85 patients), 57% (12 of 21 patients), 73% (77 of 106 patients), 86% (65 of 74 patients), and 38% (12 of 32 patients), respectively, based on asymmetric vascular maps. More recently, Carriero et al (28) administered 0.20 mmol/kg gadopentetate dimeglumine in 101 patients—78 with unilateral malignant lesions and 23 with unilateral benign lesions—and obtained sensitivity, specificity, accuracy, positive predictive, and negative predictive values of 72% (56 of 78 patients), 100% (23 of 23 patients), 78% (79 of 101 patients), 100% (56 of 56 patients), and 51% (23 of 45 patients), respectively.

Although the prevalences of breast cancer in our study and in the Mahfouz et al (27) and Carriero et al (28) studies were similar (approximately 75% vs 80% and 77%, respectively), greater overall accuracy was achieved in our study (87% vs 73% and 78%, respectively) mainly owing to our higher sensitivity and negative predictive values: 88% and 70%, respectively, in our study versus 77% and 38%, respectively, in the Mahfouz et al study and 72% and 51%, respectively, in the Carriero et al study. It seems reasonable to assume that this improved accuracy is related to the greater angiographic effect of gadobenate dimeglumine (compared with the angiographic effect of gadopentetate dimeglumine), which in turn is related to the greater T1 relaxivity of gadobenate dimeglumine. The improved performance of gadobenate dimeglumine–enhanced MR imaging compared with the performance of gadopentetate dimeglumine–enhanced MR imaging in the detection and characterization of breast lesions was observed previously in the original phase II clinical trial (29).

Several limitations of our study should be taken into account. First, the evaluation of breast vascular maps was performed without masking the enhancing lesions. Although this may have introduced bias in terms of the assessment of side-based prevalence of vascularity when the vascular asymmetry was around the cutoff point, it should be noted that the evaluation procedure was similar to that performed daily in routine clinical practice. Second, our series did not include any patients with pure in situ carcinomas, which are known to have reduced angiogenesis compared with invasive carcinomas (35). However, five invasive carcinomas with large intraductal components were included, and four (80%) of them were considered to be true-positive for asymmetric breast vascularity. Nevertheless, the predictive value of asymmetric breast vascularity for in situ carcinomas remains to be investigated.

A further limitation, which was due to the retrospective nature of this study, was the absence of follow-up information on one-sided decreased vascularity to confirm the true-negative cases of breast vascular asymmetry. Similarly, no information on the remaining portion of the breast after surgery or core-needle biopsy was available to exclude the possible true-positive cases of one-sided increased vascularity in those patients who were deemed to have false-positive lesions on the basis of the benignity of the suspicious lesion after either surgery or biopsy. For these reasons, our observation of a frequent association between one-sided increased breast vascularity and ipsilateral breast cancer cannot be immediately applied to the general female population.

It should also be noted that the percentages of false-negative breast vascular maps in the presence of invasive cancers of minimal (≤10 mm in diameter) or nonminimal (>10 mm in diameter) size suggest a 25% probability of false-negative (symmetric) breast vascular maps when the invasive cancer is minimal and a 14% probability of false-negative (symmetric) breast vascular maps when the
Invasive cancer is not minimal; these probabilities should be taken into account in clinical practice. At present, we do not know if the associated finding of one-sided increased vascularity increases the sensitivity and/or the specificity of contrast-enhanced breast MR imaging, in which accuracy is based on standard morphologic and dynamic criteria, in a large population of women. This would be an extremely interesting finding, particularly if the rate of positive cases observed by using conventional criteria was lower than that observed in the current study.

Finally, it would be interesting to determine whether the association between one-sided increased vascularity and ipsilateral breast cancer is also apparent with the use of gadopentetate dimeglumine and other conventional MR imaging contrast agents or whether it is more apparent with the use of gadobenate dimeglumine owing to this agent’s higher in vivo relaxivity and preferential properties for MR angiography (16–19). The small number of cases studied with single-dose gadopentetate dimeglumine (n = 23) in the current study precluded a meaningful comparative evaluation with this agent. On the other hand, results of previous studies (27,28) have suggested that gadopentetate dimeglumine also may be appropriate for depicting ipsilateral higher vascularity.

In conclusion, our results suggest that gadobenate dimeglumine, as compared with gadopentetate dimeglumine, may have preferential properties for MR imaging evaluation of breast vascularity and that one-sided increased breast vascularity is frequently associated with ipsilateral invasive breast cancer. However, further work is clearly warranted to assess the additional value of vascular asymmetry as a sign of breast malignancy, as compared with the value of established dynamic and morphologic criteria. Additional work should be focused not only on patients with invasive cancer but also on patients with in situ cancers and in a larger population of women with benign lesions or no lesions.

References

Right Ventricular Dysfunction and Pulmonary Obstruction Index at Helical CT: Prediction of Clinical Outcome during 3-month Follow-up in Patients with Acute Pulmonary Embolism

**Purpose:** To retrospectively quantify right ventricular dysfunction (RVD) and the pulmonary artery obstruction index at helical computed tomography (CT) on the basis of various criteria proposed in the literature and to assess the predictive value of these CT parameters for mortality within 3 months after the initial diagnosis of pulmonary embolism (PE).

**Materials and Methods:** Institutional review board approval was obtained, and informed consent was not required for retrospective study. In 120 consecutive patients (55 men, 65 women; mean age ± standard deviation, 59 years ± 18) with proved PE, two readers assessed the extent of RVD by quantifying the ratio of the right ventricle to left ventricle short-axis diameters (RV/LV) and the pulmonary artery to ascending aorta diameters (PA/AO). Right ventricular short-axis diameters were measured with the right ventricular free wall being parallel to a line drawn from the point of maximal closure of the pulmonary valve to the midpoint of the pulmonary valve annulus. Pulmonary artery diameter was defined as the anteroposterior diameter at the true bifurcation level of the main pulmonary arteries. Ascending aortic diameter was defined as the largest transverse diameter of the ascending aorta at the level of the right pulmonary artery origin. The extent of pulmonary vascular obstruction was assessed using a virtual pulmonary artefact index that is the ratio of the diameter of the pulmonary artery to the diameter of the ascending aorta. The obstruction index was computed as [(diameter of pulmonary artery − diameter of ascending aorta)/diameter of ascending aorta] × 100.

**Results:** CT signs of RVD (RV/LV ratio, >1.0) were seen in 69 patients (57.5%). During follow-up, seven patients died of PE. Both the RV/LV ratio and the obstruction index were shown to be significant risk factors for mortality within 3 months (P < .04 and .01, respectively). No such relationship was found for the ratio of the pulmonary artery to ascending aorta diameters (P = .66) or for the shape of the interventricular septum (P = .20). The positive predictive value for PE-related mortality with an RV/LV ratio greater than 1.0 was 10.1% (95% confidence interval [CI]: 2.9%, 17.4%). The negative predictive value for an uneventful outcome with an RV/LV ratio of 1.0 or less was 100% (95% CI: 94.3%, 100%). There was a 11.2-fold increased risk of dying of PE for patients with an obstruction index of 40% or higher (95% CI: 1.3, 93.6).

**Conclusion:** Markers of RVD and pulmonary vascular obstruction, assessed with helical CT at baseline, help predict mortality during follow-up.

**Correspondence:** m.v.huisman@lumc.nl

**Authors:** Menno V. Huisman, MD; Albert de Roos, MD; Hein Putter, PhD; Ieneke J. C. Hartmann, MD; Peter M. T. Pattynama, MD; Rutger W. van der Meer, MD; Annette A. van den Berg-Marcou, MD; Marco J. L. van Strijen, MD; Menno V. Huisman, MD

**Published in:** Radiology 2005;235:798-803

**Keywords:** Pulmonary embolism; Right ventricular dysfunction; Pulmonary artery obstruction index; Helical CT; Mortality.
which may progress to right ventricular failure and circulatory collapse (4). It has been reported that a substantial proportion (40%) of normotensive patients with acute PE presents with echocardiographic signs of RVD. These patients with latent hemodynamic impairment have a 10% rate of PE-related shock and a 5% rate of in-hospital mortality (5). In contrast, normotensive patients without RVD have a more benign short-term prognosis (0% rate of PE-related shock or in-hospital mortality) (5). Thus, the presence of RVD is a marker for adverse clinical outcome in patients with acute PE. Echocardiography is recommended as a first-line technique to diagnose PE, assessment of RVD at CT would facilitate patient diagnosis.

In small studies with helical CT, the ratio of the right ventricle to left ventricle short-axis diameters (RV/LV) has been proposed as an accurate sign for the presence of RVD (6–9). In addition, other criteria have been proposed, including deviation of the interventricular septum and the ratio of the pulmonary artery to ascending aorta diameters (10). Also, the extent of PE (i.e., the arterial thrombus load in the pulmonary arteries) has been proposed, including deviation of the interventricular septum (11) and the pulmonary artery obstruction index (12). Therefore, results of CT in 120 patients (mean age, 59 years ± 18 [± standard deviation]; range, 18–89 years) were available for our retrospective analysis. There were 55 men (mean age, 58 years ± 16; range, 18–86 years) and 65 women (mean age, 60 years ± 20; range, 20–89 years). There was no significant difference in age between the male and female patients (P = .635).

CT Signs of RVD

The hard-copy CT images in patients with positive diagnosis of PE were selected and read by two observers (P.M.T.P. and R.W.v.d.M., with 9 years and less than 1 year of experience in examining thoracic CT scans, respectively) in a consensus reading who were unaware of the clinical signs and symptoms and the patient’s condition at the time of initial presentation. They were unaware of the clinical outcome, as well. Although P.M.T.P. was one of the investigators in the ANTELOPE study, he did not interpret the CT scans for that study. R.W.v.d.M. was a resident in radiology. The observers were aware that PE had been previously diagnosed.

The scans were evaluated by measuring the minor axes of the right and left ventricles of the heart in the transverse plane at their widest points between the inner surface of the free wall and the surface of the interventricular septum (Fig 1). These maximum dimensions may be found at
Degree of Vascular Obstruction

We quantified the vascular obstruction index (ie, the percentage of vascular obstruction of the pulmonary arterial tree caused by PE) by using the scoring system of Qanadli et al (11). In brief, this index is defined as the number of segmental artery branches that are blocked and corrected by a factor of one for partial blockage or a factor of two for completely obstructive PE. With this scoring system, the highest possible score is 40 (thrombus completely obstructing the pulmonary trunk), which corresponds to a 100% obstruction index.

Three-month Follow-up

All patients with acute PE were administered intravenous unfractionated heparin for at least 5 days; the aim of this treatment was a prolongation of activated partial thromboplastin time by a factor of 1.5–2.5. Vitamin K antagonists were given for a period of at least 3 months with the aim of an international normalized ratio of 2.0–3.0. During the follow-up period, all patients received routine clinical care from their physicians. All patient deaths during the 3 months of follow-up had been prospectively registered and evaluated by an independent adjudication committee that had full access to all available clinical and diagnostic patient data. The adjudication committee determined and recorded whether a patient death should definitely or most probably be attributed to PE or whether it should be attributed to a cause unrelated to PE.

Statistical Analysis

Statistical analysis was performed with commercially available software (SPSS version 10.0 for Windows; SPSS, Chicago, Ill). The association between death and the RV/LV ratio, the pulmonary artery-to-ascending aorta ratio, and the obstruction index are expressed as mean ± standard deviation, and the differences in the shape of the interventricular septum are expressed as frequencies. Differences in the shape of the interventricular septum, the pulmonary artery-to-ascending aorta ratio, and the obstruction index among the patients, who were divided into groups according to their RV/LV ratio, were tested with one-way analysis of variance. The least significant difference test was used as correction for multiple post hoc testing.

RESULTS

Clinical Follow-up

Three-month follow-up was completed in all patients. During the follow-up period, 18 (15%) patients died. According to the independent adjudication committee, death was related to PE in seven patients (definitely related to PE in six patients, most probably related to PE in one patient), none of whom had a history of cardiac failure or chronic obstructive pulmonary disease. Death in the remaining 11 patients was assumed to be directly related to malignancy (n = 6), myocardial infarction (n = 2), respiratory failure (n = 1), diverticulitis (n = 1), or neurologic damage after cardiopulmonary resuscitation (n = 1). Table 1 lists the duration of survival after presentation for these 18 patients.

RV/LV Ratio

The mean RV/LV ratio was 1.17 ± 0.37 (Table 2). There was a statistically significant relationship between the RV/LV ratio and PE-related mortality, with a regression coefficient of 1.55 (P = .04, Table 3). The mean RV/LV ratio was 1.17 ± 0.37 (Table 2). When the patients were divided into three groups (patients who died of PE, patients who survived, and patients who died of a cause unrelated to PE), results of the overall test for equality for the three groups together showed a statistically significant difference in the mean RV/LV ratios (P = .018). Results of post hoc testing showed that in patients who died of PE, the mean RV/LV ratio was significantly higher than that in patients who survived (1.54 ± 0.18 vs...
In our study population, an RV/LV ratio of 1.0 or less was predictive value for an uneventful outcome with an RV/LV ratio greater than 1.5 (three deaths from PE, 11.2% of patients had an RV/LV ratio greater than 1.0 but less than or equal to 1.5 (four deaths from PE, 8%), and 18 (15%) patients had an RV/LV ratio greater than 1.5 (three deaths from PE, 17%). Results of analysis of variance testing with least significant difference as a post hoc test showed a significant increase in age (P < .001) for patients in both groups with an elevated RV/LV ratio compared with patients with a RV/LV ratio of 1.0 or less. There was no significant difference in age distribution between the two groups with elevated RV/LV ratios (P = .325). There was no significant difference in sex distribution (X² test, P = .559) between the three groups. The positive predictive value for PE-related mortality with an RV/LV ratio greater than 1.0 was 10.1% (95% confidence interval: 2.9%, 17.4%), and the negative predictive value for an uneventful outcome with an RV/LV ratio of 1.0 or less was 100% (95% confidence interval: 94.3%, 100%). In our study population, an RV/LV ratio of 1.0 or less excluded mortality due to PE.

CT-derived Vascular Obstruction Index

The mean value of the vascular obstruction index was 31.8% ± 22.9 (Table 2); in one case, a patient had undergone complete pneumectomy of one lung, and so this case was censored for this measurement. Results of Cox regression analysis showed a statistically significant relationship between the vascular obstruction index and PE-related mortality (P = .01, Table 3).

When the patients were divided into three groups (patients who died of PE, patients who survived, and patients who died of a cause unrelated to PE), results of the overall test for equality for the three groups together showed a statistically significant difference in the mean obstruction indexes (P = .001). Results of post hoc testing showed that patients who died of PE during follow-up had, on average, a significantly higher vascular obstruction index than did patients who survived (60.4% ± 28.4 vs 29.3% ± 20.8, P < .001) or patients who died of a cause unrelated to PE (36.8% ± 26.3, P = .027). There was no significant difference in the average obstruction index between patients who died of causes unrelated to PE and patients who survived (P = .280) (Fig 3).

When the previously recommended cutoff value of 40% for the pulmonary vascular obstruction index was used (11), patients with an index of 40% or greater had an 11.2-fold increased risk of dying of PE (despite anticoagulant treatment) relative to patients with an index of less than 40% (hazard ratio, 11.2: 95% confidence interval: 1.3, 93.6). Six of 37 patients with an obstruction index of 40% or greater died of PE versus one of 83 patients with an obstruction index of less than 40%. The positive predictive value for PE-related mortality with an obstruction index of 40% or greater was 16.2%

---

**TABLE 2**

CT Measurements in Patients with Proved PE

<table>
<thead>
<tr>
<th>Parameter</th>
<th>All Patients (n = 120)</th>
<th>RV/LV Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 120</td>
<td>&gt;1.5 (n = 18)</td>
</tr>
<tr>
<td>Maximum diameter (mm)†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right ventricle</td>
<td>46.9 ± 8.3</td>
<td>57.8 ± 7.4</td>
</tr>
<tr>
<td>Left ventricle</td>
<td>42.1 ± 7.8</td>
<td>31.4 ± 4.5</td>
</tr>
<tr>
<td>RV/LV ratio†</td>
<td>1.17 ± .37</td>
<td>1.86 ± .31</td>
</tr>
<tr>
<td>Interventricular septum‡</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>71 (59)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Displacement grade 2</td>
<td>24 (20)</td>
<td>4 (22)</td>
</tr>
<tr>
<td>Displacement grade 3</td>
<td>25 (21)</td>
<td>14 (78)</td>
</tr>
<tr>
<td>PA/Ao ratio‡</td>
<td>0.92 ± .14</td>
<td>0.92 ± .13</td>
</tr>
<tr>
<td>Obstruction index (%)‡</td>
<td>31.8 ± 22.9</td>
<td>62.8 ± 20.8</td>
</tr>
<tr>
<td>Obstruction index categorical‡</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;40%</td>
<td>83 (69)</td>
<td>2 (11)</td>
</tr>
<tr>
<td>≥40%</td>
<td>37 (31)</td>
<td>16 (89)</td>
</tr>
</tbody>
</table>

* Refers to overall test for equality of means or proportions.
† Data are mean ± standard deviation.
‡ Data are numbers of patients, and numbers in parentheses are percentages.
§ Grade 2 = flattened, grade 3 = deviated toward the left ventricle.
¶ PA/Ao ratio = ratio of inner lumen pulmonary artery diameter to aorta diameter.
∥ In one patient, the obstruction index could not be determined because of complete pneumectomy of one lung.

**TABLE 3**

Univariate Cox Regression Analysis of Death due to PE during 3-month Follow-up in Patients Treated with Anticoagulant Therapy

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Regression Coefficient</th>
<th>Standard Error</th>
<th>P Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>RV/LV ratio</td>
<td>1.55</td>
<td>0.74</td>
<td>.035</td>
</tr>
<tr>
<td>Obstruction index†</td>
<td>0.041</td>
<td>0.016</td>
<td>.009</td>
</tr>
<tr>
<td>PA/Ao ratio‡</td>
<td>1.36</td>
<td>3.08</td>
<td>.66</td>
</tr>
</tbody>
</table>

* PA/Ao ratio = ratio of inner lumen pulmonary artery diameter to aorta diameter.
The negative predictive value for an uneventful outcome with an obstruction index of less than 40% was 98.8% (95% confidence interval: 96.4%, 100%).

There was a significant ($P < .001$) but weak correlation between the RV/LV ratio and the vascular obstruction index, OBS, determined with the regression equation: $RV/LV = 0.0099$ \times OBS + 0.85 (Pearson $r^2 = 0.38$). The positive predictive value for PE-related mortality among patients with an RV/LV ratio greater than 1.0 and an obstruction index of 40% or greater was 18.8% (95% confidence interval: 5.0%, 32.6%), and the negative predictive value for an uneventful outcome among patients with an RV/LV ratio of 1.0 or less or an obstruction index of less than 40% was 98.9% (95% confidence interval: 96.7%, 100%).

**Figure 2.** Graph of mean values of the RV/LV ratio relative to clinical outcome. Bars represent means and lines represent 95% confidence intervals. The mean RV/LV ratio was significantly higher in patients who died of PE than in the other patients (analysis of variance, $P = .018$). Brackets show significant differences between groups.

**Figure 3.** Graph of mean values of obstruction index relative to clinical outcome. Bars represent means and lines represent 95% confidence intervals. The mean obstruction index was significantly higher in patients who died of PE than in the other patients (analysis of variance, $P = .001$). Brackets show significant differences between groups.

**DISCUSSION**

Results of this study showed that the presence and severity of RVD and the extent of obstruction of the pulmonary arterial tree as assessed at helical CT help predict mortality within 3 months of clinical presentation in initially hemodynamically stable patients with PE. It is important to note that both the absence of RVD as assessed by means of an RV/LV ratio of 1.0 or less and an obstruction index of less than 40% were predictive for an uneventful outcome. In contrast, the deviation of the interventricular septum and the ratio between the pulmonary artery and ascending aorta diameters seemed to have no prognostic relevance in acute PE. It is of interest that clinically relevant parameters of RVD can be obtained even from CT examinations that are not gated to the electrocardiogram.

Various diagnostic techniques or laboratory tests have been proposed to stratify patients with PE at clinical presentation into groups with higher or lower risk for fatal PE, with the ultimate aim of identifying those patients who might benefit from more aggressive fibrinolytic therapy. Results of previous studies have shown that echocardiography might be a useful method to predict RVD and clinical outcome (4,5,14). As an alternative, blood tests, including measurement of plasma brain natriuretic peptides (15,16) or cardiac troponins T and I (17), have been proposed as prognostic indicators for benign versus complicated courses. A decrease in cardiac troponins T and I (17), have been proposed as prognostic indicators for benign versus complicated courses. Ten Wolde and colleagues (15) showed that the adjusted odds ratio of a brain natriuretic peptide level greater than 21.7 pmol/L for death related to PE was 14.1 (95% confidence interval: 1.5, 131.1), while in another study, a cutoff level of a brain natriuretic peptide level of less than 50 pg/L enabled identification of 95% of patients with a benign clinical course.

802 • Radiology • June 2005 van der Meer et al
Konstantinides et al. (17) found that 35%–40% of patients with PE had elevated cardiac troponin levels that were associated with overall mortality and a complicated course.

Some limitations of our study should be considered. First, imaging was performed with single–detector row helical CT scanners. At the time of investigation, multi–detector row CT scanners were not yet available in the participating hospitals. To our knowledge, there currently are no reports of studies available in the literature in which RV/LV ratios assessed with single–detector row technology have been compared with those assessed with multi–detector row CT scanners. We admit that this newer CT technology could have been helpful for better imaging of the subsegmental arteries. It remains to be studied whether the results will be different at multi–detector row CT, since subsegmental arteries have only a minimal influence on the obstruction index. Second, the number of deaths related to PE that were observed during the follow-up period was somewhat limited and, as a consequence, the confidence limits around the point estimates are wide.

Regardless of the small numbers, our findings are in agreement with those of Wu et al. (12), who suggested that the quantification of a clot at CT pulmonary angiography is an important predictor of patient death in the setting of PE, whereas authors of a previous study with an even smaller number of PE-related deaths (18) did not find a significant relationship between RVD or the extent of vascular obstruction of the pulmonary artery circulation and clinical outcome. Our study should be followed with a larger study to confirm our findings. Future studies on the prognostic role of RVD and obstruction index assessment at helical CT for PE should also include the determination of brain natriuretic peptide and troponin measurements in individual patients to determine their complementary role in stratifying the heterogeneous group of patients with PE. On the basis of these tests, a patient group suitable for a shorter stay in the hospital or even for full treatment outside of the hospital (16) could be selected. Furthermore, more aggressive therapy, including thrombolytic therapy (19), may be warranted in those patients with RVD or a high degree of vascular obstruction, although additional evidence from properly designed clinical trials is still awaited.

We conclude that both the RV/LV ratio and the pulmonary vascular obstruction index as assessed at helical CT are potentially useful tools to predict mortality in patients with initially hemodynamically stable PE at clinical presentation.

Acknowledgments: Results of this study are part of the results of the ANTELLOPE study, a Dutch prospective multicenter trial on PE. The following investigators were part of the ANTELLOPE study group: B. J. Sanson, J. G. Lijmer, M. ten Wolde, M. H. Prins, H. R. Büller (Academic Medical Center, Amsterdam); W. de Monyé, M. V. Huisman, P. M. T. Pattynama (Leiden University Medical Center, Leiden); M. J. L. van Strijen, G. J. Kieft (Leyenburg Hospital, The Hague); M. R. MacGillavry, F. Turkstra, D. P. M. Brandjes (Sloter-vaart Hospital, Amsterdam); P. J. Hagen, R. J. Roeleveld, O. S. Hoekstra, P. E. Postmus (Vrije Universiteit Medical Center, Amsterdam); I. J. C. Hartmann, P. F. G. M. van Waes, J. D. Banga (Universiteit Medical Center, Utrecht).

References

2. van Beek EJ, Kuijjer PM, Büller HR, Brandjes DP, Bossuyt PM, ten Cate JW. The clinical course of patients with suspected pulmonary embolism. Arch Intern Med 1997; 157:2593–2598.
Purpose: To investigate whether two-phase contrast material–enhanced computed tomographic (CT) findings serve as predictors of changes in left ventricular (LV) function and wall thickness (WT) after acute myocardial infarction (MI) and successful angioplasty.

Materials and Methods: Ethics committee approval and informed consent were obtained. In 58 patients (51 men and seven women; mean age, 62 years ± 12 [standard deviation]) who had experienced an acute MI and undergone successful angioplasty, two-phase (acquisitions at 45 seconds and 7 minutes) contrast-enhanced CT was performed in the acute (mean interval between treatment and CT, 37 hours ± 4) and intermediate (mean interval, 28 days ± 4) periods and for long-term (mean interval, 12 months ± 4) follow-up. CT images were reviewed for an early perfusion defect (ED) at 45 seconds and for late enhancement (LE) and a residual perfusion defect (RD) at 7 minutes. Myocardial enhancement patterns and WT were assessed, and LV ejection fraction (LVEF) and percentage decrease in WT were calculated. The patient group was subdivided into three groups according to enhancement pattern: Group 1 included patients with LE but no ED or RD; group 2, patients with ED and LE but no RD; and group 3, patients with ED, LE, and RD. Fisher exact testing was used to measure categorical response. Paired and unpaired t tests were used for comparison between two groups (points); Tukey-Kramer multiple comparison and repeated-measures analysis of variance were used for comparisons between the three groups. P < .05 was considered to indicate a significant difference.

Results: In group 3 (n = 36), WT in infarcted area was significantly reduced at the intermediate and long-term CT examinations (P < .001). At the intermediate and long-term examinations, percentage decrease in WT was greater in group 2 (n = 10) than in group 1 (n = 12) (P < .05 for intermediate and P < .001 for long-term examination) and was greatest in group 3 (P < .001 for both examinations). LVEF was poorest in group 3 and best in group 1.

Conclusion: Two-phase contrast-enhanced CT proved useful in predicting LV functional recovery and WT in patients who had experienced acute MI and undergone successful angioplasty.

© RSNA, 2005

Previous researchers (1,2) have proposed the use of a reconstruction algorithm with an electrocardiographically (ECG)-gated or non–ECG-gated technique at single-section helical computed tomography (CT) (3) for cardiac imaging. Multi–detector row CT has been developed in recent years as the temporal resolution of helical CT has improved, and CT has thus become a useful tool for coronary artery imaging (4–6).
Because the prognosis of patients who have experienced an acute myocardial infarction (MI) depends on their myocardial reperfusion status, the assessment of microvascular flow after reperfusion therapy is of great importance.

Conventionally, myocardial perfusion is evaluated mainly by using nuclear imaging techniques (7,8). Contrast material–enhanced magnetic resonance (MR) imaging (9–12) and myocardial contrast echocardiography (13–15) are also common procedures that can be employed to assess myocardial perfusion. Electron-beam CT also provides reliable information on myocardial perfusion (16). In a case report, contrast-enhanced helical CT revealed an early myocardial perfusion defect and late enhancement in acute MI, and these findings agree with those observed at nuclear imaging and contrast-enhanced MR imaging (17), suggesting a possible use for contrast-enhanced helical CT in myocardial perfusion assessment.

The goal of this study was to investigate whether two-phase contrast-enhanced CT findings serve as predictors of changes in left ventricular function and wall thickness in patients who have experienced acute MI and have undergone successful reperfusion therapy.

MATERIALS AND METHODS

Initial Study Population

Our study initially included 65 consecutive patients enrolled between October 1999 and March 2001: 59 patients who had experienced acute MI and had undergone successful reperfusion therapy and six patients who had experienced acute MI, had undergone successful reperfusion therapy, and had previously undergone part of the study protocol that was later used with the other 59 patients. These six patients agreed to complete the protocol used with the other 59 patients. The study protocol was approved by the ethics committee of Ehime Prefectural Imabari Hospital, and all patients gave their informed consent.

Inclusion Criteria

Inclusion criteria consisted of the following:

1. The presence of typical symptoms of acute MI associated with ECG changes and a serum concentration of creatine kinase (CK) of more than twice the upper limit of normal with more than 5% of the isoenzyme CK-MB in the serum.

2. A first acute MI that was related to a single coronary artery.

3. Successful coronary angioplasty of the totally or subtotally occluded infarct-related artery (Thrombolysis in Myocardial Infarction [TIMI] grade 0 or 1) within 12 hours after the onset of chest pain. In brief, TIMI grade 0 perfusion indicates that there is no antegrade flow beyond the point of occlusion, grade 1 indicates minimal incomplete perfusion of a contrast medium around the clot, grade 2 (partial perfusion) indicates complete but delayed perfusion of the distal coronary bed with a contrast medium, and grade 3 (complete perfusion) indicates antegrade flow to the entire distal coronary bed at a normal rate.

4. Residual stenosis of less than 50% after angioplasty.

Exclusion Criteria

Exclusion criteria consisted of the following:

1. Renal failure (serum creatinine level > 1.5 mg/dL [132.6 μmol/L] and/or blood urea nitrogen level > 21 mg/dL [7.5 μmol/L]).

2. Restenosis of 50% or greater at coronary angiography during follow-up.

3. Inadequate CT imaging and/or cine angiographic results.

4. Lipid degeneration or calcification in the myocardium on non–ECG-gated 10-mm-thick plane scans obtained before the contrast-enhanced CT examination was performed.

Seven patients were excluded from the analysis because of the following reasons: Their cineangiograms were inadequate for evaluation of TIMI flow grade (n = 2), they had incomplete coronary recanalization (residual stenosis ≥ 50% owing to distal embolization) (n = 1), restenosis was detected at coronary angiography during follow-up (n = 2), or their CT images were inadequate because of atrial fibrillation (n = 1) or body movement (n = 1).

Final Study Population

Therefore, this study is based on data from 58 evaluable patients (mean age, 62 years ± 12 [standard deviation]; age range, 39–84 years): 51 men (mean age, 62 years ± 11; age range, 39–83 years) and seven women (mean age, 66 years ± 14; age range, 46–84 years). According to results of unpaired t testing, there were no significant differences between the male and the female patients in terms of age. The left anterior descending artery was involved in 31 patients, the right coronary artery was involved in 23 patients, and the circumflex artery was involved in four patients. Fifty-seven patients (98%) underwent stent placement, and one patient (2%) underwent balloon angioplasty. Regarding the final TIMI grade, 56 patients (97%) had TIMI grade 3 reflow, and two patients (3%) had TIMI grade 2 reflow.

Study Protocol

Coronary angiography was performed in all 58 patients who underwent angioplasty. In the acute phase study, conventional left ventriculography was performed immediately after coronary angioplasty, which was performed by one of three cardiologists (T.I., H.M., and J.H., with 25, 23, and 30 years of clinical practice, respectively), to assess end-diastolic volume (EDV), end-systolic volume (ESV), and ejection fraction (EF). Two-phase contrast-enhanced CT was performed within 48 hours (mean interval, 37 hours ± 4) after direct angioplasty. In the intermediate phase study, both coronary angiography and left ventriculography were performed a mean of 27 days ± 3 after direct angioplasty, and two-phase contrast-enhanced CT was performed a mean of 28 days ± 4 after direct angioplasty. In the long-term study, two-phase contrast-enhanced CT was performed 12 months ± 4 after direct angioplasty. The biplanar angiographic system used for coronary angiography and conventional left ventriculography was an Integris V3000 (Philips Medical Systems, Best, the Netherlands). A Cardio 500 analysis system (Kontron Electronik, Munich, Germany) was used for performing quantitative coronary angiography and assessing cardiac function. During the infusion of the contrast medium (Optiray [320 milligrams of iodine per milliliter]; Yamanouchi, Tokyo, Japan) at a rate of 8 mL/sec to a total of 35 mL, biplanar images were obtained at a filming rate of 30 images per second.

Two-Phase Contrast-enhanced CT

The helical CT scanner used was a single-detector row Proceed SA (GE-Yokogawa Medical Systems, Tokyo, Japan) with a gantry rotation speed of 0.8 second. Patients were asked to lie supine on the CT table and inhale oxygen at a rate of 3 L/min while the scanning parameters were prepared.

The scanning protocol was based on that used in a previous study involving electron-beam CT (16) and was as fol-
Nonionic iodinated contrast medium (iopamidol, Iopamiron [300 milligrams of iodine per milliliter]; Nihon Schering, Osaka, Japan) was intravenously administered at a rate of 1.5 mL/sec for the first acquisition for the early image. The early image was obtained 45 seconds after the start of contrast medium administration; the same contrast medium was then infused at a rate of 0.1 mL/sec for the second acquisition for the late image. The late image was obtained 7 minutes after the start of contrast medium administration; a total of 150 mL of the contrast medium was used (Fig 1).

The ECG trace was recorded during scanning so that we could determine triggers and identify the diastolic image data set. Patients were asked to hold their breath during whole-heart scanning. Scanning parameters were as follows: collimation, 3 mm; table feed, 3 mm per rotation; and number of rotations, 40 (at a rate of 12 cm every 32 seconds). The tube current and voltage were 200 mA and 140 kV, respectively.

A partial 180° reconstruction algorithm was used, and the temporal resolution of one transaxial image was 0.4 second. The field of view was 180 mm, and the matrix was 512 × 512. Transaxial images were generated by using overlap reconstruction at a pitch of 0.1, and 391 3-mm-thick transaxial images, at intervals of 0.3 mm (0.08 second), including various cardiac phases, were obtained (1). The weighted CT dose index, as defined in reference 18, was 9.4 mGy for one acquisition.

The diastolic image data set, which was extracted with reference to the R wave of the recorded ECG trace, was used to assess myocardial enhancement pattern and wall thickness. The 3-mm-thick cardiac short-axis images were reconstructed by using the double-oblique method (Fig 2). All measurements were performed by using a ZIO-M900 workstation (ZIO Software, Tokyo, Japan).

Assessment of Myocardial Enhancement with Two-Phase Contrast-enhanced CT

All patient-identifying information on the CT images was obscured; the images were then randomized. Two cardiologists (T.I. and Y.K., with 25 and 10 years of clinical practice, respectively) and two radiologists (H.H. and T.M., with 20 and 25 years of clinical practice, respectively), who were blinded to patient treatment information, visually judged the myocardial enhancement pattern together. Disagreements were solved by consensus.

We defined a myocardial perfusion defect (a dark zone) on the early-phase images (those obtained 45 seconds after contrast material administration) as an early perfusion defect, the presence of smaller dark regions in the subendocardium surrounded by partially enhanced myocardium on the late-phase images (those obtained 7 minutes after contrast material administration) as a residual perfusion defect, and the presence of an enhanced zone on the late-phase images as late enhancement.

Regions of interest were created for one-third of the area of each finding (early perfusion defect, residual defect, and late enhancement) and were placed over the center of each area. A region of interest of the same size, which was at least greater than 50 mm², was also placed over a remote noninfarcted area on the opposite side of the infarct-related area so that we could measure the mean attenuation. This was performed by the same four investigators (T.I., Y.K., H.H., and T.M.) working in consensus.
Enhancement Patterns

Several enhancement patterns theoretically exist; however, all enhancement could actually be classified into one of the following three patterns: group 1, in which there is an absence of early perfusion defect in the early phase and a presence of late enhancement without residual perfusion defect in the late phase; group 2, in which there is a presence of early perfusion defect in the early phase and a presence of late enhancement without residual perfusion defect in the late phase; and group 3, in which there is a presence of early perfusion defect in the early phase and a presence of both late enhancement and residual perfusion defect in the late phase (Fig 3).

Wall Thickness

Wall thickness was measured by the same four investigators (T.I., Y.K., H.H., and T.M.) in consensus by using an early-phase end-diastolic short-axis image. The measured points in the acute phase, intermediate phase, and long-term studies were the infarct-related and the noninfarcted areas. We calculated the percentage decrease in wall thickness, or \( WT_{pd} \), in the intermediate phase and long-term studies as follows: \( WT_{pd} = [(WT_a - WT_t)/WT_t] \times 100 \), where \( WT_a \) is the wall thickness observed during the acute phase study, \( WT_t \) is that observed during the intermediate phase study, and \( WT_l \) is that observed during the long-term study.

In the long-term study, a scar was defined according to published criteria based on autopsy data (19). If the myocardial segment was thin on the early-phase images (wall thickness, <6 mm), the patient was given a diagnosis of transmural scar formation.

Analysis of Conventional Coronary Angiographic and Left Ventriculographic Results

All patient-identifying information on the coronary angiograms was obscured; the angiograms were then randomized. Two cardiologists (H.M. and H.K., with 23 and 14 years of clinical practice, respectively), who were blinded to patient clinical outcome, classified in simultaneous consensus the antegrade contrast material flow in the infarct-related artery on the final coronary angiograms according to the criteria established by the TIMI study group. All patient-identifying information on the conventional left ventriculograms was obscured; the ventriculograms were then randomized. Two observers (J.H. and J.A., with 30 and 14 years of clinical practice, respectively) measured EDV and ESV and calculated the left ventricular EF by using the area-length method (20) and working in consensus.

Determination of Peak CK Level and CK-MB Fraction

Immediately after reperfusion therapy, blood was collected every 4 hours, and J.A. determined the maximum values of serum CK and the isoenzyme CK-MB.

Determination of Ischemic Time

Ischemic time was measured by J.A. and was defined as the interval from the onset of the symptoms to the time at which the first balloon inflation was performed.

Statistical Analysis

All data are expressed as means ± standard deviations. The Fisher exact test was used for comparing the frequencies of sex and culprit artery among the groups (group 1, group 2, and group 3). The Tukey-Kramer multiple comparison test was used for comparing age among the groups. An unpaired t test was used for comparing mean CT values in regions of interest of noninfarcted area, and the Tukey-Kramer multiple comparison test was used for comparing CT values between three areas (the region of interest of the noninfarcted area, the region of interest of the late enhancement area, and the region of interest of the residual perfusion defect area). The Tukey-Kramer multiple comparison test was also used for comparing ages, CK levels, CK-MB fractions, and ischemic time between the groups.

A paired t test was used for two-point comparisons (between the acute and the intermediate phase study) of EDV, ESV, and EF in each group. The Tukey-Kramer multiple comparison test was also used for group comparisons of EDV, ESV, and EF at two points (the acute and intermediate phase studies).

Regarding the wall thickness of the noninfarcted area, repeated analysis of variance measurement of groups, time points (the acute phase, intermediate phase, and long-term studies), and their interaction terms was performed. As a covariance structure among the time
Characteristics of 58 Patients according to Group

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients</td>
<td>12 (21)</td>
<td>10 (17)</td>
<td>36 (62)</td>
</tr>
<tr>
<td>Sex*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>10 (83)</td>
<td>7 (70)</td>
<td>34 (94)</td>
</tr>
<tr>
<td>Female</td>
<td>2 (17)</td>
<td>3 (30)</td>
<td>2 (6)</td>
</tr>
<tr>
<td>Age (y)</td>
<td>66 ± 11</td>
<td>68 ± 11</td>
<td>59 ± 11</td>
</tr>
<tr>
<td>Culprit artery*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right coronary</td>
<td>5 (42)</td>
<td>5 (50)</td>
<td>13 (36)</td>
</tr>
<tr>
<td>Left anterior descending</td>
<td>5 (42)</td>
<td>3 (30)</td>
<td>23 (64)</td>
</tr>
<tr>
<td>Circumflex</td>
<td>2 (17)</td>
<td>2 (20)</td>
<td>0</td>
</tr>
<tr>
<td>Clinical treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stent placement</td>
<td>12 (100)</td>
<td>10 (100)</td>
<td>35 (97)</td>
</tr>
<tr>
<td>Balloon angioplasty</td>
<td>0</td>
<td>0</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Final TIMI grade</td>
<td>2</td>
<td>0</td>
<td>2 (6)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>12 (100)</td>
<td>10 (100)</td>
</tr>
</tbody>
</table>

Note.—Unless otherwise stated, data are numbers of patients, with percentages in parentheses. Percentages may not add up to 100% owing to rounding.

* No significant differences (P > .05) between groups (Fisher exact test).
† Data are mean values ± standard deviations. There were no significant differences (P > .05) in age between the groups (Tukey-Kramer multiple comparison test).

Figure 4. Short-axis contrast-enhanced helical CT images show the typical enhancement patterns in the three groups. Images obtained in 56-year-old man in group 1 with acute MI (arrows) in the territory of the left anterior descending coronary artery show no early perfusion defect and late enhancement without a residual defect. Images obtained in 62-year-old woman in group 2 with acute MI (arrows) in the territory of the right coronary artery show an early perfusion defect and late enhancement without a residual defect. Images obtained in 60-year-old man in group 3 with acute MI (arrows) in the territory of the circumflex coronary artery show an early defect and late enhancement with a residual defect.

EDV, ESV, and EF in Acute and Intermediate Phase Studies: Intra- and Intergroup Comparisons

In group 1, between the time of the acute phase and the time of the intermediate phase study, the EDV did not change significantly (value at acute phase study, 129 mL ± 24; value at intermediate phase study, 121 mL ± 29 [P = .069]), the ESV decreased (value at acute phase study, 21 mL ± 14; value at intermediate phase study, 18 mL ± 13 [P < .05]). Between the intermediate phase and the long-term study, the ESV also decreased (value at intermediate phase study, 18 mL ± 13; value at long-term study, 15 mL ± 9 [P < .01]).

RESULTS

Myocardial Enhancement Pattern at Two-Phase Contrast-enhanced CT

There were 12 patients in group 1, 10 patients in group 2, and 36 patients in group 3. Patient characteristics were not significantly different between the three groups (Table).

We found that the attenuation of early perfusion defects (30.2 HU ± 11) was significantly lower than that of the noninfarcted areas (102.1 HU ± 9.0, P < .001) on the early-phase (45-second) image. Areas with late enhancement (112.9 HU ± 18.5) had higher attenuation compared with areas with residual perfusion defect (59.3 HU ± 11, P < .001) and noninfarcted areas (88.6 HU ± 13.7, P < .001) vs areas with late enhancement, P < .001 vs areas with residual perfusion defect) on the late-phase (7-minute) images (Fig 4).

Peak CK Level, Peak CK-MB Fraction, and Ischemic Time

Peak CK (2555 U/L ± 1055) and CK-MB (194 U/L ± 89) values in group 2 were higher than those in group 1, which had a peak CK value of 904 U/L ± 270 (P < .01) and a peak CK-MB value of 87 U/L ± 43 (P < .05). Peak CK and CK-MB values, respectively, were highest in group 3 at 4111 U/L ± 1301 (P < .001) and 285 U/L ± 109 (P < .05). The ischemic time (the time elapsed from the first onset of symptoms to the first balloon inflation) was 282 minutes ± 93 in group 3 and was significantly longer than that in group 1 (170 minutes ± 59, P < .001). The ischemic time in group 2 (234 minutes ± 44) did not significantly differ from that in groups 1 and 3.

points, compound symmetry was assumed and robust variance was used.

Regarding the wall thickness of the infarcted area, heterogeneity among individuals in the degree of damage due to infarction was expected, so we used the wall thickness at the acute phase study as an adjusting factor; repeated-measures analysis of variance with groups, time points (the intermediate phase and long-term studies), and their interaction terms was then performed. As a covariance structure among the time points, compound symmetry was assumed and robust variance was used.

The Tukey-Kramer multiple comparison test was used for group comparisons of percentage decrease in wall thickness at two points (the intermediate phase and long-term studies). A paired t test was used for two-point comparisons (between the intermediate phase and the long-term study) of percentage decrease in wall thickness in each group. P < .05 was considered to indicate a statistically significant difference. The statistical analyses were performed by using SAS analyses (SAS Institute, Cary, NC).
study, 51 mL ± 11; value at intermediate phase study, 28 mL ± 8 [P < .001], and the EF improved (value at acute phase study, 60% ± 11; value at intermediate phase study, 76% ± 8 [P < .001]). In group 2, EDV, ESV, and EF did not change significantly: EDV decreased from 122 mL ± 24 to 115 mL ± 26, ESV decreased from 46 mL ± 12 to 43 mL ± 14, and EF increased from 62% ± 7 to 63% ± 8. In group 3, EDV increased from 121 mL ± 35 to 148 mL ± 37 [P < .001], ESV increased from 44 mL ± 21 to 72 mL ± 30 [P < .001], and EF decreased from 64% ± 11 to 52% ± 13 [P < .001]. In the acute phase study, no significant differences were noted in EDV, ESV, or EF among the three groups.

In the intermediate phase study, the EDV in group 3 was the largest among the three groups (P < .05), the ESVs in groups 1 and 2 were smaller than the ESV in group 3 (P < .001 and P < .01, respectively), and the EF in group 1 was higher than that in group 3 (P < .001) and group 2 (P < .05). The EF in group 2 was higher than that in group 3 (P < .05).

**Wall Thickness and Percentage Decrease in Wall Thickness**

The wall thickness values in the non-infarcted area at three time points—the acute phase, intermediate phase, and long-term studies, respectively—were 12.0 mm ± 1.7, 12.2 mm ± 1.7, and 11.8 mm ± 1.7 in group 1; 11.0 mm ± 1.7, 11.1 mm ± 1.6, and 11.0 mm ± 1.7 in group 2; and 11.5 mm ± 2.1, 11.5 mm ± 2.0, and 11.2 mm ± 1.8 in group 3. There were no significant differences among the three groups (P = .058), or their interactions (P = .181) according to results of repeated-measures analysis of variance.

The wall thickness values in the infarcted area at three time points—the acute phase, intermediate phase, and long-term studies, respectively—were 11.7 mm ± 1.3, 11.6 mm ± 1.3, and 11.6 mm ± 1.5 in group 1; 11.1 mm ± 2.1, 9.1 mm ± 1.7, and 8.1 mm ± 1.7 in group 2; and 11.1 mm ± 1.9, 6.1 mm ± 1.7, and 4.8 mm ± 1.5 in group 3. There were no significant differences (P = .657) in wall thickness at the acute phase study among the groups. However, according to results of repeated-measures analysis of variance, there were significant differences within each group (P < .001) and among time points (P = .001), their interactions (P < .001), and adjusting factors (P = .001).

In group 3, 21 patients (58%) with transmural early perfusion defect were found to have scar formation. Figure 5 shows typical images obtained at the long-term study (in the same patients as in Fig 4) in these three groups.

As regards the infarcted area, because groups, time points (intermediate phase and long-term studies), and their interactions differed significantly, we tried to investigate the percentage decrease in wall thickness, which was based on the wall thickness at the acute phase study, at two points (the intermediate phase and long-term studies).

In terms of percentage decrease in wall thickness of the infarcted areas, there were no significant differences between group 1 and group 2 at the intermediate phase and long-term studies. The percentage decrease in wall thickness in group 3 at the long-term study was greater than that at the intermediate phase study (P < .001).

At the intermediate phase study, the percentage decrease in wall thickness of the infarcted areas in group 1 (0.4% ± 1.7) was the lowest among the three groups (group 2: 17.6% ± 9.8, P < .05; group 3: 44.4% ± 15.9, P < .001). The percentage decrease in wall thickness in group 3 was greater than that in group 2 (P < .001).

At the long-term study, the percentage decrease in wall thickness in group 3 (56.1% ± 4.2) was the highest among the three groups (group 1: 0.9% ± 3.6, P < .001; group 2: 26.5% ± 10.7, P < .001). The percentage decrease in wall thickness in group 2 was greater than that in group 1 (P < .001).

**DISCUSSION**

Our results show that, in patients with acute MI, the myocardial enhancement pattern at two-phase contrast-enhanced CT performed after reperfusion therapy could serve as a predictor of left ventricular functional recovery and wall thickness. Previous researchers have reported that mortality is reduced in patients with TIMI grade 3 flow (21–26). Gibson et al (27) reported that patients with both normal epicardial (TIMI grade 3) flow and normal tissue-level perfusion (TIMI myocardial perfusion grade 3) have an extremely low mortality risk.

Although revascularization of the epicardial coronary arteries (improvement of TIMI grade flow) is necessary for the myocardium to be salvaged, this is not enough to ensure myocardial recovery. That is, successful revascularization of epicardial coronary arteries is not equal to successful reperfusion at the microvascular level. In this study, we detected three enhancement patterns among the 56 patients with TIMI grade 3 flow; the variability in enhancement pattern indicates that there is variability in the extent of microvascular damage. The least extensive recovery of left ventricular function was observed in patients who had both an early and a residual perfusion defect, namely, those in group 3. That is, they might not have complete reperfusion at the microvascular level, while those patients who did not have an early or a residual perfusion defect (group 1) might experience a complete recovery of left ventricular function, which might be...
indicative of successful reperfusion at both the epicardial and the microvascular level.

**Importance of an Early Defect**

Contrast medium is thought to reach the microvascular bed in the early phase after intravenous administration. A study involving electron-beam CT (16) revealed that myocardial enhancement in the early phase reflected the volume of the vascular bed.

Reduced signal intensity on first-pass contrast-enhanced MR images has been shown to indicate reduced blood flow (28). Therefore, an early perfusion defect observed by using CT would also reflect a decrease in the volume of the vascular bed—that is, a decrease in the myocardial blood flow.

Using an experimental infarct-reperfusion model in dogs, Braunwald and Kloner (29) classified the condition of myocardial tissue into four layers beginning from the endocardial side. The first layer corresponded to a viable and very thin myocardium that received oxygen directly from the left ventricle; the second layer, to myocardial necrosis with extensive capillary (microcirculation) disorder; the third layer, to myocardial necrosis in which blood supply was preserved to some extent; and the fourth layer, to stunned myocardium that had escaped necrosis. The early perfusion defect in our study may correspond to the second layer (myocardial necrosis with extensive capillary disorder) of the Braunwald classification and that late enhancement might correspond to the third layer, where blood supply is preserved to some degree, indicating possible residual myocardial viability.

In this study, the percentage decrease in wall thickness in group 3 was significantly greater than that in group 2—that is, the residual perfusion defect in group 3 indicated that there was less antegrade microvascular flow beyond the point of microvascular obstruction in group 3 than in group 2. As a result of incomplete perfusion at the microvascular level, the percentage decrease in wall thickness in group 3 was greater than that in group 2 in the intermediate phase and long-term studies.

Given our results, we may conclude that a residual perfusion defect indicated a necrotic area caused by severe microvascular obstruction—the so-called no-reflow phenomenon (31) that is caused by the presence of red blood cells and necrotic debris (32) in the “wavefront” of ischemic necrosis (33,34).

In general, there is a consensus that delayed hyperenhancement at MR imaging reflects nonviable myocardium (30). However, other studies involving humans revealed that 3–5 days after a reperfused MI, some regions of enhancement recovered function 3 months later (10,35). In our study, late enhancement was also observed in both group 1 and group 2, indicating that the area of late enhancement includes viable myocardium, at least in part, at examinations performed within 48 hours of reperfusion therapy. An MR imaging–based study in rats that involved occluding the coronary artery for 30 minutes and for 2 hours revealed that the enhanced zone was time dependent (ie, it decreased in size over time), and the true infarct size corresponded to the enhancement size at 21 minutes ± 4 (36). On late images, which were acquired 7 minutes after the start of the administration of the contrast medium in that study, the true infarct size might have been overestimated; however, the contrast medium had been slowly injected at 0.1 mL/sec (to a total of 37.5 mL) after the bolus injection at 1.5 mL/sec (to a total of 112.5 mL). The total dose and injection rate of the bolus and/or the continuous injection of the contrast medium might have had an effect on the size of the enhancing lesion, although these factors were not investigated in that study.

After successful reperfusion therapy, both the wall thickness and the microvasculature change dynamically during the acute healing stage. Therefore, in the clinical setting, we assume that the timing of the CT examination after reperfusion therapy is important in assessing late enhancement with contrast-enhanced CT.

However, this concept of these enhancement patterns and their combinations at CT has not attained wide acceptance, and because CT can also help define the depth and extent of early and residual perfusion defects and late enhancement, the sizes of these parameters in comparison with the infarct size should be investigated in future studies.

**Study Limitations**

In this study, because the images were read concurrently by four observers and we could not compare interobserver and intraobserver agreement, interobserver and intraobserver reliabilities were not confirmed.

Regarding the x-ray exposures during two-phase contrast-enhanced CT, the radiation dose for one acquisition was 9.4 mGy. In our protocol, we performed three two-phase contrast CT studies, for a radiation dose of 18.8 mGy for each study and a total of 56.4 mGy, which does not include the fluoroscopy dose at angioplasty and angiography. Although x-ray exposure dose was high in this study, our results suggest that if two-phase contrast-enhanced helical CT was performed once within 48 hours after reperfusion therapy, the enhancement patterns could be evaluated and these enhancement patterns would serve to predict left ventricular function and wall thickness, while the x-ray exposure at this examination would be 18.8 mGy—exactly the same exposure incurred when nonoverlapping reconstructions are used. Additionally, the exposed range (12 cm) for a two-phase cardiac CT examination is smaller than that for a three-phase abdominal CT examination.

**Importance of Residual Defect and Late Enhancement**

After the contrast medium reaches the microvascular bed, it gradually flows into the interstitium (extracellular space), stays there, and is then washed out slowly. Therefore, myocardial enhancement in the late phase mainly reflects the characteristics of the interstitium—that is, the volume of the interstitial space (16). When myocardial cells are damaged and the cell number decreases after an acute MI, the volume of the interstitial space increases.

In the present study, when a residual perfusion defect was detected, as it was in group 3, functional recovery was not observed. However, when an early perfusion defect turned into late enhancement, as happened in group 2, deterioration of left ventricular function was minimal or less than that observed in group 3. We speculate that the area of residual perfusion defect might correspond to the second layer (myocardial necrosis with extensive capillary disorder) of the Braunwald classification and that late enhancement might correspond to the third layer, where blood supply is preserved to some degree, indicating possible residual myocardial viability.

In this study, the percentage decrease in wall thickness in group 3 was significantly greater than that in group 2—that is, the residual perfusion defect in group 3 indicated that there was less antegrade microvascular flow beyond the point of microvascular obstruction in group 3 than in group 2. As a result of incomplete perfusion at the microvascular level, the percentage decrease in wall thickness in group 3 was greater than that in group 2 in the intermediate phase and long-term studies.

Given our results, we may conclude that a residual perfusion defect indicated a necrotic area caused by severe microvascular obstruction—the so-called no-reflow phenomenon (31) that is caused by the presence of red blood cells and necrotic debris (32) in the “wavefront” of ischemic necrosis (33,34).

In general, there is a consensus that delayed hyperenhancement at MR imaging reflects nonviable myocardium (30). However, other studies involving humans revealed that 3–5 days after a reperfused MI, some regions of enhancement recovered function 3 months later (10,35). In our study, late enhancement was also observed in both group 1 and group 2, indicating that the area of late enhancement includes viable myocardium, at least in part, at examinations performed within 48 hours of reperfusion therapy. An MR imaging–based study in rats that involved occluding the coronary artery for 30 minutes and for 2 hours revealed that the enhanced zone was time dependent (ie, it decreased in size over time), and the true infarct size corresponded to the enhancement size at 21 minutes ± 4 (36). On late images, which were acquired 7 minutes after the start of the administration of the contrast medium in that study, the true infarct size might have been overestimated; however, the contrast medium had been slowly injected at 0.1 mL/sec (to a total of 37.5 mL) after the bolus injection at 1.5 mL/sec (to a total of 112.5 mL). The total dose and injection rate of the bolus and/or the continuous injection of the contrast medium might have had an effect on the size of the enhancing lesion, although these factors were not investigated in that study.

After successful reperfusion therapy, both the wall thickness and the microvasculature change dynamically during the acute healing stage. Therefore, in the clinical setting, we assume that the timing of the CT examination after reperfusion therapy is important in assessing late enhancement with contrast-enhanced CT.

However, this concept of these enhancement patterns and their combinations at CT has not attained wide acceptance, and because CT can also help define the depth and extent of early and residual perfusion defects and late enhancement, the sizes of these parameters in comparison with the infarct size should be investigated in future studies.
tion. Therefore, when one considers the useful information obtained from the myocardial studies, the radiation dose seems to be acceptable for patients with heart disease that may threaten their lives.

In conclusion, in patients with acute MI, the myocardial enhancement pattern at two-phase contrast-enhanced CT performed after reperfusion therapy can serve as a predictor of left ventricular functional recovery and wall thickness.

Acknowledgments: We are grateful to Tsuyoshi Matsunaka, MD, Kazuhiro Nishimura, MD, Katsuji Inoue, MD, Kana Sakamoto, MD, and Junko Kato, MD, for their excellent assistance in data analysis. We are very grateful to Yoshinobu Kubota (Toward, Tokyo, Japan) for technical support in data computing and to GE Yokogawa Medical Systems for analyzing and calculating the weighted CT dose index in this study.

References


31. Reimer KA, Ganote CE, Jennings RB. The “wavefront phenomenon” of myocardial ischemic cell death. II. Transmural progression of necrosis within the framework of ischemic bed size (myocardium at risk) and collateral flow. Lab Invest 1979; 40:633–644.


Anomalous Coronary Arteries in Adults: Depiction at Multi–Detector Row CT Angiography

PURPOSE: To retrospectively determine the imaging features of anomalous coronary arteries depicted at multi–detector row computed tomographic (CT) angiography in 18 patients seen at four institutions.

MATERIALS AND METHODS: Eighteen patients underwent imaging with a four- or 16-section multi–detector row CT unit by using retrospective electrocardiographic (ECG) gating after infusion of 120–150 mL of intravenous contrast material. Section thicknesses of 0.8–3.0 mm were achieved during breath holding, and images were reconstructed with a 50% overlap. Volumetric reconstructions were obtained for each patient. Each study was assessed retrospectively for the origin and course of the anomalous coronary artery by two thoracic radiologists; decisions were made in consensus. Institutional review board exemption and informed consent waiver was granted at each institution. The study was compliant with the Health Insurance Portability and Accountability Act.

RESULTS: Seventeen patients were referred because of equivocal findings at cardiac catheterization or echocardiography; in one, the anomalous coronary artery was incidental. A total of 20 anomalous vessels were found. Twelve patients with 14 variant vessels had an anomalous origin of a left coronary artery (right cusp, 13; noncoronary cusp, one). In four patients, an anomalous right coronary artery originated from the left side; one patient had a single coronary artery arising from the right cusp. In one patient, a left coronary artery-to-vein fistula was observed. In 10 patients, the anomalous vessel passed between the aorta and the main pulmonary artery or right ventricular outflow track. In each case, the origin of the anomalous coronary artery and its course in relationship to the great vessels were unequivocally demonstrated. Volumetric images were useful for showing the three-dimensional orientation of the anomalous coronary artery with respect to the great vessels and cardiac chambers.

CONCLUSION: Multi–detector row CT angiography provided accurate depiction of vessel origin and course in this review of 20 anomalous coronary arteries. The results of this study suggest that CT is a viable noninvasive modality for delineating coronary arterial anomalies, particularly if findings at coronary angiography are equivocal.

Coronary artery anomalies are potentially life-threatening anatomic variants that occur in approximately 1% of patients (1,2). Most of these anomalous vessels are not clinically important. However, it has been recognized for more than 3 decades that patients in whom the aberrant vessel passes between the aorta and the main pulmonary artery are at risk for sudden death, particularly if the vessel supplies the left coronary artery distribution (3). Coronary artery bypass grafting may be indicated for such patients (1).

Conventionally, evaluation of coronary artery anomalies is performed by using catheter-based angiography. However, the precise course of the vessel may not be adequately defined with this technique. Magnetic resonance (MR) imaging has often been used to
delineate the anomalous coronary artery in equivocal cases; however, MR imaging can be limited by low spatial resolution and artifacts and can be technically challenging (4). Recently, the development of multi–detector row computed tomography (CT) has permitted better definition of the coronary vessels with CT. Thus, the purpose of our study was to retrospectively determine the imaging features of anomalous coronary arteries depicted at multi–detector row CT angiography in 18 patients seen at four institutions.

MATERIALS AND METHODS

Patients

A retrospective evaluation was performed at four institutions to identify all patients who underwent retrospective electrocardiographically (ECG) gated cardiac multi–detector row CT angiography, in whom an anomalous coronary vessel was found. These 18 patients constitute the study group. The patients ranged in age from 18 to 74 years (mean age, 52 years); there were 14 men and four women. The patients underwent imaging at one of four institutions (Vanderbilt University, Nashville, Tenn; University of Maryland Hospital, Baltimore, Md; University Hospitals, Cleveland, Ohio; Indiana University, Indianapolis, Ind) between 2001 and 2003.

Record Review

The medical records were reviewed by the lead investigator at each institution (Vanderbilt University, J.D.; University of Maryland Hospital, C.S.W.; University Hospitals, R.C.G.; Indiana University, C.A.M.). Information was recorded with regard to medical history and whether prior cardiac catheterization or ECG had been performed. In addition, the results of any surgical procedures were noted. The decision as to whether to recommend coronary artery bypass graft surgery was documented. An institutional review board exemption and informed consent waiver was granted for this study at each institution. Our study was compliant with the Health Insurance Portability and Accountability Act.

Imaging

CT scans were obtained in each patient by using four- (n = 11) or 16- (n = 7) section multi–detector row CT scanners (MX8000 or MX8000iDT; Phillips Medical Systems, Best, the Netherlands). For all but one patient, retrospective ECG-gated images were obtained through the heart during one or two breath holds. Details of the scanning protocols for the four- and 16-section CT scanners are provided in Table 1. The average scanning time was 30 seconds, with 3–4 additional minutes for preprocedural placement and adjustment of ECG leads. Reconstructions at various phases of the cardiac cycle were performed. Among patients who underwent ECG gating, the 75% phase during diastole was found to be optimal in the analysis of anomalies of the left coronary artery. Anomalies of the right coronary artery were evaluated at the 37.5%, 50%, or 75% phase of the cardiac cycle, depending on which showed the least amount of motion. Between 120 and 150 mL of iodinated contrast material was injected through an 18–20-gauge intravenous catheter into an antecubital vein at a rate of 3–4 mL/sec. A timing bolus was performed with 20 mL of contrast material, or automated bolus timing was used, as described in Table 1. β-Blockers were not used to control heart rate.

Table 1: Scanning Protocols for Four- and 16-Section CT

<table>
<thead>
<tr>
<th>Parameter</th>
<th>16-Section CT</th>
<th>Four-Section CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tube voltage (kV)</td>
<td>140</td>
<td>120</td>
</tr>
<tr>
<td>Tube current (mA)</td>
<td>350–500</td>
<td>350–400</td>
</tr>
<tr>
<td>Section thickness (mm)</td>
<td>1.0</td>
<td>1–3</td>
</tr>
<tr>
<td>Increment (mm)</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>Imaging time (sec)</td>
<td>0.42</td>
<td>0.5</td>
</tr>
<tr>
<td>Pitch</td>
<td>0.2–0.3</td>
<td>0.375</td>
</tr>
<tr>
<td>Field of view (mm)</td>
<td>220–250</td>
<td>220–250</td>
</tr>
<tr>
<td>ECG gating</td>
<td>Retrospective</td>
<td>Retrospective</td>
</tr>
<tr>
<td>Algorithm</td>
<td>Soft tissue</td>
<td>Soft tissue</td>
</tr>
<tr>
<td>Bolus timing</td>
<td>Automatic with region of interest on descending aorta. Threshold set at 150 HU and a 10-sec delay to scan</td>
<td>Automatic with region of interest on ascending aorta. Threshold set at 100 HU and a 4-sec delay to scan</td>
</tr>
</tbody>
</table>

Image Review

All images were reviewed in a transverse projection at the operator’s console in a cine stack mode, which provided the primary evaluation of the anomalous vessel. In addition, the study data were transferred to an off-line workstation (MXView; Phillips Medical Systems), and further reconstructions were obtained with sliding thick- and thin-slab multiplanar reformation (5–15 mm), with thin-section maximum intensity projection (5 mm), and in a three-dimensional volumetric mode (volume rendering). A qualitative assessment was made by the lead investigator at each site, with regard to the utility of transverse images as compared with the various reconstructions in establishing the diagnosis. The images that showed the anomaly best were reviewed by a minimum of two thoracic radiologists who were blinded to the specific results of coronary angiography and echocardiography (C.S.W. and J.D., R.C.G., or C.A.M.). The determination of coronary artery origin and course (anterior to the main pulmonary artery, posterior to the aorta, intervascular or intraseptal course) of the anomalous vessel was arrived at by consensus. An intraseptal course was defined as passage of the vessel through septal muscle well below the plane of the pulmonic valve. An intervascular course was designated as a vessel extension near the level of or higher than the pulmonic valve. An assessment was made as to the presence of the kinking of the anomalous vessel near its origin, which was defined as angulation greater than 90° of the artery at CT.

RESULTS

Of 18 patients with 20 anomalous coronary arteries, 17 were referred owing to equivocal findings at cardiac catheterization (n = 16) or echocardiography (n = 1). Each of these patients was being evaluated for chest pain. In the patients referred after cardiac catheterization and angiography, the angiogram demonstrated the exact site of origin of the anomalous vessel, but the referring service was uncertain of the precise course. In one patient, the aberrant vessel was found incidentally during nongated multi–detector row CT angiography performed for unrelated reasons. In the case of the arteriovenous fistula, the location of the fistula could not be identified at conven-
ational angiography owing to the high flow state.

Twelve patients (14 vessels) had an anomalous origin of a left coronary artery (13 vessels from the right cusp, and one from a noncoronary cusp) (Fig 1). In two of these 12 patients, there were separate anomalies of the left anterior descending and circumflex arteries. Four patients demonstrated an anomalous origin of the right coronary artery from the left side. One patient had a single coronary artery arising from the right cusp (Fig 2). In the last patient, a left coronary artery-to-vein fistula was identified.

Ten anomalous vessels in 10 patients passed between the aorta and the main pulmonary artery (Table 2). Seven of these 10 patients had a left coronary vessel that originated from the right cusp (Figs 3, 4). In the remaining three patients, the right coronary artery originated from the left cusp and extended in an intervascular course between the aorta and the main pulmonary artery to reach the anterior atrioventricular groove (Fig 5). The left anterior descending artery was affected in three of these seven left-sided vessels, and the left main coronary artery arose from the right cusp in the other four. Among four of the seven left-sided vessels, the left coronary artery followed an intervascular course between the aorta and the main pulmonary artery. In the other three vessels, an intraseptal course of an anomalous left coronary artery was observed.

Of the 10 patients in whom the anomalous artery (10 coronary arteries) had an intervascular or intraseptal course, six underwent bypass grafting (four left and two right coronary arteries). One of the three patients with an intraseptal course had severe atherosclerotic disease and was deemed not to be a candidate for surgery. In the other, a decision was made not to perform surgery on the basis of a negative myocardial perfusion study. Surgery was recommended in one patient; however, this patient declined surgery. In the remaining patient, surgery was recommended, but the patient was

### TABLE 2
Profile of Patients with Anomalous Coronary Arteries Coursing between the Great Vessels

<table>
<thead>
<tr>
<th>Patient No./Age (y)/Sex</th>
<th>No. of Sections</th>
<th>Gated</th>
<th>Anomalous Vessel*</th>
<th>Origin*</th>
<th>Course†</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/47/M</td>
<td>16</td>
<td>Yes</td>
<td>LMC</td>
<td>Right</td>
<td>Intervascular</td>
<td>Surgery</td>
</tr>
<tr>
<td>2/53/F</td>
<td>16</td>
<td>Yes</td>
<td>LMC</td>
<td>Right</td>
<td>Intervascular</td>
<td>Surgery</td>
</tr>
<tr>
<td>3/61/M</td>
<td>4</td>
<td>Yes</td>
<td>LMC</td>
<td>Right</td>
<td>Intervascular</td>
<td>Surgery</td>
</tr>
<tr>
<td>4/59/M</td>
<td>4</td>
<td>Yes</td>
<td>LMC</td>
<td>Right</td>
<td>Intraseptal</td>
<td>None‡</td>
</tr>
<tr>
<td>5/44/M</td>
<td>4</td>
<td>Yes</td>
<td>LAD</td>
<td>Right</td>
<td>Intervascular</td>
<td>Surgery</td>
</tr>
<tr>
<td>6/46/M§</td>
<td>16</td>
<td>Yes</td>
<td>RCA</td>
<td>LAD</td>
<td>Intervascular</td>
<td>Surgery</td>
</tr>
<tr>
<td>7/64/M</td>
<td>16</td>
<td>Yes</td>
<td>RCA</td>
<td>LAD</td>
<td>Intervascular</td>
<td>Surgery</td>
</tr>
<tr>
<td>8/28/M</td>
<td>4</td>
<td>Yes</td>
<td>RCA</td>
<td>Left</td>
<td>Intervascular</td>
<td>Surgery</td>
</tr>
<tr>
<td>9/62/M</td>
<td>16</td>
<td>Yes</td>
<td>RCA</td>
<td>Left</td>
<td>Intraseptal</td>
<td>Surgery</td>
</tr>
<tr>
<td>10/43/M§</td>
<td>16</td>
<td>Yes</td>
<td>LAD</td>
<td>Right</td>
<td>Intraseptal</td>
<td>Surgery</td>
</tr>
</tbody>
</table>

* LAD = left anterior descending artery, LMC = left main coronary artery, RCA = right coronary artery.
† Intravascular course is between the aorta and pulmonary artery.
‡ This patient was not a candidate for surgery.
§ Patient also had a left circumflex artery from the right cusp that coursed posterior to the aorta.

---

**Figure 1.** Images obtained in a 45-year-old woman with recurrent chest pain and palpitations. (a) Coronary angiogram obtained with a right anterior oblique projection shows an anomalous right coronary artery (arrow) originating from the proximal left anterior descending artery and extending to the anterior atrioventricular groove. (b, c) Transverse CT scans obtained with a four-section scanner show the anomalous right coronary artery (arrow) arising from the left anterior descending artery and coursing anterior to the main pulmonary artery (P) into the anterior atrioventricular groove. A = aorta.
excluded for noncompliance with medical regimen.

In each case, multi-detector row CT angiography unequivocally demonstrated the origin of the anomalous coronary artery and its course in relation to the great vessels. CT was also effective in delineating the angulation or kinking of the vessel with respect to its point of origin (Fig 3).

Although the transverse images were sufficient to determine the arterial origin and course, reconstructed images provided useful supplemental information. Maximum intensity projection, thin-slab multiplanar reformations, and volumetric images were useful for depicting the three-dimensional spatial orientation of the anomalous vessel with respect to the great vessels and cardiac chambers.

In the patient with the arteriovenous fistula, CT showed the junction between the dilated proximal left circumflex artery and the great cardiac vein, as well as the more normal-appearing left circumflex artery distal to the arterial–venous junction. A volume-rendered image demonstrated the tortuous and ectatic vessel coursing along the posterior surface of the heart (Fig 6). The arteriovenous fistula was presumed to be congenital.
DISCUSSION

Coronary artery anomalies occur in about 1% of patients undergoing cardiac catheterization (1). Three types of ectopic anomalies have been described: (a) ectopic origin from a coronary sinus, (b) absent coronary artery, and (c) ectopic origin from a main pulmonary artery. In adults, the first two types of anomaly are most common. Another type of anomaly is a coronary arteriovenous fistula, which accounts for approximately 13% of anomalies (1,5).

Approximately 20% of coronary artery anomalies produce life-threatening symptoms, including arrhythmias, syncope, myocardial infarction, or sudden death. Indeed, congenital coronary artery anomalies are the second most common cause of sudden death due to structural heart disease in young athletes (6). In particular, an anomalous vessel that crosses between the aorta and the main pulmonary artery, either a left coronary artery originating from the right sinus or a right coronary artery emanating from the left sinus, may be associated with a poorer outcome (7,8). It is postulated that the acute angulation or slit-like ostium associated with this type of anomaly may lead to myocardial ischemia. During intense physical activity, the aortic wall stretches and dilates to support the increased cardiac output, which further compresses the slit-like ostium of the anomalous vessel against the main pulmonary artery (9). Demonstration of traversal of the coronary artery between the aorta and the main pulmonary artery may be an indication for coronary artery bypass surgery (10). It is important to differentiate an intervascular course, in which the anomalous vessel courses directly between the aorta and the main pulmonary artery, from an intraseptal course, which may appear similar but in which the anomalous vessel passes more inferiorly within the muscular septum. The latter course is associated with a comparatively benign course. CT can often help differentiate these two entities by showing that the anomalous vessel passes within the septal muscle inferior to the plane of the pulmonic valve.

At angiography, the precise course of the anomalous vessel may be difficult to delineate due to its complex three-dimensional geometry displayed in two di-

Figure 5. Images obtained in a 28-year-old man with hypertension, recurrent chest pain, and an anomalous right coronary artery. (a) Coronary angiogram obtained in a left anterior oblique projection shows the anomalous right coronary artery (arrow) originating from the left coronary artery. (b, c) Contiguous transverse thick-slab reformations (10-mm-thick section) from a four-section CT unit show the right coronary artery (arrow in b) originating from the left cusp (arrowhead) and extending between the aorta (A) and the main pulmonary artery (P) to reach the anterior atrioventricular groove. (d) Volume-rendered image demonstrates the anomalous vessel (arrow) extending from the left cusp anteriorly. A = aorta, P = main pulmonary artery.

Figure 6. Images obtained in a 69-year-old man with chest pain. Results of coronary angiography were suggestive of an arteriovenous fistula. (a) Volumetric image shows the markedly enlarged and tortuous left circumflex vessel (arrows) coursing along the posterior cardiac surface. (b) Parasagittal thin-slab reformatted image (5-mm-thick section) shows the junction of the left circumflex artery (CX) and the great cardiac vein (GCV). Note the more normal-appearing distal (dist) circumflex artery. prox = proximal
mensions fluoroscopically. In the hands of an experienced angiographer, the proper diagnosis can be suggested at cardiac catheterization. The rarity of these anomalies results in limited experience for many angiographers, and this, in turn, causes a large percentage of coronary artery anomalies to be categorized incorrectly at coronary catheterization (11). Thus, noninvasive imaging may prove valuable for confirming the diagnosis.

Noninvasive imaging of the coronary arteries presents substantial obstacles. The proximal coronary vessels measure only 4 mm in diameter, with tapering of the more distal vessels. The arteries are tortuous and are subject to both respiratory and cardiac motion by virtue of their epicardial location. The two major techniques used to noninvasively image the coronary arteries are MR angiography and CT angiography.

Substantial investigative work has been performed with regard to the use of MR imaging and MR angiography in the evaluation of the coronary arteries (12). Multiple two- and three-dimensional strategies have been attempted, but the spatial resolution of the technique has been insufficient to permit its widespread adoption in the setting of atherosclerotic coronary disease. As shown in several studies, MR angiography provides an accurate assessment of the course of anomalous coronary arteries and can be used as an adjunctive technique when results of coronary angiography are equivocal (4,13,14).

Nevertheless, MR imaging has several disadvantages in the assessment of the diminutive coronary arteries. The technique cannot be performed for patients with pacemakers or defibrillating devices, and it may be difficult to perform for claustrophobic patients. MR imaging is best performed with ECG gating, which may not be achievable in patients with certain arrhythmias. Finally, the spatial resolution of MR imaging is substantially inferior to that of the newest generation of CT scanners. Notwithstanding these limitations, MR imaging continues to be a plausible noninvasive alternative for imaging anomalous coronary arteries, as evidenced by a recent MR angiography study with a three-dimensional free-breathing sequence (15).

Early reports of the use of CT for coronary artery evaluation have emphasized electron-beam technology. Ropers et al (16) used ECG-gated electron-beam CT with 3-mm collimation and three-dimensional reconstructions to assess 60 patients, 30 of whom had coronary artery anomalies. Two observers successfully categorized each patient as to the presence or absence of coronary anomalies. Twenty-nine of the 30 anomalies were characterized accurately. Although this study demonstrated the feasibility of CT, electron-beam CT is not widely available and is limited by relatively poor z-axis resolution.

More recently, a number of developments in helical CT technology have substantially enhanced its utility in the evaluation of the coronary arteries (17). Current multi–detector row CT technology, consisting of at least four to 16 detector rows, enables rapid coverage of the coronary territory, generally in a single breath hold. The more rapid coverage permits thinner collimation, often with sections of 1 mm or less. A faster gantry rotation (500 msec or less), combined with the use of partial reconstruction algorithms, has led to more rapid imaging. Currently, a temporal resolution of 50–200 msec is achievable depending on heart rate. With overlapping reconstruction during the cardiac cycle, eight to 10 cardiac phases can be obtained. The better temporal resolution and the ability to choose the optimal cardiac phase result in a near cessation of physiologic motion.

An accurate diagnosis may be achievable, even in the absence of an optimal CT study. As demonstrated with one case in this study, proximal coronary artery images of diagnostic quality may be acquired without ECG gating. Intravenous contrast material was used in each case in this series, but it may not be required to depict coronary anomalies if there is sufficient epicardial fat to provide intrinsic tissue contrast. These issues may be clinically relevant because adequate ECG gating is not possible in all patients, and allergy or renal insufficiency may prevent the use of intravenous contrast material. To date, and to our knowledge, the use of multi–detector row CT in the detection of anomalous coronary arteries has been described only in case reports (18,19).

As demonstrated in the present series, multi–detector row CT angiography was able to delineate clearly the origin and course of the anomalous coronary artery. In no case was the proximal course of the coronary vessel ambiguous. In general, the entire study, including lead placement, was completed within 5 minutes, which is substantially faster than is typically achieved with MR imaging. The small section thickness permitted volumetric reconstructions of high quality.

Although these images were not crucial in the diagnosis of the anomaly, they were valuable for depicting the relationships among the coronary vessel, great vessels, and ventricles. Such images give the surgeons a better understanding of the complex anatomy before repair.

An important consideration is the high effective radiation dose (9 mSv) to which a patient undergoing multi–detector row CT angiography is exposed. By comparison, the estimated radiation dose with electron-beam CT and diagnostic coronary arteriography is 1–2 mSv and 3–10 mSv, respectively (20,21). Although the radiation dose is of concern, investigators have shown the potential to substantially decrease the radiation exposure with multi–detector row CT angiography by reducing the dose during the systolic portion of the cardiac cycle (22).

Our study is limited both by its retrospective nature and a relatively small number of patients. Another drawback is the lack of a reference standard technique to prove that our consensus CT findings were correct. In all but two of our patients, an equivocal coronary angiogram, the purported reference standard, was the stated indication for CT. Although direct visualization at surgery is confirmatory, the anomalous coronary course is often obscured by epicardial fat (Cardarelli M, oral communication, 2004). Thus, surgery cannot be considered a consistently reliable standard reference technique. Because of the indeterminate coronary angiogram and lack of an external reference standard, a direct comparison between multi–detector row CT angiography and coronary angiography was not possible.

In the present series, multi–detector row CT angiography provided accurate depiction of vessel origin and course in this review of 20 anomalous coronary arteries. Most patients were referred after an equivocal coronary angiogram. The results of this study suggest that CT is a viable noninvasive modality in the delineation of coronary arterial anomalies, particularly if results of coronary angiography are equivocal.

References


Safety and Effectiveness of Gadolinium-enhanced Multi–Detector Row Spiral CT Angiography of the Chest: Preliminary Results in 37 Patients with Contraindications to Iodinated Contrast Agents

PURPOSE: To prospectively evaluate the safety and effectiveness of gadolinium-enhanced multi–detector row spiral computed tomographic (CT) angiography of the pulmonary circulation by using two gadolinium doses in patients with contraindications to iodinated contrast agents.

MATERIALS AND METHODS: Study was approved by the Ethics Committee, and written informed consent was obtained. Thirty-seven patients (20 men, 17 women) with contraindications to iodinated contrast agents (allergic reactions, n = 27; impaired renal function, n = 10) underwent CT angiography of the pulmonary circulation in search of acute pulmonary embolism (n = 28) or for management of tumoral disease (n = 9). CT angiography was performed (a) with four–detector row (n = 19) or 16–detector row (n = 18) scanners; (b) at randomly assigned gadolinium doses of either 0.3 mmol per kilogram of body weight (n = 19) or 0.4 mmol/kg (n = 18); and (c) with a systematic evaluation of clinical and biologic tolerance of gadolinium. Comparison of percentages between group 1 and group 2 scans was performed with the χ2 or the Fisher exact test. An unpaired Wilcoxon rank sum test was used for numeric variables. P < .05 was considered to indicate a significant difference.

RESULTS: The mean (± standard deviation) volume of gadopentetate dimeglumine administered in the overall study group was 48 mL ± 9.6 (range, 29–65 mL). The level of maximal enhancement in the pulmonary arteries was significantly higher in group 2 than in group 1 (215.8 HU ± 95 vs 141.3 HU ± 44) (P = .02) and was maintained throughout the entire region of interest in a greater number of examinations in group 2 than in group 1 (n = 16 [89%] vs n = 2 [10.5%]) (P < .0001). The number of diagnostic CT angiograms was significantly higher in group 2 than in group 1 (n = 17 [94%] vs n = 13 [68%]) (P = .007). Significant but transient reduction of creatinine clearance was observed in one patient with preexisting moderate chronic renal failure (0.3 mmol/kg gadolinium dose).

CONCLUSION: High-quality gadolinium-enhanced CT angiograms require the use of 16–detector row CT technology; the doses administered did not alter the renal function except transiently in one patient.

© RSNA, 2005

Iodinated contrast media, while safe in most patients, are associated with a substantial incidence of life-threatening adverse events, particularly in high-risk patients (1). During
the past few years, gadolinium-based con-
trast agents have been used as an alternative
to iodinated contrast agents for patients with
absolute or relative contraindications to io-
dinated contrast agents in examinations
such as intravenous urography, digital
subtraction angiography of various parts
of the body, and numerous interventional
procedures, with resulting adequate image
quality and no side effects (2–6).

The use of gadolinium in computed
tomography (CT) was described in 1989
by Bloem and Wondergem (7), who dem-
strated that the appearance of the uri-
nary collecting system at gadolinium-
enhanced imaging was similar to that at
iodine-enhanced CT in two patients.

Since that initial study, authors of several
articles have evaluated the image quality
and tolerance of gadolinium-based con-
trast material for CT angiography of the
aorta, liver, and cervical and intracranial
vessels in experimental (8–11) and clini-
cal (5,11–15) studies. No adverse events
were reported, and sufficient enhance-
ment for diagnostic purposes was ob-
tained in the investigated subsets of pa-
tients. With regard to CT angiography of
the pulmonary circulation, the use of
gadolinium was reported in a single case
report by Coche et al (16), who per-
formed spiral CT with a dual-detector ar-
ray with successful detection of acute
pulmonary embolism. However, the use
of gadolinium for chest CT requires the
administration of higher doses than are
usually recommended for magnetic reso-
nance (MR) imaging, and the safety of
such doses has not been studied in a large
population of patients. Therefore, the
purpose of the present investigation was
to prospectively evaluate the safety and
effectiveness of gadolinium-enhanced
multi–detector row spiral CT angiogra-
phy of the pulmonary circulation by us-
ing two doses, 0.3 and 0.4 mmol per ki-
ogram of body weight, of gadolinium in
patients with contraindications to iodin-
ated contrast agents.

MATERIALS AND METHODS

Study Protocol

This study was initiated as a random-
ized, monocentric clinical trial aimed at
investigating the effectiveness and safety
of gadolinium-enhanced multi–detector
row spiral CT angiography of the pulmo-
nary circulation. This study was sup-
ported by Schering (Berlin, Germany),
who provided the contrast agent gado-
pentate dimeglumine (Magnevist). A
power injector (Spectris) was provided by
Medrad (Pittsburgh, Pa), which enabled
the authors to perform the injection of
gadopentetate dimeglumine followed by
a saline flush. Those authors who are
not employees of Schering (M.R.J., P.D.,
I.T.L., J.B., A.D., J.R.) had control of in-
cclusion of all data and information that
might present a conflict of interest for
the author (O.E.) who is an employee of
Schering.

Patients who fulfilled the following in-
cclusion criteria were eligible for the
present investigation: indication for a
spiral CT angiographic examination of the
pulmonary circulation for a diagnos-

tic or follow-up purpose and a contrain-
dication to the administration of iodin-
ated contrast agents. Exclusion criteria
included clinical instability, pregnancy,
breast feeding, patients younger than 18
years, impossibility of peripheral vein
catheterization with an 18-gauge cathe-
ter necessary for the high-flow admin-
istration of gadopentetate dimeglumine,
history of hemolytic anemia, or severely
impaired renal function with no possibil-
ity of hemodialysis. Vital signs, various
laboratory serum parameters, and gen-
teral tolerance were monitored to assess
safety and tolerance of the doses of gado-
linium administered in this protocol.

This protocol received the approval of
our Institutional Review Board and Eth-
ics Committee. Written informed con-
sent was obtained from all patients be-
fore they were admitted to the trial.

Population Studied

Between April 2002 and March 2003,
39 consecutive patients were referred to
the Department of Radiology for gadolin-
ium-enhanced spiral CT angiography of
the pulmonary circulation. Two patients
were excluded because of poor venous
access, which led to a final study group of
37 patients (20 men and 17 women). The
mean (± standard deviation) age of the
study group was 56.3 years ± 16 (range,
18–81 years). All patients had relative or
absolute contraindications to iodine-con-
taining contrast media, including known
intolerance to iodinated contrast agents
(n = 19); preexisting impaired renal func-
tion (n = 10); or history of anaphylactic
reactions after drug administration (eg,
aspirin, penicillin), a situation known to
be associated with a substantial incidence
of life-threatening adverse events after
administration of iodinated contrast me-
da (n = 8). Previous history of intolerance
to iodinated contrast media included severe allergic skin reactions (n =
10), laryngeal edema (n = 7), and ana-
phylactic reactions (n = 2). The 10 pa-
tients at risk of iodinated-based contrast
agent–induced nephrotoxicity had chroni-
cally elevated serum creatinine levels
(>1.5 mg/dL [133 μmol/L]) and decreased
creatinine clearance values (<70 mL/min
[1.17 mL/sec]). According to the experi-

ence of the nephrologists involved in this
study (including P.D., with 25 years of
experience), chronic renal failure was
graded as moderate (creatinine cle-
arence of 30–50 mL/min [0.500–0.834 mL/
sec] in four patients, and severe (creati-
inine clearance of <30 mL/min [<0.500
mL/sec]) in four patients. None of these
patients required chronic hemodialysis
treatment. The cause of renal insufficiency
was diabetic nephropathy in four patients,
nephrosclerosis in four patients, myeloma
in one patient, and unknown in one pa-
tient.

According to a computer-generated ran-
domization list, patients were assigned to
one of two dose groups: 0.3 or 0.4 mmol/kg
of gadolinium (0.5 mmol/mL gado-
pentetate dimeglumine). Because of the
upgrade of our equipment during the in-
vestigation, multi–detector row spiral CT
angiography was performed with a four-
detector row scanner for the first 19 con-
secutive patients (group 1: nine women
and 10 men) and with a 16–detector row
scanner for the remaining 18 consecutive
patients (group 2: eight women and 10
men). Indications for multi–detector row
spiral CT angiography included suspi-
ction of acute pulmonary embolism (n =
27) or pretherapeutic evaluation or fol-
low-up of tumoral disease in close con-
tact with the hilum (n = 10). Gadolinium-
enhanced CT angiography was preferred
to ventilation-perfusion scintigraphy for
the diagnosis of acute pulmonary embolism
in our population owing to the presence
of underlying respiratory disease and a
poor diagnostic value of ventilation-per-
fusion scintigraphy in this category of
patients. Patients with a history of an
allergic-like syndrome to iodinated con-
trast agents received the usual preventive
antihistamine and corticosteroid chemoi-
therapy 48 hours prior to the examina-
tion.

CT Examination

Scanning protocol.—The limited maxi-
mum volume of gadolinium that could
be administered required the investiga-
tors to reduce the z-axis coverage to the
middle third of the thorax and to scan
this region in the shortest period of time.
The region surveyed extended from the level of the aortic arch to the level of the inferior pulmonary veins, as previously recommended for single–detector row spiral CT of the pulmonary circulation (17). The scanning parameters for CT angiography were dependent on the multi–detector row spiral CT technology used. With the four–detector row scanner (Sensation 16; Siemens Medical Solutions, Forchheim, Germany), 120 kV, 60–100 mAs, a 0.5-second rotation time, a 4 × 1-mm collimation (four detectors with 1-mm section thickness), and a pitch of 2 with reconstruction of 2-mm-thick scans were used. With the 16–detector row scanner (Sensation 4; Siemens Medical Solutions), 80–120 kV, 70–120 mAs, a 0.5-second rotation time, a 16 × 1.5-mm collimation, and a pitch of 1.5 with reconstruction of 2-mm-thick scans were used. In group 2, the kilovoltage was chosen according to the patient’s body morphology and was set at 80 kV when the patient’s weight was less than or equal to 70 kg (n = 8) and at 120 kV when the patient’s weight was greater than 70 kg (n = 10), with subsequent adaptation of the milliamperage. No dose reduction system was applied for gadolinium-enhanced spiral CT examinations.

CT angiography was systematically preceded by a nonenhanced CT over the entire thorax to provide a complete chest CT examination. Both nonenhanced and gadolinium-enhanced spiral CT scans were acquired in the craniocaudal direction. Contiguous mediastinal and lung images were systematically reconstructed by using high spatial frequency and soft algorithms, respectively. All CT scans were photographed at two window settings appropriate for viewing the lung parenchyma (window width, 1600 HU; window level, −600 HU) and the mediastinum (window width, 350 HU; window level, 30 HU).

Administration of gadolinium.—Each patient received a dose of 0.3 or 0.4 mmol/kg. The corresponding volume of contrast material was administered with a power injector at a flow rate of 6 ml/sec. The choice of the 6-ml/sec flow rate was dictated by the following findings in the literature: (a) Gadopentetate dimeglumine concentration of 0.5 mmol/ml shows approximately the same CT attenuation as does iodine contrast medium at 150 mg/ml. (13) and (b) in a preliminary experience with single–detector row spiral CT angiography of the pulmonary circulation, Remy-Jardin et al (17) reported the administration of 120 mg/ml of iodinated contrast agent at a rate of 6 mL/sec with good to excellent quality of vascular enhancement. The injection of gadopentetate dimeglumine was followed by a saline flush in all cases (15 ml of saline serum administered at 3 ml/sec). Because of the limited amount of gadolinium administered, use of the automatic bolus triggering software program available with the CT units (Care bolus software; Siemens Medical Solutions) could be applied only to group 2 patients. A circular region of interest (4–7 mm in diameter) was positioned (M.R.J.) at the level of the main pulmonary artery, and a threshold for triggering data acquisition was preset at 50 HU. In group 1, the start delay was empirically chosen. We initiated this protocol with a start delay comparable to that recommended for CT angiography of the pulmonary circulation with use of iodine, namely varying between 12 and 15 seconds, which was rapidly found to be unsuited for an adequate opacification of the pulmonary arteries from the first part of the volume scanned. Because of the short transit time of gadopentetate dimeglumine, the start delay was progressively shortened to 6–7 seconds.

Evaluation of the Safety and Effectiveness of Gadolinium-enhanced Spiral CT Examinations

Evaluated parameters.—Because adverse reactions have been reported after administration of gadolinium-based contrast agents (18), we systematically evaluated the patients’ clinical and biologic tolerance of gadolinium. The following three categories of clinical parameters were recorded twice, at baseline (before injection) and 24 hours after administration: (a) pulse rate (beats per minute); the systolic and diastolic blood pressures, both measured after the patient maintained a supine position for at least 5 minutes; and the occurrence of any adverse event, such as sensation of warmth at the injection site, local pain (in case of gadolinium extravasation), or mild to severe allergic reactions that included nausea, vomiting, headaches, flush, allergic skin reaction, severe anaphylactic reaction, or convulsive attack.

Biologic tolerance of gadolinium was assessed with blood samples obtained within 24 hours before (baseline) and 24 hours after administration of gadopentetate dimeglumine. The parameters monitored in the serum were creatinine level, electrolytes, iron, haptoglobin level, blood count, and reticulocytes. For the purpose of the evaluation of the renal tolerance of gadolinium, the primary variable was the change in serum creatinine level after contrast material injection for patients with normal renal function and the change in creatinine clearance after contrast material injection for patients with preexisting impaired renal function. For patients with normal renal function prior to CT angiography, an acute contrast agent–induced reduction in renal function was defined as an increase in the serum creatinine concentration of at least 0.5 mg/dl (44 μmol/L) 24 hours after administration of the contrast agent. For patients with preexisting renal failure, a 10% or greater decrease in the creatinine clearance value within 24 hours after the examination indicated an acute contrast agent–induced reduction in renal function.

The CT parameters evaluated in the present investigation included the following: (a) measurement of the peak level of enhancement within the pulmonary arteries by using a circular region of interest positioned successively in seven central pulmonary arteries—the pulmonary artery trunk, right and left interlobar pulmonary arteries, right and left upper lobe pulmonary arteries, and right and left lower lobe pulmonary arteries; (b) search for a gradient in pulmonary artery enhancement from top to bottom of the volume scanned, which was divided into three zones—the upper zone, extending from the aortic arch to the level of the carina, the middle zone, extending from the carina to the level of the right middle lobe bronchus, and the lower zone, extending from the level of the right middle lobe bronchus to the level of the inferior pulmonary veins; a difference of at least 50 HU between the CT numbers measured in the central pulmonary arteries of the upper and lower lung zones defined the presence of a gradient of arterial enhancement in the volume scanned; and (c) evaluation of the overall image quality, graded as excellent when the peak CT number within the pulmonary arteries was greater than 150 HU with a constant level of enhancement from top to bottom of the volume scanned, good when the peak CT number within the pulmonary arteries was between 100 and 150 HU with a constant level of enhancement from top to bottom of the volume scanned, or poor when the peak CT number within the pulmonary arteries was less than 100 HU and/or a gradient of arterial enhancement was observed in the volume scanned.

The main effectiveness parameter for the comparison of image quality between
groups 1 and 2 was the overall image quality, which led the readers to classify the examinations into two main categories: diagnostic CT angiograms (when the overall image quality was rated as good or excellent) and nondiagnostic CT angiograms (when the overall image quality was rated as poor).

Conditions of evaluation of clinical, biologic, and CT parameters.—During this investigation, the clinical and biologic tolerance of gadolinium-enhanced multidetector row spiral CT angiography was first investigated by the radiologist in charge of the examination and then by the referring physician, with each patient’s clinical and biologic results systematically recorded on a case report form by the principal investigator (M.R.J.). To assess the immediate general clinical tolerance, the patients were closely observed during the examination and were nonsuggestively questioned about their well-being at the end of the CT examination and approximately 30 minutes after the injection. The patients were blinded to the dose of gadolinium administered. All CT examination scans were prospectively read during the course of clinical work-up by the principal investigator, with the objective of participating in the patient’s clinical management. The diagnosis of endoluminal clots was based on the recognition of segmental filling defects (72 HU) surrounded by contrast material–enhanced blood (145 HU).

For the purpose of the present study, results of gadolinium-enhanced CT examinations were reinterpreted in consensus between two faculty radiologists (M.R.J., J.R., with 15 and 20 years of experience, respectively, in interpreting chest CT). The two readers were not blinded to the differences in CT technology. However, they were not aware of the dose of gadolinium administered nor of clinical data, in particular the presence of renal dysfunction at the time of CT angiography. A retrospective analysis of the biologic tolerance of gadolinium was undertaken by a nephrologist (P.D.) involved in the study design, who was blinded to the injected dose of gadolinium.

Statistical Analysis

Statistical analysis was performed with commercially available software (SAS Institute, Cary, NC). Data were expressed as the mean ± standard deviation for continuous variables and as frequencies or percentages for categoric variables. Comparison of percentages between group 1 and group 2 scans was performed by using the χ² or the Fisher exact test. An unpaired Wilcoxon rank sum test was used for numeric variables. P < .05 was considered to indicate a statistically significant difference.

RESULTS

Image Quality of Gadolinium-enhanced Spiral CT

In the overall study group, the mean volume of gadopentetate dimeglumine administered was 48 mL ± 9.6 (range, 29–65 mL), the mean z-axis coverage was 130.94 cm ± 41.73 (range, 64–236 cm), the mean duration of data acquisition was 5.7 seconds ± 2.91 (range, 2.2–15 seconds), the mean value of the peak level of enhancement within the pulmonary arteries was 183.53 HU ± 83.9 (range, 79–363 HU), and 30 (81%) CT angiograms were graded as diagnostic and seven (19%) as nondiagnostic. Table 1 summarizes the technical conditions of CT angiographic examinations in the two groups of patients. No statistically significant difference was found in the mean volume of gadolinium administered nor in the mean start delay between group 1 and group 2. Significant differences were observed between the two groups in the mean z-axis coverage and the duration of data acquisition, which were longer and faster, respectively, in group 2 compared with group 1.

The quality of arterial enhancement on gadolinium-enhanced spiral CT angiograms is summarized in Table 2. The mean value of the peak level of enhancement of pulmonary arteries was significantly higher in group 2 than in group 1 (P = .02). A constant degree of pulmonary artery enhancement within the volume scanned was observed in 16 (89%) of

### TABLE 1

<table>
<thead>
<tr>
<th>Technical Conditions of CT Angiographic Examinations</th>
<th>Group 1 (n = 19)</th>
<th>Group 2 (n = 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gadolinium dose (mmol/kg)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>11 patients</td>
<td>8 patients</td>
</tr>
<tr>
<td>0.4</td>
<td>8 patients</td>
<td>10 patients</td>
</tr>
<tr>
<td><strong>Volume of gadopentetate dimeglumine (mL)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48.38 ± 9.5 (29–60)</td>
<td></td>
<td>48.0 ± 10 (30–65)</td>
</tr>
<tr>
<td><strong>Duration of start delay (sec)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.27 ± 3.8 (2–15)</td>
<td></td>
<td>8.72 ± 3.8 (2–14)</td>
</tr>
<tr>
<td><strong>Duration of data acquisition (sec)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.64 ± 3.2 (2.7–15)</td>
<td></td>
<td>4.4 ± 1.4 (2.2–7.4)</td>
</tr>
<tr>
<td><strong>Z-axis coverage (cm)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>102.0 ± 28.6 (64–154)</td>
<td></td>
<td>159.0 ± 33.6 (97–236)</td>
</tr>
</tbody>
</table>

* Data are the mean ± standard deviation. Data in parentheses are the range.  † P < .001.
‡ P < .0001.

### TABLE 2

<table>
<thead>
<tr>
<th>Quality of Pulmonary Artery Enhancement on Gadolinium-enhanced Spiral CT Angiograms</th>
<th>Group 1 (n = 19)</th>
<th>Group 2 (n = 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peak level of pulmonary artery enhancement (HU)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>141.3 ± 44 (79–222)</td>
<td></td>
<td>215.9 ± 95 (88–363)</td>
</tr>
<tr>
<td><strong>Degree of arterial enhancement from top to bottom of the volume scanned</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>2 (10.5)</td>
<td>16 (89)</td>
</tr>
<tr>
<td>Presence of a gradient</td>
<td>17 (89.5)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Lower enhancement in lower third of the volume scanned</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Lower enhancement in upper third of the volume scanned</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td><strong>Overall image quality</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Good</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Excellent</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td><strong>No. of diagnostic CT angiograms</strong></td>
<td>13 (68)</td>
<td>17 (94)</td>
</tr>
</tbody>
</table>

* Data are the mean ± standard deviation. Data in parentheses are the range.  † P = .02.
‡ Data are the number of angiograms. Data in parentheses are percentages.  † P < .0001.
§ Data are the number of patients.
‡ Data in parentheses are percentages.  † P = .007.
among the 27 patients referred because of suspicion of acute pulmonary embolism, seven patients had a nondiagnostic CT scan (six patients in group 1 and one patient in group 2) because of poor level of contrast enhancement within the pulmonary arteries. In the remaining 20 patients, CT scans were positive for pulmonary embolism in one patient and negative in 19 patients. The 10 patients referred for pretherapeutic evaluation or follow-up of tumoral disease had a diagnostic CT scan.

Clinical and Biologic Tolerance of Gadolinium

No adverse events occurred after intravenous injection of the contrast agent at either dose, including allergic-like reactions in patients with history of intolerance to iodinated contrast agents. The examination was considered painless, and there was no extravasation at the site of venous administration of gadolinium. Table 3 summarizes the renal tolerance of the doses of gadolinium administered. A total of 18 patients (nine patients in both groups) received the 0.3 mmol/kg dose. A total of 19 patients (10 in group 1 and nine in group 2) received the 0.4 mmol/kg dose. Twenty-seven patients had normal renal function prior to gadolinium-enhanced CT angiography, and 10 patients had preexisting renal insufficiency. No impairment in renal function was observed except in one patient with preexisting diabetic nephropathy who received a dose of 0.3 mmol/kg. The creatinine clearance decreased from 5.3 mL/min (0.088 mL/sec) to 4.3 mL/min (0.072 mL/sec) 24 hours after the administration of the contrast agent and returned to baseline values within 3 days. No variations in hematologic and clinical chemistry results were found in the studied population during the observation period.

DISCUSSION

Results of this study demonstrate that the use of gadolinium-enhanced contrast medium in place of iodinated contrast medium in high-risk patients provides a quality of pulmonary artery enhancement sufficient to give diagnostic information in most cases, thus confirming in a relatively large series of patients the preliminary findings reported by Coche et al (16). In 30 (81%) of the 37 patients examined in our study, gadolinium-enhanced CT angiograms were graded as diagnostic on the basis of a good to excellent level of arterial enhancement, with a constant level of enhancement from top to bottom of the volume scanned. However, we observed that the rate of diagnostic examinations was highly dependent on the multi–detector row CT technology used. Because of the upgrade of our equipment during the study, multi–detector row spiral CT angiography was performed with a four–detector row scanner for the first 19 patients (group 1) and with a 16–detector row scanner for the remaining 18 patients (group 2). The number of diagnostic CT angiograms was significantly higher in group 2 (17 of 18 examinations, 94%) than in group 1 (13 of 19 examinations, 68%) (P = .007) mainly because of a higher level of pulmonary artery enhancement, otherwise kept constant from top to bottom of the volume scanned in a majority of patients. The rapidity of data acquisition available with a 16–detector row CT equipment is the major determinant for such an improvement in image quality.

First, the mean duration of data acquisition was significantly shorter in group 2 than in group 1 (4.4 vs 7.6 seconds, P < .001), a key element when considering the constraint represented by the limited amount of gadopentetate dimeglumine.
administered to the patient; this amount varied between 29 and 65 mL. Contrast medium delivery with a power injector is mandatory, but the small total volume of undiluted gadolinium places a restriction for its use with four–detector row CT equipment. Second, the speed of 16–detector row CT made possible the use of an automatic bolus triggering software program for each patient in group 2, whereas the start delay had to be empirically chosen in all group 1 patients. Consequently, an adequate and constant level of enhancement within pulmonary arteries was obtained from the first anatomic level scanned throughout the entire region of interest in all but two patients in group 2 (89%), while this was possible only in two (10.5%) of the 19 patients in group 1. In the remaining 17 patients in group 1, the gradient of arterial enhancement observed within the volume scanned was easily related to an inadequate selection of the scanning delay. This scanning delay was too short in 10 cases, with the highest level of opacification reached only at the end of data acquisition, and too long in seven cases, resulting in an insufficient amount of contrast material for a correct opacification of the lower portion of the region of interest.

In the absence of significant differences in the mean volume of gadopentetate dimeglumine administered between the two groups of patients, the speed of the 16–detector row CT equipment is the most likely explanation for the higher mean peak enhancement level within the pulmonary arteries in group 2 (215.9 HU) compared with group 1 (141.3 HU). The enhancement level with gadopentetate dimeglumine was otherwise in the same range as that achievable with iodinated contrast media. It should be emphasized that the mean CT numbers within the central pulmonary arteries in groups 1 and 2 were higher than those reported by Coche et al (16), who reported CT numbers reaching 103 HU at the level of the main pulmonary artery after administration of 60 mL of gadodiamide. In that study, the gadolinium-enhanced CT examination was performed by using a spiral CT scanner with a dual-detector array with a 2 × 2.7-mm collimation, a table increment of 7.5 mm per rotation, and a 1-second rotation time from the level of the aortic arch down to the diaphragm, which led to a longer scanning time compared with that in our study. It should be noted that eight patients from group 2 also benefited from another technologic improvement, namely the availability of scanning at 80 kV. By keeping in mind that the attenuation by gadopentetate dimeglumine increases with decreasing tube voltage when measured from 80 to 137 kV (19), scanning patients at 80 kV may have also contributed to the improvement in the overall image quality of gadolinium-enhanced CT scans in group 2.

Use of gadolinium as an alternative contrast agent for multi–detector row spiral CT examinations of the chest requires the administration of higher doses than are usually recommended for MR imaging, which vary between 0.1 and 0.2 mmol/kg of a 0.5 mmol/mL gadolinium-based contrast agent. Whereas the triple dose of 0.3 mmol/kg of gadolinium-based contrast agent can be used for MR imaging of brain metastases (20) and active lesions in multiple sclerosis of the brain and spinal cord (21), it has not received formal approval for other MR indications nor for use in CT. The debate over the administration of high doses of gadolinium is centered on the possibility of providing satisfactory diagnostic x-ray studies, which are often questioned in the literature, and on the safety of such doses, especially in patients with underlying renal insufficiency (22).

The present investigation helps provide preliminary answers to these important clinical questions. In the first part of this study, we demonstrated that it was possible to provide diagnostic images when using gadolinium-based contrast agent for multi–detector row spiral CT examinations of the chest, provided that the 16–detector row technology was available. The second objective of our study was to evaluate the clinical and biologic tolerance of gadolinium doses of 0.3 and 0.4 mmol/kg. With regard to clinical tolerance, we observed no adverse events after intravenous injection of the contrast agent
at either dose, including allergic-like reactions in patients with a history of intolerance to iodinated contrast agents. The examination was considered painless by all patients, and there was no extravasation at the site of venous administration of gadopentetate dimeglumine, despite the administration at a rate of 6 mL/sec. With regard to the renal tolerance of the two doses of gadolinium administered, we failed to observe any increase in the serum creatinine level among the 27 patients with normal renal function prior to the gadolinium-enhanced CT examination. Among the 10 patients with preexisting renal insufficiency graded as moderate in two cases, marked in four cases, and severe in four cases, the doses of 0.3 and 0.4 mmol/kg resulted in no adverse effects on renal function except in one patient in whom a transient reduction of creatinine clearance was observed. This patient had preexisting moderate chronic renal failure related to diabetic nephropathy and received the 0.3 mmol/kg dose. The creatinine clearance returned to baseline values within 3 days.

This preliminary clinical experience in a population of 37 patients suggests that gadolinium-based contrast media are well tolerated, even in patients with underlying renal insufficiency, when gadolinium is administered at doses of 0.3 or 0.4 mmol/kg. Although intravenous injection of gadopentetate dimeglumine has been proved to be safe for doses as high as 0.5 mmol/kg even in patients with renal insufficiency (14,23–26), the safety of such doses had not been previously studied in a large population of patients. Nevertheless, it should be pointed out that the limited number of patients with renal failure in the present investigation is not sufficient to draw definitive conclusions about the renal tolerance of gadolinium in this specific population.

Several limitations of this study should be mentioned. First, we tailored the scanning protocol to that proposed for single–detector row spiral CT of the pulmonary circulation by using a low concentration of iodinated contrast material. This led us to administer the contrast agent at a rate of 6 mL/sec and to exclude two patients with poor venous access to avoid venous damage and extravasation of contrast medium. Second, the upgrade of our CT equipment led us to modify the scanning parameters not only between the groups but also among patients in group 2, a few of them having been able to benefit from the availability of a lower kilovoltage. These changes in the scanning protocol over time make a strict comparison between the two groups of patients difficult. Third, in our protocol, we investigated gadolinium-induced reduction in renal function with biologic parameters controlled 24 hours after CT angiography, whereas the time interval often considered in the literature dealing with acute contrast agent–induced renal failure is 48 hours (27–30). The risk of underestimation of contrast agent–induced renal failure in our protocol is outweighed by the fact that our population was composed of inpatients who underwent periodic biologic evaluation. None of the evaluations obtained within 1 week of the gadolinium-enhanced CT examination as part of the clinical management revealed changes in the patients’ renal function, especially in patients with underlying renal insufficiency.

In conclusion, gadolinium-based contrast agents can provide diagnostic CT angiograms of the pulmonary circulation, but high-quality examinations require the use of 16–detector row CT technology. The doses of 0.3 and 0.4 mmol/kg were found to be well tolerated, even in patients with underlying renal insufficiency.

References
12. Quinn AD, O’Hare NJ, Wallis FJ, Wilson GF. Gd-DTPA: an alternative contrast me-
contrast agent for use in CT during in-
traarterial administration. J Comput As-
14. Pena CS, Kaufman JA, Geller SC, Walt-
man AC. Gadopentate dimeglumine: a
possible alternative contrast agent for CT
angiography of the aorta. J Comput Assist
15. Karcaaltincaba M, Foley WD. Gadolinium-
enhanced multidetector CT angiography
of the thoracoabdominal aorta. J Comput
Assist Tomogr 2002; 26:875–878.
16. Coche EE, Hammer FD, Gofette PP. Dem-
onstration of pulmonary embolism with
dynamic gadolinium-enhanced spiral CT.
17. Remy-Jardin M, Remy J, Wattiinne L, Gi-
raud F. Central pulmonary thromboem-
bolism: diagnosis with spiral volumetric
CT with the single breath hold tech-
nique—comparison with pulmonary an-
18. Shellock FG, Kanal E. Safety of magnetic
resonance imaging contrast agents. J
19. Schmitz SA, Wagner S, Schuhmann-Gi-
ampieri G, Wolf KJ. Evaluation of gadobu-
trol in a rabbit model as a new lanthanide
contrast agent for computed tomography.
20. Kuhn MJ, Hammer GM, Swenson LC,
Youssef HT, Gleason TJ. MRI evaluation of
“solitary” brain metastases with triple-
dose gadoteridol: comparison with con-
trast-enhanced CT and conventional-dose
gadopentate dimeglumine MRI studies in
the same patients. Comput Med Imaging
multi-centre longitudinal study compar-
ing the sensitivity of monthly MRI after
standard and triple dose gadolinium-DTPA
for monitoring disease activity in multiple
2020.
22. Thomsen HS, Almen T, Morcos SK. Gad-
olinium-containing contrast media for
radiographic examinations: a position
Safety and tolerance after intravenous ad-
ministration of 0.3 mmol/kg Gd-DTPA:
results of a randomized, controlled clini-
cal trial. Invest Radiol 1991; 26(suppl 1):
S221–S223.
al. Renal tolerance of Gd-DTPA dime-
glumine in patients with chronic re-
156.
25. Prince MR, Arnoldus C, Frisoli JK. Neph-
rotoxicity of high-dose gadolinium com-
pared with iodinated contrast. J Magn Re-
cokinetics, safety, and tolerability of
gadoversetamide injection (Optimark) in
subjects with central nervous system or
liver pathology and varying degrees of
renal function. J Magn Reson Imaging
27. Katzberg R. New and old contrast agents:
physiology and nephrotoxicity. Urol Ra-
28. Solomon R. Contrast-medium induced
29. Lasser EC, Lyon SG, Berry CC. Reports on
contrast media reactions: analysis of data
from reports of the U.S. Food and Drug
Administration [published correction ap-
ppears in Radiology 1997; 204:876]. Ra-
30. Tublin ME, Murphy ME, Tesler FN. Cur-
rent concepts in contrast media-induced
nephropathy. AJR Am J Roentgenol 1998;
171:933–939.
Case 87

HISTORY

A 59-year-old man with non-insulin-dependent diabetes mellitus underwent laparoscopic cholecystectomy for biliary stones. He was able to return to work 8 days after surgery. The patient started to experience numbness and tingling in all four limbs 2–3 weeks later. The symptoms progressed to the point that he had difficulty driving, walking, and using his hands. Bladder function was intact. The patient was hospitalized for a complete neurologic evaluation 8 weeks after surgery. A cervical spine magnetic resonance (MR) imaging examination (Figs 1–3) was performed at that time.

Figure 1. Sagittal T2-weighted fast spin-echo MR image (repetition time msec/echo time msec, 3894/130) of the cervical spinal cord.

E-mail the most likely diagnosis to dxplease@rsna.org (use only for submission of diagnosis). Include case number, your name (as you would want it to appear in the journal), address, phone and fax numbers. Only one case, one name, and one diagnosis per e-mail submission. Multiple diagnoses, multiple submissions, submissions without a case number will not be considered. Deadline: Midnight U.S. Central Time, August 15, 2005. Answer will appear in the October issue. Authors wishing to submit cases for Diagnosis Please should first write to the Editor to obtain approval for the case and further information.
Figure 2. Transverse T2-weighted fast spin-echo MR images (5703/130). (a–c) Images were obtained through the cervical spinal cord at three separate levels from C3 through C4.

Figure 3. Transverse T1-weighted fast spin-echo MR image (747/12) obtained after contrast material administration through the cervical spinal cord at the level of C4.
Case 83: Multifocal Fibrosclerosis with Mediastinal-Retroperitoneal Involvement

HISTORY

A 72-year-old man with a history of congestive heart failure, malignant hypertension, and peripheral and coronary vascular disease was admitted to the hospital 4 days after onset of acute dyspnea and lower-extremity edema. He underwent coronary catheterization; after this procedure, his serum creatinine levels increased from 1.6 to 5.9 mg/dL (14.1–52.1 μmol/L). He was subsequently referred for renal and aortic magnetic resonance (MR) imaging to allow more comprehensive evaluation of renal and vascular disease.

Prior to admission, the patient had undergone coronary artery bypass graft placement for triple-vessel coronary artery disease (5 years previously) and right carotid endarterectomy (1 year previously). Pathologic specimens of the carotid artery showed dense fibrous tissue and chronic inflammation surrounding the vessel wall.

IMAGING FINDINGS

T1- and T2-weighted MR images of the abdomen demonstrated intermediate soft-tissue signal intensity in the subcapsular and perirenal space that circumferentially envelops the kidneys (Figs 1, 2). There was also tissue with the same MR imaging characteristics surrounding the abdominal aorta, left renal artery (Fig 3), and aortic arch (Fig 4).

DISCUSSION

Multifocal fibrosclerosis is a rare syndrome of unknown cause that is characterized by fibrosis involving multiple organ systems. Comings et al (1) reported the occurrence of different combinations of retroperitoneal fibrosis (RPF), mediastinal fi-
In the case presented, and RPF is the most common manifestation of the mediastinal anatomy (2,4).

Depending on the involved organ systems, different clinical and radiologic manifestations are observed in patients with multifocal fibrosclerosis. In the case presented, there is extensive soft-tissue proliferation, most extensively surrounding the kidneys, but there is also involvement of the abdominal aorta, renal arteries, and thoracic aorta. This patient additionally had soft-tissue fibrosis around the carotid arteries (not shown, but stated in the history). While the abdominal findings alone are typical of perirenal fibrosis and RPF, their combination with the same pattern of disease in the mediastinal and cervical areas denotes generalized systemic involvement; this is termed multifocal fibrosclerosis. These findings were confirmed by means of core biopsy of the tissue surrounding the kidneys and examination of the endarterectomy specimen. In this patient, symptoms of congestive heart failure and coronary artery disease were caused by atherosclerotic disease; atherosclerotic narrowing of the renal ostia also likely contributed to the patient’s renal failure and hypertension.

Mediastinal fibrosis in multifocal fibrosclerosis is characterized by hypointense tissue on T1- and T2-weighted MR images in the mediastinum, especially around the aorta, as in this case. Mediastinal fibrosis can also be associated with superior vena cava obstruction, stenosis of the pulmonary veins, tracheal obstruction (eg, symptoms similar to asthma), esophageal stricture, pulmonary artery obstruction with cor pulmonale, and coronary artery occlusion (1,4). Histologically, it has a similar appearance to that observed in patients with RPF (1,13). Cross-sectional imaging findings are those of an infiltrative soft-tissue process, with variable degrees of deformation of the mediastinal anatomy (2,4).

RPF was a prominent component of multifocal fibrosclerosis in the case presented, and RPF is the most common manifestation of multifocal fibrosclerosis. RPF is characterized by fibrous tissue proliferation surrounding the aorta and extending into periaortic tissue as a plaque-like infiltrative soft-tissue process (14). In most cases, it is localized around the lower aorta and common iliac arteries in the retroperitoneal space (12), but the pelvic cavity can also be involved (5). About two-thirds of the cases are considered idiopathic; in the remaining one-third, an inciting factor (eg, methysergide, malignancies, radiation therapy, surgery, hemorrhage, retroperitoneal infection, autoimmune diseases) can be identified (15). It has also been suggested that RPF could be secondary to a local autoallergic reaction to components of atherosclerotic plaques, occurring in areas in which the aorta has severe atherosclerotic plaque (13). There is a male-to-female ratio of approximately 3:1. In 70% of patients, the age at diagnosis is between 30 and 60 years, but the disease has been described in children (16).

Two histologic patterns of RPF have been described. One pattern consists of chronic fibrosis that is relatively avascular, cellular, and often calcified. The second pattern demonstrates collagen bundles interspersed with an equal or greater volume of inflammatory cells and mucopolysaccharide. In the later group, small blood vessels are numerous, and the infiltrative process is pleomorphic, consisting of lymphocytes, plasma cells, eosinophils, Russell bodies, fibroblasts, and mast cells. Neutrophils are absent (12). It is hypothesized that the more active type of inflammation matures into the more chronic histologic pattern.

The diagnosis of retroperitoneal fibrosis is often suggested by imaging studies performed to enable the evaluation of nonspecific clinical signs and symptoms. Medial deviation and narrowing of the ureters near the lower lumbar spine (17) demonstrated by excretory urography represents a classic radiologic presentation of RPF. Renal failure may be profound, despite low-grade obstructive uropathy. One proposed mechanism of obstruction is interference with ureteral peristaltic activity rather than mechanical obstruction (18).

Sonographic findings of RPF consist of hydronephrosis and visualization of fibrotic tissue as a hypoechoic mass in the para-aortic region (19). Computed tomography (CT) shows retroperitoneal plaque-like, infiltrative soft-tissue surrounding vessels and ureter. The soft-tissue attenuation seen on CT scans is similar to that of muscle (20). Enhancement is variable after intravenous administration of iodinated contrast material. MR
images show an infiltrative soft-tissue process with homogeneously low signal intensity compared with that of the adjacent psoas muscle seen on T1-weighted images (21). However, RPF can show moderate to high signal intensity on T2-weighted images, depending on the degree of associated inflammatory response. Extensive inflammation early in the disease process results in increased T2 signal intensity (21,22). In a series of 17 patients, malignant RPF was characterized by ill-defined margins, increased signal intensity on T2-weighted images, and heterogeneous MR signal intensity (23). MR imaging characteristics, however, are not specific for distinguishing benign from malignant cause. Therefore, biopsy is essential to confirm the diagnosis. Fluorodeoxyglucose positron emission tomography can play a role in establishing the diagnosis and sites of involvement in patients with multifocal fibrosclerosis, and it could be helpful in evaluating disease activity and patient response to corticosteroid therapy by showing high fluorodeoxyglucose uptake in involved areas (likely caused by active inflammation) (24).

Localized extensions of RPF (19,22,25–32) may be associated with more typical presentations (eg, periaortic disease) that may lead to the correct imaging diagnosis. In the case presented, subtle periaortic abnormalities were present and suggest the diagnosis of RPF.

The primary diagnostic consideration in this case would be treated lymphoma. The other differential diagnoses usually considered for RPF, including sarcoma, metastatic lymphadenopathy, amyloidosis, and extramedullary hematopoiesis (22,33,34), would not appear as hypointense soft tissue on T1- and T2-weighted images around the kidneys and thoracic and abdominal aorta. A particular characteristic of RPF is the tendency to infiltrate and envelop surrounding structures without displacing them. Malignancies, such as lymphomas and sarcomas, typically cause displacement of surrounding structures, especially the aorta (21,35). In addition, the anterior margin of RPF is clearly delineated, respecting peritoneal boundaries, while the posterior margin is poorly defined and not easily separated from subjacent structures (20).

Therapy for multifocal fibrosclerosis includes removal of any identifiable inciting agents, including a search for occult tumors, and suppression of inflammatory process. Steroid therapy is of value in most cases when there is no associated malignancy. In steroid-resistant cases, other forms of immunosuppression may be useful. Medroxyprogesterone acetate, progesterone, and tamoxifen have been described as alternative forms of treatment. Surgical intervention is considered for steroid-resistant cases and for patients who do not tolerate steroids. Long-term follow-up is indicated, as recurrences are unpredictable and may occur from 3 months to more than 10 years after the initial diagnosis and treatment (15).

Acknowledgment: The authors gratefully acknowledge the assistance provided by Tracy L. McCracken, RN, CCRC, in the preparation of the manuscript.

References
Congratulations to the 20 individuals and one resident group that submitted the most likely diagnosis (multifocal fibrosclerosis with mediastinal-retroperitoneal involvement) for Diagnosis Please, Case 83. The names and locations of the individuals and of the resident group, as submitted, are as follows:

**Individual responses**
Albert J. Alter, Madison, Wis
Juan Ramón Ayuso, Barcelona, Spain
Richard Beedle, Auckland, New Zealand
Michael P. Buetow, MD, Okemos, Mich
Antonio Cavalcanti, MD, São Paulo, Brazil
Ann S. Fulcher, MD, Richmond, Va
Myeong-Jin Kim, MD, PhD, Seoul, Korea
Yu-Ting Kuo, MD, Kaohsiung, Taiwan
David A. Lisle, Brisbane, Australia
Manabu Minami, MD, Ibaraki, Japan
Eduardo Mondello, MD, Buenos Aires, Argentina

**Resident group response**
Oregon Health and Science University Radiology Residents, Portland, Ore

Mizuki Nishino, MD, Boston, Mass
Michael T. O’Loughlin, MD, West Hartford, Conn
Taro Shimono, MD, Osaka, Japan
Kouichi Sugiya, Hamamatsu, Japan
Norio Takahashi, MD, Fuku, Japan
Hiroyuki Ueda, Kyoto, Japan
P. M. Vos, Vancouver, British Columbia, Canada
Satoru Yoshida, MD, Muroran City, Japan
Joe Yut, Olathe, Kan
Intermittent Claudication: Functional Capacity and Quality of Life after Exercise Training or Percutaneous Transluminal Angioplasty—Systematic Review

**PURPOSE:** To systematically review published data about the short- and long-term effects of exercise training and angioplasty on functional capacity and quality of life of patients with intermittent claudication.

**MATERIALS AND METHODS:** Articles published between January 1980 and February 2003 were included if patients had intermittent claudication treated with exercise training or angioplasty and if both functional capacity and quality-of-life scores from Medical Outcomes Study 36-Item Short Form health survey were reported for at least 3 months of follow-up. Data were pooled by using a random effects model and weighted means. Pooled results were compared between the treatment groups by using the $\chi^2$ test and the Student t test ($\alpha = .05$, two sided).

**RESULTS:** In the analyses, five studies (202 patients) were included in the exercise group, and three studies (470 patients), in the angioplasty group. At 3 months of follow-up, the ankle-brachial index was significantly improved in the angioplasty group (mean change, 0.18; $P < .01$) but not in the exercise group (mean change, 0.01; $P = .29$). At 3 months, quality of life was significantly improved with regard to ratings of physical functioning and bodily pain in the exercise group (mean change, 18 and 10, respectively; $P < .01$) and physical role functioning in the angioplasty group (mean change, 30; $P = .03$). Mean change in ankle-brachial index significantly differed between the two treatment groups at 3 and 6 months ($P < .01$); mean change in quality-of-life scores did not.

**CONCLUSION:** Improvement in quality of life was demonstrated after both exercise training and angioplasty, whereas functional capacity showed significant improvement after angioplasty. The ankle-brachial index significantly differed between the two treatment groups at 3 and 6 months, whereas the quality-of-life scores did not.

Intermittent claudication is the mildest manifestation (ie, Rutherford stage I, II, or III) of peripheral arterial disease, and its prevalence is approximately 5% in men older than 50 years (1,2). The overall aging of populations in Western societies indicates that the prevalence of this condition will increase over the next decades (3). Various treatment strategies, such as exercise training and percutaneous transluminal angioplasty (PTA) with or without stent placement, have been proposed in order to ameliorate symptoms. The choice between exercise training and PTA in patients with intermittent claudication, however, has been the subject of controversy. To our knowledge, there has been no direct comparison of exercise training with PTA to determine the relative effects of these two therapies both on functional capacity (usually expressed as ankle-brachial index [ABI] or as walking distance) and on quality of life.
In 1957, Foley (4) described the beneficial effects of daily walks and physical training in patients with intermittent claudication. Several subsequent reports (5–7) confirmed that exercise training led to an improvement in the walking ability of patients with intermittent claudication. It has been shown to significantly improve both the onset of leg discomfort (initial walking distance) and the point at which the pain becomes unbearable and forces the patient to stop walking (maximum walking distance) (8,9). PTA became available in 1964 and is currently an established treatment option for patients with intermittent claudication (10). Low complication rates (4%–5%) and 5-year patency rates of 80% for aortoiliac disease and 60% for femoropopliteal disease were reported for PTA (11–14). This technique has been popularized as an inexpensive, effective, and simple method (15–17).

In recent years, quality of life has been extensively used as a parameter for assessing and expressing the success of the outcome in patients with intermittent claudication (18). The TransAtlantic Inter-Society Consensus stated that if quality of life could be accurately assessed, it would be the ideal primary end point (19). Furthermore, the relationship between functional capacity and quality of life is complex and contradictory, and both therefore should be considered in the assessment of outcome (20,21).

There remains uncertainty surrounding the effectiveness of the treatment strategies in patients with intermittent claudication. Thus, the objective of our study was to systematically review the published data about the short- and long-term effects of exercise training and angioplasty on functional capacity and quality of life of patients with intermittent claudication.

**MATERIALS AND METHODS**

**Data Sources and Data Extraction**

Publications were initially selected by the first author (S.S.), and the selection was verified by a coauthor (J.L.B.). The final selection was made by both authors together. Both authors evaluated all reports independently for inclusion and exclusion criteria.

A literature search was performed by using MEDLINE and the reference lists in articles. Further searches were performed by using the same keywords in the Cochrane Central Register of Controlled Trials, Cochrane Anesthesia Group Specialized Register, Cumulative Index to Nursing and Allied Health Literature, and PiCarta.

The reviewers used the English-language medical literature published in 1980 or later (22). Published studies of PTA dated from 1980, and to find contemporary controls between the exercise group and the PTA group, we restricted our review to this period. Randomized controlled trials, prospective cohort studies, and retrospective cohort studies were included, while case reports and reviews were excluded. Key words were (“intermittent claudication” OR “claudicants”) AND (“angioplasty” OR “PTA” OR “stent placement” OR “balloon dilation” OR “exercise” OR “gymnastics” OR “walking” OR “training”) AND (“quality of life” OR “Qol” OR “health status”) AND (“functional capacity” OR “ankle brachial index” OR “walking distance”) AND (“vascular peripheral” OR “arterial peripheral”). The studies were included if they met the following criteria: Patients with intermittent claudication were included and were treated with either exercise training (walking or gymnastics) or with PTA with or without stent placement in lesions in the aortoiliac or femoropopliteal arterial segments. Both functional capacity as demonstrated by walking distance (on a treadmill or as reported by the patient) or ABI, and quality-of-life scores (generic or disease specific) for at least 3 months of follow-up, were reported. Reports about PTA studies that also included patients with critical ischemia were excluded from our analysis if the outcome results were not reported separately for patients with critical ischemia and for those with intermittent claudication (23,24). When the same group of investigators reported their results in various journals, all reports were examined for similarities and completeness, and the results were treated as those of a single study in our review (25,26).

Data were abstracted independently by the same two authors who selected the publications. The following parameters were recorded: study design; patient characteristics; measures of severity of claudication, such as Rutherford classification, walking distance, and ABI; type of exercise training or intervention; outcomes; and follow-up results reported by using standardized forms. The exercise training programs were evaluated by the same two authors, and discrepancies between the evaluations were resolved with consensus.

**Functional Capacity and Quality of Life**

The ABI at rest and the maximum treadmill walking distance are the most frequently reported outcomes of functional capacity. The ABI is defined as the ratio of the systolic blood pressure measured at one of the ankle arteries to the systolic blood pressure measured at the brachial artery (9). The most efficient and reliable test for the determination of the maximum walking distance is considered to be a treadmill test (27).

The Medical Outcomes Study 36-Item Short Form (SF-36) health survey was most frequently used in studies in which investigators evaluated quality of life after exercise training or PTA in patients with intermittent claudication. The SF-36 survey was developed to evaluate the physical, social, and physical role functioning of patients and to elicit their perceptions of their general health and well-being in eight dimensions (28). Each dimension is subjectively rated by the respondent and assigned a score on a scale of 0–100 in which 100 indicates best functioning or well-being. In previous studies, it was demonstrated that the dimensions of physical functioning, bodily pain, and physical role functioning showed the most substantial improvement after treatment in patients with intermittent claudication (29). In this study, therefore, we concentrated on the following three dimensions: physical functioning (limitation of usual activities because of a physical problem), physical role functioning (limitation of usual role activities because of a physical problem), and bodily pain (pain or discomfort in the body). In addition, we added the dimension of general health. If data from the SF-36 survey were not reported but data from SF-20 (a shorter version of the same survey) were available, the SF-20 data were used. For studies of exercise training, data beyond 6 months of follow-up were not available, and for studies of PTA, data at 6 months were poor. The weighted means for functional capacity and the quality-of-life scores in the PTA studies did not change significantly between 3 months and 12 months of follow-up. Therefore, to improve the comparison between the exercise training and PTA studies at 6 months of follow-up, we used the data obtained at 12 months of follow-up to calculate the weighted means at 6 months.
Statistical Analysis of Data

To pool data for discrete variables such as male sex, cardiac disease, hypertension, diabetes mellitus, pulmonary disease, hyperlipidemia, stroke, history of smoking, and current smoking, the random effects model described by Laird and Mosteller (30) was used. The random effects model takes into account the variance between studies, as well as that within studies. Because of the lack of reported data for variance around the means of continuous variables such as age, ABI, walking distance, and quality-of-life score, we were unable to use the random effects model to pool the continuous variables. Therefore, we calculated the weighted mean and 95% confidence interval (CI) for these variables.

Within the treatment study groups of exercise and PTA, the dichotomous baseline characteristics of all patients in the included studies were tested for homogeneity by using the χ² test (statistical significance level, α = .05). Differences in pooled patient characteristics between the studies of PTA and those of exercise rehabilitation group and nonexercise (control) group.

Results

Selected Articles and Comparisons

From the literature search, a total of 380 abstracts published between January 1980 and February 2003 were selected. Many studies (n = 348) were excluded on the basis of the abstract for the following reasons: There was no relation with peripheral arterial disease (n = 191); drug use in patients with intermittent claudication was investigated (n = 78); only patients with ischemia were included in the study (n = 23); patients with both intermittent claudication and ischemia were included in the study, without differentiation according to disease severity in the reported results (n = 2); or the follow-up period was shorter than 3 months (n = 54). Of the remaining 32 studies, 22 were excluded because patient characteristics were missing (n = 5), no data on both functional capacity and quality of life were reported (n = 16), or dichotomous outcomes were used for walking distances (n = 1). Of the 10 eligible articles, two (25,26) reported data from the same study population and were therefore considered as one study. All publications from studies that met our selection criteria reported either SF-36 or SF-20 survey data, with two exceptions: In one article (32), data from the Nottingham Health Profile questionnaire were reported, and in another article (33), data from a disease-specific questionnaire, PAVK-86, were reported instead. To improve the comparison of quality-of-life scores between the studies, we excluded these two studies from our systematic review. Thus, seven studies (12,25, 26,34–38) with a total of 202 patients who underwent exercise training therapy and 470 patients who underwent PTA (Tables 1 and 2) were included in our review.

### Table 1

Characteristics of Studies of Exercise Training or PTA for Intermittent Claudication

<table>
<thead>
<tr>
<th>Study and Reference</th>
<th>Year of Publication</th>
<th>Study Location</th>
<th>No. of Institutions</th>
<th>Study Period*</th>
<th>Treatment</th>
<th>Type of Study</th>
<th>Quality-of-Life Data Collection†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bosch et al (12)</td>
<td>1999</td>
<td>The Netherlands</td>
<td>6</td>
<td>1993–1996</td>
<td>PTA</td>
<td>Prospective cohort and randomized controlled trial</td>
<td>SF-36, EuroQol-5D, HUI, rating scale, time trade-off, standard gamble</td>
</tr>
<tr>
<td>Gardner et al (35)</td>
<td>2001</td>
<td>United States</td>
<td>1</td>
<td>ND</td>
<td>Exercise</td>
<td>Prospective cohort and randomized controlled trial</td>
<td>SF-36</td>
</tr>
<tr>
<td>Patterson et al (36)</td>
<td>1997</td>
<td>United States</td>
<td>1</td>
<td>ND</td>
<td>Exercise</td>
<td>Prospective cohort</td>
<td>SF-36, WIQ</td>
</tr>
<tr>
<td>Regenstein et al (37)</td>
<td>1997</td>
<td>United States</td>
<td>1</td>
<td>ND</td>
<td>Exercise</td>
<td>Prospective cohort and randomized controlled trial</td>
<td>SF-36</td>
</tr>
<tr>
<td>Savage et al (38)</td>
<td>1997</td>
<td>United States</td>
<td>1</td>
<td>ND</td>
<td>Exercise</td>
<td>Prospective cohort</td>
<td>SF-36</td>
</tr>
</tbody>
</table>

* ND = no data.
† For a description of time trade-off and standard gamble (preference measures), see reference 44. EuroQol-5D = five-dimensional quality-of-life survey developed by EuroQol Group, HUI = Health Utilities Index, SF-20 = Medical Outcomes Study 20-Item Short Form, WIQ = Walking Impairment Questionnaire.
‡ Comparison between primary stent placement and primary angioplasty with subsequent selective stent placement.
§ Comparison between exercise rehabilitation group and nonexercise (control) group.
# Comparison between supervised hospital-based exercise program and unsupervised home-based exercise program.

---

To pool data for discrete variables such as male sex, cardiac disease, hypertension, diabetes mellitus, pulmonary disease, hyperlipidemia, stroke, history of smoking, and current smoking, the random effects model described by Laird and Mosteller (30) was used. The random effects model takes into account the variance between studies, as well as that within studies. Because of the lack of reported data for variance around the means of continuous variables such as age, ABI, walking distance, and quality-of-life score, we were unable to use the random effects model to pool the continuous variables. Therefore, we calculated the weighted mean and 95% confidence interval (CI) for these variables.

Within the treatment study groups of exercise and PTA, the dichotomous baseline characteristics of all patients in the included studies were tested for homogeneity by using the χ² test (statistical significance level, α = .05). Differences in pooled patient characteristics between the studies of PTA and those of exercise rehabilitation group and nonexercise (control) group.
Four of the seven studies were randomized controlled trials (Table 1). However, in none of these randomized controlled trials was exercise training compared with PTA directly: In one study, primary stent placement was compared with primary PTA, and in three randomized controlled trials, different types of exercise training were compared. The remaining three studies were prospective cohort studies; in one of these, exercise training and PTA were compared. Follow-up periods varied from 3 to 24 months; in only one study was the mean follow-up period (14.7 months) reported (12).

Many baseline demographic and patient characteristics were not reported (Table 3). The percentage of male patients was 69% in the exercise training group and 73% in the PTA group (P = .56). The mean age of patients in the exercise training group also was not significantly different from that in the PTA group (68 vs 63 years, respectively; P = .33).

Across the studies in each treatment group, homogeneity was found in the proportion of the study population with male sex, coronary artery disease, hypertension, pulmonary disease, hyperlipidemia, stroke, and smoking behavior. Among the studies in the exercise training group, heterogeneity was found in regard to the percentage of patients with diabetes mellitus and the percentage of patients who were current smokers. Among the studies in the PTA group, no heterogeneity was found for these comorbidities.

Between the two treatment groups, patient characteristics were similar, except in regard to the percentages of patients with hypertension and stroke, which were significantly higher in the exercise training group than in the PTA group (64% vs 31%, P = .01, for hypertension; 23% vs 11%, P = .03, for stroke).

The weighted mean ABI at baseline in the exercise training group was lower than that in the PTA group (0.64 vs 0.71), but the difference was not statistically significant (P = .13). Also, no statistically significant difference was found in the weighted mean for maximum walking distance at baseline between the exercise training group and the PTA group (289 m vs 155 m, respectively; P = .11).

To detect publication bias, we constructed two funnel plots: one for the PTA group, and one for the exercise training group. Neither plot appeared symmetric. There were too few studies, however, to permit a proper evaluation of publication bias.

Exercise Training Programs

The exercise training programs differed across the studies (Table 4). These differences included variations in supervision, intensity, frequency, duration, and type of exercise. Of the five studies included in the exercise training group, one study involved supervised training, one study involved unsupervised training, and three studies each included one group with supervised training and one group with unsupervised training. Supervision implied that patients participated in a structured hospital-based exercise program.

In all programs, the frequency of supervised exercise training was at least three times per week, and walking was the mode of exercise used. In addition to walking, one study used a combination of walking, standing on tiptoe, and cycling (33,36). The duration of the supervised exercise training session varied from 15 to 60 minutes. The duration of the exercise training programs varied from 12 to 24 weeks. The mean percentage of compliance among patients in the supervised programs varied from 73% to 100%. (In one study, compliance was not reported.)

In the four unsupervised programs, the type and intensity of contact with the patient varied from weekly or monthly conversations by telephone to weekly lectures and instructions or no contact during the exercise program. In one study, patients were asked to maintain a weekly log. The mean compliance of the patients in the unsupervised programs

---

### Table 2: Criteria for Inclusion or Exclusion of Patients in Each Study

<table>
<thead>
<tr>
<th>Study and Reference</th>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bosch et al (12)</td>
<td>Intermittent claudication or critical ischemia caused by stenosis or occlusion in the iliac arteries</td>
<td>Stenosis of more than 10 cm in length; arterial occlusion of more than 5 cm, or of 5 cm or less in length and not allowing the passage of a guidewire; stenosis involving the distal aorta; severe comorbidity (eg, severe cardiac or cerebrovascular abnormality, malignant disease); and nonmedical factors such as inability to understand Dutch, or expected poor compliance</td>
</tr>
<tr>
<td>Chetter et al (25,26)</td>
<td>Treatment with PTA for intermittent claudication</td>
<td>Critical ischemia as defined by the European Consensus document</td>
</tr>
<tr>
<td>Curne et al (34)</td>
<td>Intermittent claudication and lower limb arterial disease potentially suitable for vascular intervention</td>
<td>Not reported</td>
</tr>
<tr>
<td>Gardner et al (35)</td>
<td>Positive response on Rose et al (45) questionnaire, age &gt; 60 years, ABI &lt; 0.97</td>
<td>Fontaine stage I or III disease; exercise tolerance limited by factors other than claudication; poorly controlled hypertension; pulmonary disease; hemiparetic gait; severe arthritis; poorly controlled diabetes mellitus; active major medical problem, including cancer, renal or liver disease, anemia, substance abuse, and dementia</td>
</tr>
<tr>
<td>Patterson et al (36)</td>
<td>Symptoms &gt; 3 months, ABI &lt; 0.9 at rest, or systolic blood pressure decrease of 15 mm Hg or more after exercise test</td>
<td>Ischemic rest pain or tissue loss, inability to participate in an exercise program because of limitations of comorbid illness, failed cardiac screening</td>
</tr>
<tr>
<td>Regensteiner et al (37)</td>
<td>Stable intermittent claudication &gt; 3 months; ABI &lt; 0.94 at rest, decreased to &lt; 0.73 after exercise test</td>
<td>Pain at rest, no ischemic ulceration or gangrene, inability to walk on the treadmill at a speed &lt; 2 mph, symptoms of angina, congestive heart failure, chronic obstructive pulmonary disease, arthritis, diabetes, previous vascular surgery or angioplasty within previous year</td>
</tr>
<tr>
<td>Savage et al (38)</td>
<td>Age &gt; 50 years; Rutherford stage I, II, or III disease; unilateral intermittent claudication</td>
<td>Unstable cardiopulmonary disease; severe extremity arthritis; tobacco use; weight &gt; 40 kg above ideal; renal insufficiency; use of beta-blocker, pentoxifylline, or cilostazol within 8 weeks of entry; functioning lower extremity bypass; severe cognitive impairment</td>
</tr>
</tbody>
</table>
## TABLE 3
Baseline Characteristics of Patients Undergoing Exercise Training or PTA

<table>
<thead>
<tr>
<th>Study, Reference, and Patient Group</th>
<th>Exercise Training Studies</th>
<th>PTA Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male Sex (%)</td>
<td>Age (y)*</td>
</tr>
<tr>
<td>Currie et al (34), group 1</td>
<td>78</td>
<td>76</td>
</tr>
<tr>
<td>Gardner et al (35)</td>
<td>28</td>
<td>89</td>
</tr>
<tr>
<td>Patterson et al (36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>46</td>
</tr>
<tr>
<td>Regensteiner et al (37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>ND</td>
</tr>
<tr>
<td>Bosch et al (12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>143</td>
<td>71</td>
</tr>
<tr>
<td>Group 2</td>
<td>136</td>
<td>73</td>
</tr>
<tr>
<td>Chetter et al (25)</td>
<td>117</td>
<td>78</td>
</tr>
<tr>
<td>Currie et al (34), group 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>74</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>470</td>
<td>73.3</td>
</tr>
<tr>
<td>All of the above</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean†</td>
<td>NA</td>
<td>73.3</td>
</tr>
<tr>
<td>95% CI</td>
<td>NA</td>
<td>73.3</td>
</tr>
<tr>
<td>Upper limit</td>
<td>76.1</td>
<td>73.7</td>
</tr>
</tbody>
</table>

Note.—Data are in patients unless otherwise specified. NA = not applicable, ND = no data.

* Data are mean unless otherwise indicated.
† Data are numbers of lesions.
‡ Median is given.
† Current smoking or diabetes mellitus was an exclusion criterion.
§ Based on random effects model, except where otherwise indicated.
§ Weighted mean is given.
** A statistically significant difference (P < .05) was found between pooled mean for exercise training study group and pooled mean for PTA study group.
†† Values obtained from Tetteroo et al (46).
838

䡠

Radiology

䡠

June 2005

ND
24†
Savage et al (38), group 1

Walking
Yes

Regensteiner et al (37), group 1

Note.—ND ⫽ no data.
* Mean compliance for both groups (1 and 2) is given.
† After 12 weeks, patients in the hospital exercise program were transferred to the home exercise program.

3

15–40

100
12

Yes, with weekly
lectures
Yes

Gardner et al (35)

Patterson et al (36), group 1

Arm and leg ergometry,
stationary cycling
Walking

2 mph, 0% grade for 3 min,
increased to 3.5% grade every
3 min to maximal claudication
pain
2 mph, 0% grade with 2%
increase every 2 min until
maximal claudication pain

3

35–50
(including rest
periods)

78*
12
60
3

73
24
ND
3

2 mph, 0% grade with 2%
increase every 2 min until
maximal claudication pain
1–2 mph, 2%–3% grade

3
No treadmill used
Walking to point of intense
pain

No; contact by
telephone once a
month
Yes
Savage et al (38), group 2

Regensteiner et al (37), group 2

Treadmill walking to near
claudication pain

ND
24

100
12
No treadmill used
Walking

3

35–50
(including rest
periods)
15–40

78*
12
20–40
3
No treadmill used

92*
12
ND
7
No treadmill used

No; weekly lectures,
instruction, and logs
No; weekly contact by
telephone

Mode of Exercise

Different exercises while lying,
sitting, and standing, and
daily walking
Walking

Length of Exercise
Training Program
(wk)
Duration of
Exercise Session
(min)
Frequency
of Exercise
(d/wk)
Treadmill Exercise Speed
and Grade
Supervision

Patterson et al (36), group 2

In the exercise training group, the
mean scores for the dimensions of physical functioning, physical role functioning, bodily pain, and general health demonstrated improvement at 3 and 6
months (Fig 2, Table 6). However, the
improvement was statistically significant
only for the dimensions of physical functioning and bodily pain at 3 months and

No

Quality of Life

Study, Reference, and Patient
Group

All studies reported the ABI during rest
at baseline (Table 5). The maximum
walking distance was determined in two
of the three studies in the PTA group and
in four of the five studies in the exercise
training group. It should be noted that
the test duration and the speed and grade
settings used during the treadmill test to
determine the maximum walking distance varied between studies. Grades of
0%–10% were used with speeds of 1–2
mph (1.6 –3.2 km/h), either fixed or increased in increments of 1 or 2 mph.
In the exercise training group, 1- or
2-year follow-up data were not available
(Table 5). Improvement was shown in
the mean ABI at rest between baseline
and 3- or 6-month follow-up, but the difference was not statistically significant
(P ⫽ .29 and P ⫽ .73, respectively) (Fig 1).
At 3 months, there was no statistically
significant improvement in the mean
maximum walking distance (P ⫽ .09), but
a statistically significant improvement in
the mean maximum walking distance
was demonstrated between baseline and
6-month follow-up (P ⫽ .02).
In the PTA group, a significant improvement in the mean ABI at rest was
demonstrated at 3 and 6 months of follow-up (P ⬍ .01 for both) (Fig 1). In addition, a statistically significant improvement in the mean maximum walking
distance was demonstrated at 3 months
(P ⫽ .01) and 6 months (P ⬍ .01).
The difference between the two treatment groups in ABI at 3 and 6 months
was significant (P ⬍ .01 and P ⫽ .02,
respectively). Although the weighted
means of the maximum walking distance
at baseline were not significantly different between the treatment groups, we
did not compare this outcome of functional capacity during follow-up, because
of the variation in measurements of the
maximum walking distance among the
studies.

TABLE 4
Components of Exercise Training Programs for Patients with Intermittent Claudication in the Included Studies

Radiology

Functional Capacity

Currie et al (34), group 2

Mean
Compliance
(%)

varied from 78% to 100% (in one study,
compliance was not reported).

Spronk et al


for the dimension of bodily pain at 6 months. The effect of treatment was highest for the dimension of physical functioning at 3 months ($P < .01$).

In the PTA group, mean scores for the dimensions of physical functioning, physical role functioning, bodily pain, and general health showed improvement at 3 and 6 months. Improvements were statistically significant for the dimensions of physical role functioning and general health at 3 months (Fig 2, Table 6). The effect of treatment was greatest for the dimension of physical role functioning after 3 months ($P < .01$).

There were no significant differences at baseline between the two treatment groups with regard to the dimensions of physical functioning, physical role functioning, bodily pain, and general health. In addition, no significant differences in the mean change of quality of life were demonstrated between the two treatment groups during follow-up.

**DISCUSSION**

In the current systematic review, functional capacity and quality-of-life outcomes after exercise training or PTA were reviewed and compared. Results demonstrated that at 3 and 6 months of follow-up the ABI was significantly improved in the PTA group but not in the exercise group. Quality of life was significantly improved in the dimensions of physical functioning and bodily pain in the exercise group and in the dimensions of physical role functioning and general health in the PTA group at 3 months. ABI significantly differed between the two treatment groups at 3 and 6 months, whereas the mean change in the quality-of-life scores did not significantly differ between the groups during follow-up.

To our knowledge, no review has been published in which exercise training and PTA are compared with regard to both functional capacity and quality of life. In addition, results have been published from only two randomized controlled trials in which exercise training and PTA were directly compared (32,39). These studies, however, did not meet our inclusion criteria; one study did not report both functional capacity and quality of life, and the other study did not use the SF-36 survey for quality-of-life assessment. The results of one of these studies...
showed that, at 12 months of follow-up, patients with mild or moderate intermittent claudication who underwent supervised exercise training had greater symptomatic improvement than did patients who underwent PTA (39). Results of the other study showed no significant difference in quality of life between control subjects (patients after unsupervised exercise training) and the PTA group after 2 years of follow-up (32,40). The PTA group, however, included substantially fewer patients with arterial occlusion and/or severe stenosis than did the exercise group. Data from these two randomized controlled trials were collected for a review, but the trials were relatively small, and the results must be interpreted with caution (41).

Unfortunately, one of the limitations in our systematic review was that data for both functional capacity and quality of life were inconsistently reported. In particular, data about long-term functional capacity and quality-of-life outcomes for exercise training after 6 months were lacking, and, in consequence, the deleterious effect of postangioplasty restenoses must have been somewhat underestimated. Furthermore, many baseline demographic and patient characteristics that may have been relevant to a comparison between treatment groups, such as the disease location (aortoiliac vs femoropopliteal arteries), smoking history, and the use of medications (eg, cilostazol, statins), were not reported.

Another limitation, which became apparent after review of the exercise training programs, is the lack of standardization in the different components of exercise training. This limitation hampered the comparison of the reported results of the different studies. Investigators in the different exercise training studies evaluated the effectiveness of supervised or unsupervised exercise programs, but the relative effectiveness of these programs is still controversial. Some studies, particularly those performed in Europe, have shown good results with a home-based exercise program (42,43). Other studies, however, have shown little benefit for this type of program (37), whereas supervised exercise has provided consistent clinical improvement in almost every study (8,36–38). There is also a lack of standardization of the treadmill test with regard to the use of fixed loading or graded loading, different speeds, and different maximum walking times. A longer walking distance was used in the exercise studies than in the PTA studies. These different protocols could help to explain the different maximum walking distances reported at baseline and during follow-up between the two treatments. The existence of this bias hampered comparison between the two treatments.

Another limitation of our study was the potential publication bias. However, studies in the exercise training group were relatively small, and there were only four studies in the PTA group, which limits its power for detection of publication bias. Another limitation of the current systematic review is that the included studies were not randomized controlled trials in which exercise training was compared with PTA. This implies that the population undergoing exercise training may differ from the population undergoing PTA. Nevertheless, in our study, most patient characteristics were not significantly different between the treatment groups. Even though randomized controlled clinical trials are supposed to be the superior model for evaluating and comparing different treatment modalities, they too have limitations (44). In general, a randomized trial is performed in an ideal setting, with a selected, usually small, patient population, which may hamper generalization of the results.

Systematic reviews of cohort studies, on the other hand, may reflect general clinical practice and provide additional data about a larger number of patients. There is still no evidence that quality of life for patients with intermittent claudication is significantly better after PTA than after exercise training after more than 6 months. Patients with intermittent claudication are often treated conservatively; however, various exercise training programs are used (34–38). In addition, the definitions of supervised exercise and unsupervised exercise vary widely in European countries. Therefore, a standardized exercise training program with measurement of the maximum walking distance according to a universally accepted standard would be desirable. Over the past decade, PTA has been more widely used in patients with intermittent claudication, mainly because of its low associated morbidity and mortality (11–13). PTA, however, is a local form of treatment, and it is therefore not surprising that patients may experience relapse after PTA, with ipsilateral or contralateral symptoms representing progressive disease at other sites; the benefit with PTA may be a short-term effect. A randomized controlled trial of both treatments with long-term follow-up will
enable further evaluation of effectiveness and determination of functional capacity and quality of life after more than 6 months.

In conclusion, our results demonstrate an improvement in quality of life after exercise training, whereas both functional capacity and quality of life showed significant improvement after PTA. ABIs in the PTA group were significantly higher than those in the exercise training group at 3 and 6 months, whereas the mean change in quality-of-life scores was not significantly different between the two groups during follow-up. Despite the limitations outlined, our analysis has shown the best evidence that we have. This is what we know at this point in time, and more needs to be done.

References

### TABLE 6

Quality of Life in Exercise Training and PTA Study Groups at Baseline and during Follow-up

<table>
<thead>
<tr>
<th>Study, Reference, and Patient Group</th>
<th>Baseline</th>
<th>3 Months</th>
<th>6 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PF</td>
<td>PRF</td>
<td>BP</td>
</tr>
<tr>
<td>Currie et al (34), group 1</td>
<td>36</td>
<td>26</td>
<td>42</td>
</tr>
<tr>
<td>Gardner et al (35)</td>
<td>41</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Patterson et al (36)</td>
<td>43</td>
<td>45</td>
<td>53</td>
</tr>
<tr>
<td>Regensteiner et al (37)</td>
<td>52</td>
<td>65</td>
<td>ND</td>
</tr>
<tr>
<td>Group 1</td>
<td>52</td>
<td>65</td>
<td>ND</td>
</tr>
<tr>
<td>Group 2</td>
<td>61</td>
<td>65</td>
<td>ND</td>
</tr>
<tr>
<td>Savage et al (38)</td>
<td>54</td>
<td>84</td>
<td>59</td>
</tr>
<tr>
<td>Group 1</td>
<td>54</td>
<td>84</td>
<td>59</td>
</tr>
<tr>
<td>Group 2</td>
<td>45</td>
<td>47</td>
<td>50</td>
</tr>
<tr>
<td>All of the above</td>
<td>42</td>
<td>40</td>
<td>47</td>
</tr>
<tr>
<td>Weighted mean 95% CI</td>
<td>29</td>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td>Lower limit</td>
<td>55</td>
<td>74</td>
<td>58</td>
</tr>
</tbody>
</table>

Notes.—Unless otherwise specified, data are mean values from ratings on a scale of 0–100, with 0 indicating worst health and 100 indicating perfect health. BP = bodily pain, GH = general health, ND = no data, PF = physical functioning, PRF = physical role functioning.
* Values are medians obtained from Chetter et al (26) for the same study population.
† Values were obtained from Klein et al (47) for patients in the Dutch Iliac Stent Trial.
‡ Values were obtained at 12-month follow-up.


Schroeder TV. The TASC supplement: international recommendations for management of peripheral arterial disease [editorial]. Eur J Vasc Endovasc Surg 2000; 19:563.


McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form health survey (SF-36). II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993; 31:247–263.


Cytotoxic Effects of Ionic High-osmolar, Nonionic Monomeric, and Nonionic Iso-osmolar Dimeric Iodinated Contrast Media on Renal Tubular Cells in Vitro

PURPOSE: To compare the cytotoxic effects of dimeric and monomeric iodinated contrast media on renal tubular cells in vitro with regard to osmolality.

MATERIALS AND METHODS: LLC-PK1 cells were incubated with ioxithalamate, ioversol, iomeprol-300, iomeprol-150, iodixanol, iotrolan, and hyperosmolar mannitol solutions for 1–24 hours at concentrations from 18.75 to 150 mg of iodine per milliliter. Cytotoxic effects were assessed with 3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) assay. Data were analyzed with one-way analysis of variance; post hoc tests were performed.

RESULTS: At equal iodine concentrations, ioxithalamate showed stronger cytotoxic effects than did other contrast media (MTT conversion for ioxithalamate was 4% vs that for ioversol of 32%, that for iomeprol-300 of 34%, that for iodixanol of 40%, and that for iotrolan of 41% of undamaged control cells at 75 mg of iodine per milliliter, \( n = 61–90, P < .001 \)); there was no significant difference between low-osmolar monomeric and iso-osmolar dimeric contrast media (\( P < .05 \)). At equal molarity, dimeric contrast media induced significantly stronger cytotoxic effects than did low-osmolar monomeric contrast media (40% for iodixanol and 41% for iotrolan vs 64% for ioversol and 59% for iomeprol-300 at 98.5 mmol/L, \( n = 61–75, P < .001 \)). At equimolar concentrations, both dimeric contrast media showed stronger cytotoxic effects than did iso-osmolar formulation of iomeprol-150 (51% for iodixanol and 50% for iotrolan vs 77% for iomeprol-150 at 98.5 mmol/L, \( n = 35–40, P < .001 \)). Mannitol solutions induced weaker cytotoxic effects than did corresponding contrast media compounds (74% for mannitol-520 vs 34% for iomeprol-300 and 41% for mannitol-1860 vs 4% for ioxithalamate, \( P < .001 \)).

CONCLUSION: Besides hyperosmolality, direct cytotoxic effects of contrast media molecules contribute to their cytotoxic effects. Results of this study indicate that dimeric contrast media molecules have a greater potential for cytotoxic effects on proximal renal tubular cells in vitro than do monomeric contrast media molecules.

© RSNA, 2005

Contrast medium–induced nephropathy remains one of the most important clinical problems after intravascular administration of iodinated contrast media. Although the incidence of contrast-induced nephropathy is considered low, patients with preexisting renal insufficiency, whether or not it is associated with diabetes mellitus, are at high risk for developing contrast-induced nephropathy (1–3). Contrast-induced nephropathy is associated with high in-hospital mortality and a poor long-term survival, especially if nephropathy that requires dialysis occurs (3–5). The pathogenesis of contrast-induced nephropathy is poorly understood. Hemodynamic changes of renal blood flow, which lead
to hypoxia of the renal medulla, and direct cytotoxic effects of contrast media on renal cells are thought to contribute to the pathogenesis of contrast-induced nephropathy (6,7).

The osmolality of contrast media is considered to play an important role in the pathogenesis of contrast-induced nephropathy. Low-osmolar nonionic contrast media have been shown to have fewer nephrotoxic effects than do high-osmolar ionic contrast media (1,2,8). However, results of another clinical study showed no significant differences between the renal effects of ioxidanol and those of monomeric contrast media (9). This study has caused debate (10,11), however, and results of another clinical study were compared with those of high-osmolar ionic compounds, with fewer nephrotoxic effects than are low-osmolar monomeric contrast media (12).

Researchers in early in vitro studies found evidence for direct renal tubular cell toxic effects of contrast media (12,13). In vitro experiments are a way to examine the cytotoxic effects of contrast media on renal cells because of the absence of confounding variables, which can be found in vivo (eg, hypoxia due to hemodynamic changes or other systemic mechanisms) (14). The finding of fewer in vivo nephrotoxic effects of low-osmolar contrast media, as compared with those of high-osmolar ionic compounds, correlates with their fewer in vitro cytotoxic effects (14). Researchers in previous studies have compared the effects of different classes of contrast media on renal tubular cells only at equal iodine concentrations. To compare the potential for nephrotoxic effects of monomeric and dimeric contrast media molecules, however, it is necessary to compare the contrast media on a molar basis and to control the effects of osmolality. Thus, the purpose of our study was to compare the cytotoxic effects of monomeric and dimeric contrast media at equal iodine concentrations, as well as at equal molarity, on renal tubular cells in vitro, with special regard to osmolality.

**MATERIALS AND METHODS**

**Chemicals**

All chemicals were purchased from a commercial manufacturer (Sigma Chemical, Munich, Germany) unless otherwise noted.

**Cell Cultures**

LLC-PK1 cells that are from a proximal tubular epithelial cell line of porcine origin were obtained from a cell culture collection (American Type Culture Collection, Rockville, Md) and grown in modified medium 199, supplemented with 10% fetal bovine serum (FBS; Invitrogen, Paisley, Scotland), 100 U/mL penicillin, and 100 µg/mL streptomycin (Seromed, Vienna, Austria), at 37°C with a humidified atmosphere of 95% air and 5% CO₂. Cells were grown to confluence in flasks (75 cm²) over 6–7 days, treated with trypsin, and seeded into 96 microtiter plates (Greiner, Frickenhausen, Germany). Only confluent cell monolayers with dome formation were studied.

**Experimental Solutions**

Cells were incubated with either control media (serum-free and phenol red-free medium 199) or various doses of contrast media diluted in serum-free medium 199. Ready-to-use formulations of all contrast media we used were tested at concentrations of 49.25 mmol/L (low) and 98.5 mmol/L (high), respectively, and are enclosed within parentheses, for the following contrast media: ioxidanol (n = 69 and n = 75), iotrolan (n = 66 and n = 61), ioversol (n = 70 and n = 77), iomeprol-300 (n = 71 and n = 61), and ioxithalamate (n = 77 and n = 90). As a control, cells were incubated with corresponding volumes of a mannitol solution with an osmolality of 1860 mOsm/kg H₂O (n = 69 and n = 72) and with an osmolality of 520 mOsm/kg H₂O (n = 66 and n = 58) and with saline solution (n = 60 and n = 72). For the evaluation to which amount the osmolality contributes to cell injury, we performed separate experiments in which ioxidanol (n = 24 and n = 40), iotrolan (n = 24 and n = 40), and iomeprol-300 (n = 17 and n = 35) were compared with the iso-osmolar formulation iomeprol-150 (n = 18 and n = 35) at concentrations of 49.25 mmol/L (low) and 98.5 mmol/L (high) for 24 hours in the same experiments.
**Volume 235  Number 3  Radiology**

Cytotoxic Effects of Iodinated Contrast Media

**Determination of Cell Injury**

Cell viability was assessed by using the 3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) uptake assay. The conversion of MTT, a tetrazolium salt, into formazan depends on the activity of a group of mitochondrial dehydrogenases and, thus, is an indicator of cell metabolic activity. The MTT assay was performed according to Mosmann (15), with modifications. An equal volume of MTT reagent (5 mg/mL phosphate-buffered saline) was added to the cell supernatant and incubated 4 hours at 37°C. Supernatants were removed, and 100 µL lysis buffer (20% sodium dodecyl sulfate, 33.3% dimethylformamide, 2% acetic acid, pH 4.7) was added per well; plates were shaken at room temperature for 30 minutes, and absorbance was measured at 570 nm, with 630 nm as reference. The measurements were performed automatically by a microplate reader (Dynatech MR 5000; Dynex Technologies, Chantilly, Va). The incubation with the experimental solutions, MTT, and lysis buffer was performed by one author (M.C.H.).

**Statistical Analysis**

The data are reported as percentage of undamaged control cells and are presented as the mean ± standard error of the mean. To compare the effects of the contrast media and to compare the contrast media with the mannitol solutions and the NaCl solution, data were analyzed with one-way analysis of variance, followed by analysis with the Tukey post hoc test for multiple comparisons. MTT reduction after incubation with iomeprol-300 at various doses and various incubation times was compared with the reduction in untreated control cells by means of one-way analysis of variance, followed by analysis with the Dunnett post hoc test for multiple comparisons. A difference with a P value < .05 was considered significant. For statistical analysis and graphic representations, statistical software (Prism 3.03, 2002; GraphPad Software, San Diego, Calif) was used.

**RESULTS**

**Concentration- and Time-dependent Injury in LLC-PK1 Cells**

To assess concentration and time dependency of contrast-mediated cell injury on LLC-PK1 cells by measuring mitochondrial activity with MTT assay, we used the low-osmolar nonionic monomer iomeprol-300. The concentrations used ranged from 18.75 to 150 mg of iodine per milliliter, and the incubation times varied between 1 and 24 hours. As shown in Figure 1, iomeprol-300 induced a time- and concentration-dependent cell injury, as assessed with the MTT assay. A significant effect was observed even at a concentration of 18.75 mg of iodine per milliliter after an incubation time of 24 hours (93% ± 6, P < .01 vs unexposed control cells, n = 12). At 1 hour after incubation with iomeprol-300, a significant effect was observed at a concentration of greater than 75 mg of iodine per milliliter (92% ± 2, P < .01 vs unexposed control cells, n = 12).

**Comparison of Contrast Media**

Ioxithalamate, iomeprol-300, ioversol, iodoxanol, and iotrolan showed a significantly dose-dependent inhibition of MTT conversion (Fig 2). Both dimeric contrast media showed a slightly weaker effect (40% ± 3 for iodoxanol and 41% ± 2 for iotrolan; n = 61–75) than did the two monomeric contrast media at a concentration of 75 mg of iodine per milliliter (32% ± 2 for ioversol and 34% ± 2 for iomeprol-300; n = 61–77). This difference was not statistically significant (P > .05). At equal iodine concentrations, the ionic compound ioxithalamate showed a significantly stronger effect than did all other tested contrast media (P < .001). After incubation with ioxithalamate, the MTT conversion was 4% ± 0.3 of the undamaged control cells at a concentration of 75 mg of iodine per milliliter (n = 90).

When we compared the data on a molar basis, the results were different. Both dimeric contrast media showed a significantly stronger inhibition of MTT conversion, as compared with both monomeric contrast media (40% ± 3, 41% ± 2, 64% ± 2, and 59% ± 2 for iodoxanol, iotrolan, ioversol and iomeprol-300, respectively, at a concentration of 98.5 mmol/L; n = 61–75; P < .001 for the difference between the monomeric and the dimeric contrast media) (Fig 3). The effects of the dimeric contrast media were as strong as those of the ionic contrast medium, as observed with ioxithalamate (41% ± 2 at a concentration of 98.5 mmol/L; P > .05).

**Effect of Osmolarity and Dilution**

To exclude inhibition of MTT conversion caused by dilution of the tissue culture medium by the contrast media, the growth medium of the cells was diluted with 0.9% NaCl. A 1:8 dilution of the growth medium (corresponding to a contrast media concentration of 37.5 mg of iodine per milliliter) showed no significant effect on the MTT assay (98% ± 2; n = 60; P > .05), whereas a 1:4 dilution induced a slight but significant inhibition of MTT conversion (89% ± 2; n = 72; P < .001). This effect was significantly smaller, however, than that of all contrast media tested (P < .001) (Fig 4).

To evaluate the effect of hyperosmolal-
ity on the tubular cells, we tested mannitol control solutions with osmolarities that corresponded to those of ioxithalamate and iomeprol-300 (1860 mOsm/kg H₂O and 520 mOsm/kg H₂O, respectively). Hereafter, these mannitol solutions will be referred to as mannitol-1860 and mannitol-520. As shown in Figure 4, both mannitol solutions showed a significant dose-dependent inhibition of the MTT assay. This effect was markedly more pronounced with the mannitol-1860 solution when it was compared with that of the mannitol-520 solution (P < .001). The effects of the mannitol solutions were significantly weaker when they were compared with those of the corresponding contrast media (74% ± 3 and 34% ± 2 for mannitol-520 and iomeprol-300 and 41% ± 1 and 4% ± 0.3 for mannitol-1860 and ioxithalamate, respectively, at 25% vol/vol; n = 58–90; P < .001) (Fig 4).

At concentrations of 49.25 and 98.5 mM, the dimeric contrast media iodoxanol (69% ± 2 and 51% ± 4, n = 24–40) and iotrolan (71% ± 1 and 50% ± 2, n = 24–40) induced a more pronounced inhibition of MTT conversion, as compared with that of iomeprol-150 (81% ± 2 and 77% ± 2, n = 18–35, P < .001). The effects of the mannitol solutions were significantly weaker than those of the corresponding contrast media. The data were reported as percentage of undamaged control cells (% of control). Data show the mean ± standard error of the mean (n = 61–90). ∗ = P < .001 for ioxithalame versus ioversol and iomeprol-300. The data were reported as percentage of undamaged control cells (% of control). Data show the mean ± standard error of the mean (n = 54–90). ∗ = P < .001 for mannitol-520 versus iomprol-300 and mannitol-1860 versus ioxithalame, mg I/ml = milligrams of iodine per milliliter.

**DISCUSSION**

At equal iodine concentrations, our investigation showed no significant differences between the direct cytotoxic effects of nonionic monomeric and dimeric contrast media on renal proximal tubular cells in vitro, whereas the ionic contrast media demonstrated a markedly stronger effect when the effect was compared with that of the other contrast media. When we compared the data on a molar basis, however, the dimeric contrast media showed a significantly stronger effect on the tubular cells than did the nonionic monomeric contrast media. This suggests a greater cytotoxic effect of the dimeric contrast media molecules.

This finding, that at equal iodine concentrations the ionic high-osmolar monomer ioxithalamate induced a significantly stronger effect on the tubular cells than did the nonionic low-osmolar monomeric and iso-osmolar dimeric contrast media, is in agreement with results in previous in vitro studies (16–21). Findings in those previous studies indicated a greater nephrotoxic potential of ionic hyperosmolar contrast media when they were compared with nonionic contrast media. They greater in vitro toxic effects of the ionic hyperosmolar contrast media, as compared with those of the nonionic contrast media, correlate with their greater nephrotoxic effects in vivo (8,22). The hyperosmolality of ionic contrast media often is thought to play the pivotal role in the greater nephrotoxic effects of these compounds. To distinguish chemotoxic effects of the contrast media molecules from effects caused by hyperosmolality, we used as controls hyperosmolar mannitol solutions with the same osmolality as iomeprol-300 and ioxithalame.

Both mannitol solutions showed cytotoxic effects whereby the high-osmolar mannitol solution exerted more pronounced effects than did the low-osmolar solution. Since the mannitol solutions had significantly fewer toxic effects than did the corresponding contrast media, however, the cytotoxic effects of contrast media cannot solely be caused by their hyperosmolality. Furthermore, at a concentration of 98.5 mM/mL, iomeprol-300 showed a slightly stronger effect on the MTT assay than did the iso-osmolar compound iomeprol-150. At this concentration, the mannitol-520 solution induced only a slight inhibition of MTT conversion. Thus, only the small difference between iomeprol-150 and iomeprol-300 is probably caused by the hyperosmolality of iomeprol-300. Our results confirm findings in previous in vitro studies, which have shown that hyperosmolality of contrast media cannot fully explain their cytotoxic effects (18,20,21,23–25) or other contrast-induced side effects (26–29).

Hyperosmolality cannot be the only reason for contrast-induced cytotoxic effects, since in our study the iso-osmolar contrast media also showed cytotoxic effects. When the dimers iso-osmolar contrast media iodoxanol and iotrolan were compared with the monomeric low-osmolar contrast media iomeprol and ioversol at equal iodine concentrations, no significant difference in cytotoxic effects was observed. This finding is in agreement with results in previous studies of...
renal tubular cell cultures. No significant difference between iohexol and iodixano-

In our study, LLC-PK1 cells, which are from a proximal tubular epithelial cell line of porcine origin, were used to deter-
mine the toxic effects of contrast media on renal tubular cells. LLC-PK1 cells are a

The contrast media concentrations and incubation times used in our study were chosen carefully with regard to the

For investigation of contrast-induced cell injury, we used the MTT assay. MTT is a yellow tetrazolium salt that is con-

The MTT assay shows comparison of cytotoxic effects of the dimeric iso-osmolar contrast media iotrolan and iodixanol with those of low-osmolar monomer iomeprol-300 and those of the iso-osmolar formulation iomeprol-150 assessed with MTT assay at equimolar concentrations (in millimoles per liter [mmol/L]). LLC-PK1 cells were incubated with contrast media for 24 hours at concentrations ranging from 49.25 to 197 mmol/L. Dimeric contrast media iotrolan and iodixa-
nol showed a significantly stronger inhibition of MTT conversion than did monomeric contrast media iomeprol-300 and iomeprol-150. Iomeprol-150 induced a slightly weaker effect than did iomeprol-300. The data were reported as percentage of undamaged control cells (% of control). Data show the mean ± standard error of the mean (n = 17–40). * P < .001 for iodixanol versus iomeprol-300, iodixanol versus iomeprol-150, iotrolan versus iomeprol-300, and iotrolan versus iomeprol-150; + P < .05 for iomeprol-150 versus iomeprol-300.

To compare the effects of the monomeric and dimeric contrast media molecules, the different osmolarities of compounds must be taken into account when compounds with the same iodine concentrations are used. For this reason, we performed experiments with the iso-osmolar compound iomeprol-150, which contains 150 mg of iodine per milliliter. Identical volumes of iomeprol-150 and of the dimeric contrast media contain the same number of contrast media molecules and have almost the same osmolarity. Also, with these conditions, the dimeric contrast media showed a signifi-
cantly more pronounced cytotoxic effect in comparison with that of iomeprol. Therefore, the difference in toxic effects in tubular cells must be attributed to the contrast medium molecule itself or to additives in the contrast medium, because differences in osmolarity were excluded in our experiments. Investigators in previous studies (16,20), however, showed that the additives trometamol and edetate calcium disodium did not play a substantial role in the cytotoxic effects of contrast media.

To examine the toxic properties of contrast media in renal tubular cells in vitro, the dilution of the tissue culture media with the use of ready-to-use formulations must be considered. At a concentration of 75 mg of iodine per milliliter, 25% of the culture medium was replaced by contrast media. This replacement could be responsible for toxic effects. The dilution of the growth media by the same volume of isotonic NaCl solution, however, showed a reduction of MTT levels only at 58% ± 1.66. In addition, researchers in earlier studies have shown that the dilution of the growth medium does not explain the contrast-induced cytotoxic effects in cell culture experiments at the concentrations used (23,24).

In our study, LLC-PK1 cells, which are from a proximal tubular epithelial cell line of porcine origin, were used to deter-
mine the toxic effects of contrast media on renal tubular cells. LLC-PK1 cells are a

The contrast media concentrations and incubation times used in our study were chosen carefully with regard to the clinical situation. Contrast media doses administered clinically (even during angiography of the abdominal aorta or the renal arteries) result in plasma concentrations of approximately 10 mg of iodine per milliliter and usually do not exceed 20 mg of iodine per milliliter (14,20). Because contrast media are concentrated in the renal proximal tubules as a result of reabsorption of water, the concentration of contrast media in the renal tubules is much higher than it is in plasma (32). After intravascular administration of contrast media in rabbits, urinary concentra-
tions higher than 100 mg of iodine per milliliter have been measured (33). We compared different contrast media at concentrations of 37.5 and 75 mg of iodine per milliliter. These concentrations are in the range of those that have been used in previous cell culture studies (16,18–21,23,24). In some studies, concentra-
tions of as much as 100–150 mg of iodine per milliliter were used (20,21,23).

In healthy volunteers, contrast media were rapidly excreted almost exclusively by means of renal glomerular filtration (34). The elimination half-life increases progressively, however, with increasing renal impairment (35). In patients with chronic renal failure, the elimination half-life of iopamidol was about 70 hours (36). Furthermore, in patients with contrast-induced nephropathy, a persistent nephrogram after more than 24 hours is a common finding (37). We demonstrated time-dependent cell injury induced by contrast media at incubation times be-
tween 1 and 24 hours. The contrast media were compared after an incubation time of 24 hours, which also has been used in several studies on renal tubular cells before (20,38,39). In one study, an incubation time of as much as 3 days was used (25). Therefore, with regard to the high urinary contrast medium concentra-
tion and the prolonged exposure of the tubular cells to the contrast medium in patients with renal insufficiency, the concentra-
tions and incubation times used in our study were in the range of those that can occur in clinical situations.

For investigation of contrast-induced cell injury, we used the MTT assay. MTT is a yellow tetrazolium salt that is con-
verted into a blue formazan product by mitochondrial dehydrogenases when it is incubated with living cells. With the ass-
say, detection of very small numbers of living cells is allowed, and the amount of formazan generated is directly propor-
tional to the number of living cells. Only living cells with active mitochondria cleave MTT, in which the amount of formazan generated depends on the level of energy metabolism in the cell (15). Therefore, this assay is an indicator for cell viability, proliferation, and metabolic activity. The MTT assay is a well-established method for assessment of cytotoxic effects of various substances in cell cultures (30,40) and also has been used before to investigate cytotoxic effects of contrast media in some studies (17,20,25). The molecular mechanisms by which contrast media induce direct cytotoxic effects on renal cells are poorly understood. Findings in some studies have suggested that oxidative stress is involved in contrast-induced renal damage (41–44). Researchers in another recent in vitro study, however, have shown toxic effects of contrast media to be dissociated from tubular cell oxygen stress (25). There are also conflicting results with regard to contrast-induced apoptosis in renal tubular cells. Researchers in some studies demonstrated apoptosis in renal tubular cell cultures (21,24,39), whereas investigators in another study found no evidence of contrast-induced apoptosis (20).

Whether dimeric contrast media have fewer nephotoxic effects than do monomeric contrast media in vivo is a matter of debate. Findings in a clinical study suggested that contrast-induced nephropathy is less likely to develop in high-risk patients when iodixanol is used rather than when low-osmolar nonionic contrast media are used (9). Results in that study are arguably controversial for the following reasons: the small number of patients who developed acute renal failure; the significant differences in proteinuria, in the duration of diabetes, and in the use of angiotensin-converting enzyme inhibitors in the iohexol group compared with the iodixanol group; and the lack of standardized vigorous hydration of the patients (10). Moreover, in another clinical study, Baker et al (45) showed a markedly higher incidence of contrast-induced nephropathy after administration of iodixanol, as compared with the findings in the study of Aspelin et al (9). Furthermore, in a recent clinical trial, no significant differences in the renal effects of iopamidol 370 and iodixanol 320 were observed (11). Katayama et al (46) reported a similar renal tolerability of iodixanol and iomeprol after intrarterial injection. Further studies are required to investigate the effect of dimeric contrast media in comparison with that of monomeric contrast media on renal function in vivo.

Some limitations of our study have to be considered. As in all in vitro experiments, in our study an artificial model system was used to compare the cytotoxic effects of the contrast media. This design does not consider the complex, multifactorial pathogenesis of contrast-induced nephropathy in vivo. In vitro experiments with cultured renal cells, however, are an established method to examine the cytotoxic effects of drugs, especially contrast media, without confounding variables, which can be found in vivo, such as hypoxia caused by hemodynamic changes or other systemic mechanisms (14). Furthermore, the MTT assay used in our study is an indicator for cell viability, proliferation, and metabolic activity. The inhibition of MTT conversion by contrast media observed in our study could be explained not only by a reduced number of viable cells but also by reduced mitochondrial activity. Researchers in two studies have shown indications for contrast-induced mitochondrial dysfunction. Hardiek et al (20) demonstrated an increase in mitochondrial membrane potential by contrast media, and they suggested that a consecutive decrease in dehydrogenase activity could potentially explain the inhibition of MTT conversion by contrast media. Investigators in another study (25) found evidence for contrast-induced mitochondrial damage because of cytochrome c loss into the cell supernatant solution. They suggested that there may be a link between the observed contrast-induced plasma membrane damage and mitochondrial injury (25). The mechanisms by which contrast media exert their cytotoxic effects, however, remain uncertain.

In conclusion, in our study, we demonstrated that, although hyperosmolality contributes to the cytotoxic effects of hyperosmolar contrast media in renal proximal tubular cells, hyperosmolality plays only a minor role in the case of low-osmolar contrast media. Direct cytotoxic effects of the contrast media molecules seem to be the most important factor in the case of low-osmolar and iso-osmolar contrast media. The iso-osmolar dimeric contrast media showed strong cytotoxic effects, which were significantly ($P < .001$) more pronounced when they were compared with those of the low-osmolar monomeric contrast media at equimolar concentrations. This finding suggests a greater potential for nephrotoxic effects of the dimeric contrast media molecules when they are compared with the monomeric contrast media molecules.

**Practical application:** The design of our experiments, especially the use of iso-osmolar formulations of monomeric contrast media, allows comparison between the biologic activity of the molecules of different types of contrast media, with exclusion of the physicochemical properties of the solutions, such as osmolality. This may offer a further approach to study which molecular mechanisms cause contrast-induced nephropathy. For clinical practice, the fact that dimeric contrast media molecules are more toxic than monomeric nonionic contrast media molecules in vitro means that one should exercise caution when one is advising radiologists about a preference for use of dimeric contrast media in at-risk patients on the basis of findings in available studies. It seems possible that the different levels of biologic activity of the molecules are especially important in subgroups of patients with certain comorbidities or in whom certain medications are used concomitantly with contrast media and that these patients have been unequally distributed in the few studies with small numbers of these patients so far.

**References**


8. Barrett BJ, Carlisle EJ. Metaanalysis of the
 PURPOSE: To evaluate the effect of various multi–detector row computed tomographic (CT) reconstruction parameters and nodule segmentation thresholds on the accuracy of volumetric measurement of synthetic lung nodules.

MATERIALS AND METHODS: Synthetic lung nodules of four different diameters (3.2, 4.8, 6.4, and 12.7 mm) were scanned with multi–detector row CT. Images were reconstructed at various section thicknesses (0.75, 1.0, 2.0, 3.0, and 5.0 mm), fields of view (30, 20, and 10 cm), and reconstruction intervals (0.5, 1.0, and 2.0 mm). The nodules were segmented from the simulated background lung region by using four segmentation thresholds (300, 400, 500, and 600 HU), and their volumes were estimated and compared with a reference standard (measurements according to fluid displacement) by computing the absolute percentage error (APE). APE was regressed against nodule size, and multivariate analysis of variance (MANOVA) was performed with APE as the dependent variable and with four within-subject factors (field of view, reconstruction interval, threshold, and section thickness).

RESULTS: The MANOVA demonstrated statistically significant effects for threshold (P = .02), section thickness (P < .01), and interaction of threshold and section thickness (P = .04). The regression of mean APE values on nodule size indicates that APE progressively increases with decreasing synthetic nodule size (R² = 0.99, P < .01).

CONCLUSION: For accurate measurement of lung nodule volume, it is critical to select a section thickness and/or segmentation threshold appropriate for the size of a nodule.

© RSNA, 2005
for detecting volumetric changes (9,10). Because nodules that are considered malignant would be further evaluated or treated more aggressively, it is important to obtain accurate and reliable measurements of nodule volume changes and thus increase diagnostic confidence. Consequently, with increasing applications of thin-section CT and lung cancer screening, the assessment of nodule growth has been an active area of research (11–16). Computer-aided volumetric assessment of small pulmonary nodules for estimating growth is gaining wider acceptance (11,12,15).

The acquisition of appropriate three-dimensional imaging data is a prerequisite for accurate and reliable nodule volume measurement. There are many scanning and reconstruction parameters that affect the quantification of nodule volume (17,18). In general, improvement of spatial resolution by means of reduction of section thickness, field of view (FOV), and reconstruction interval should decrease errors in nodule volume measurement. The relative importance of these reconstruction parameters on the measurement of nodule volume, however, is not equal and likely depends on the size of nodules. Furthermore, with three-dimensional measurement methods, the nodules are segmented from the background lung parenchyma on CT images prior to volume measurement. The segmentation is usually based on gray-level thresholding, and the choice of threshold values influences the measured volumes of the segmented nodules. Although authors of several studies have investigated the effects of nodule size (13,14,19), thresholds (13,19), and section thickness (14) on the volumetric measurement of lung nodules, the combined effects of these parameters have not been reported. Thus, the purpose of our study was to evaluate the effects of various multi-detector row CT reconstruction parameters and nodule segmentation thresholds on the accuracy of volumetric measurement of synthetic lung nodules.

MATERIALS AND METHODS

Synthetic Lung Nodules and the Reference Standard

Acrylic spheres of four different diameters (3.2, 4.8, 6.4, and 12.7 mm) were used to simulate lung nodules. The size of our synthetic small nodules was selected such that the lower size limit was 3 mm, because this is commonly considered to be the lowest detection limit on clinical lung CT images. The upper size limit of a small nodule was chosen to be 13 mm, which is slightly smaller than half of the upper limit of a nodule as defined by the Fleischner Society (20). The mean attenuation value of these spheres was approximately 130 HU, which is recognized as higher than that of clinically noncalcified soft-tissue nodules (30–60 HU). To obtain the reference-standard volume of the nodules, we weighed the nodule on an analytical chemistry balance, with a precision of 0.1 mg. The reference-standard volume of a nodule was calculated by multiplying the measured weight by the specific gravity. The specific gravity of a nodule was computed from the weight of the largest nodule divided by the volume that was determined by using water displacement. This approach was based on the observation that the volume measurement of the largest nodule was less variable than that of the smaller nodules obtained with water displacement. The weights and volume measurements were repeated three to four times and were averaged to calculate the values used in our analyses. The apparent diameter of each nodule was calculated from the estimated nodule volume with the assumption that the nodules were spherical in shape.

A chest CT phantom (Computerized Image Reference Systems, Norfolk, Va) was used to simulate a 5-cm-thick transverse section of the thorax (Fig 1). The lung region of the chest CT phantom was filled with a sculpted slab of polystyrene plastic matrix material (mean attenuation value, −990 HU). The synthetic nodules were embedded in this matrix material.

Imaging Protocols

The phantom was placed on a CT table with its transverse imaging plane (Fig 1) perpendicular to the direction of table motion. The scans were obtained by using a 16-detector row CT scanner (Somatom Sensation 16; Siemens, Forchheim, Germany) with a 16 × 0.75-mm collimation, 120 kVp, 255 mA, and 500-msec gantry rotation time in a spiral mode. Several image sets were reconstructed from the raw data by choosing five different section thicknesses (0.75, 1.0, 2.0, 3.0, and 5.0 mm) and three FOVs (30, 20, and 10 cm). The FOVs of 30, 20, and 10 cm correspond to in-plane resolution of 0.59, 0.39, and 0.20 mm, respectively. The images were reconstructed with three intervals (0.5, 1.0, and 2.0 mm) by using a lung image reconstruction kernel (B60f sharp).

Nodule Volume Measurement on CT Images

All CT image sets were transferred to a personal computer with a 2.8-GHz Pentium IV processor (Intel, Santa Clara, Calif) and 1 GB of RAM. The images were displayed and processed by using a commercially available software program (Rapidia; Infinitt, Seoul, Korea). The program allowed the user to segment and measure the volume of the nodules at a user-defined attenuation threshold. After a nodule was identified and clicked on the screen by using a computer mouse, it was segmented in a fully automated fashion on the gray-level thresholds. All of the nodules volume measurements were performed by one radiologist (J.M.G., 13 years of experience with chest CT). We used four thresholds: −300, −400, −500, and −600 HU. These values were selected arbitrarily around the mean attenuation (−430 HU) of the synthetic nodule (130 HU) and lung (−990 HU) attenuations.

A total of 720 nodule volume measurements (five section thicknesses × three FOVs × three reconstruction intervals × four thresholds × four nodules) were evaluated.

Data and Statistical Analyses

For each CT image-based nodule volume measurement, the absolute percentage error (APE) was calculated as 100 × |Vm − Vr|/Vm, where Vm and Vr are the CT measured and reference-standard nodule volumes, respectively.
Mean APE values were calculated for the four sizes of synthetic nodules. Regression analysis was used to assess the association between the mean APE value and nodule size. A multivariate analysis of variance was performed with four within-subject factors (FOV, reconstruction interval, threshold, and section thickness). Interaction terms were included in the analysis. The Mauchley test was used to evaluate the sphericity assumption. To help assess interactions, the percentage error was plotted against nodule diameter, FOV, reconstruction interval, threshold, and section thickness. Statistical analyses were performed with Statistica software (version 6.0; StatSoft, Tulsa, Okla) and JMP statistical software (version 5.1; SAS, Cary, NC). P < .05 indicated a statistically significant difference.

RESULTS

The reference-standard volumes of the four nodules were estimated to be 16.5, 56.6, 135.0, and 1072.5 mm$^3$, which corresponded to apparent spherical diameters of 3.2, 4.8, 6.4, and 12.7 mm, respectively.

We were unable to segment and measure the volume in 13 of 720 nodule combinations, because these nodules were faintly delineated because of partial volume effects. All of the missing values occurred with nodule diameters of 3.2 mm. Five values occurred at FOVs of 30 cm, four occurred at FOVs of 20 cm, and four occurred at FOVs of 10 cm. Seven missing values occurred at reconstruction intervals of 2.0 mm, three occurred at reconstruction intervals of 1.0 mm, and three occurred at reconstruction intervals of 0.5 mm. Ten of the missing values occurred at thresholds of −300 HU and three occurred at thresholds of −400 HU. Twelve missing values occurred at a section thickness of 5.0 mm and one occurred at a section thickness of 3.0 mm. Because we could not segment these nodules, our best estimate of their volumes was zero. The APEs for these measurements were, therefore, 100.

The multivariate analysis of variance demonstrated statistically significant effects for threshold (P < .02), section thickness (P < .01), and the interaction of threshold and section thickness (P = .04). Findings of Mauchley tests of sphericity were nonsignificant for these effects (P ≥ .06); therefore, the reported P values are for the univariate F test. FOV, reconstruction interval, and other interactions (second, third, and fourth degree) were not statistically significant (P > .05).

Plots of mean APEs versus nodule sizes for five different section thicknesses are presented in Figure 2. Similar plots for four different thresholds are presented in Figure 3. The plots indicate consistently increasing APE with decreasing nodule size except for the threshold value of −600 HU for the 3.2-mm nodule size. To understand this exception, APEs for a 3.2-mm nodule and at threshold of −600 HU were tabulated (Tables 1, 2).

Regression analysis was used to assess the overall association between nodule size and APE. To do this, grand means of pooled data were calculated for APE values.

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Mean APE Measured in a 3.2-mm Nodule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Threshold (HU) and Section Thickness (mm)</td>
<td>Mean APE*</td>
</tr>
<tr>
<td>−300</td>
<td>51.7†</td>
</tr>
<tr>
<td>0.75</td>
<td>36.0</td>
</tr>
<tr>
<td>1.0</td>
<td>28.5</td>
</tr>
<tr>
<td>2.0</td>
<td>34.3</td>
</tr>
<tr>
<td>3.0</td>
<td>63.6</td>
</tr>
<tr>
<td>5.0</td>
<td>96.3</td>
</tr>
<tr>
<td>−400</td>
<td>32.0†</td>
</tr>
<tr>
<td>0.75</td>
<td>15.1</td>
</tr>
<tr>
<td>1.0</td>
<td>13.6</td>
</tr>
<tr>
<td>2.0</td>
<td>16.3</td>
</tr>
<tr>
<td>3.0</td>
<td>36.4</td>
</tr>
<tr>
<td>5.0</td>
<td>78.6</td>
</tr>
<tr>
<td>−500</td>
<td>13.9†</td>
</tr>
<tr>
<td>0.75</td>
<td>3.9</td>
</tr>
<tr>
<td>1.0</td>
<td>4.5</td>
</tr>
<tr>
<td>2.0</td>
<td>6.6</td>
</tr>
<tr>
<td>3.0</td>
<td>14.3</td>
</tr>
<tr>
<td>5.0</td>
<td>40.4</td>
</tr>
<tr>
<td>−600</td>
<td>13.6†</td>
</tr>
<tr>
<td>0.75</td>
<td>16.2</td>
</tr>
<tr>
<td>1.0</td>
<td>16.7</td>
</tr>
<tr>
<td>2.0</td>
<td>19.3</td>
</tr>
<tr>
<td>3.0</td>
<td>10.9</td>
</tr>
<tr>
<td>5.0</td>
<td>4.8</td>
</tr>
</tbody>
</table>

* Values are the means of the pooled data for reconstruction interval and FOV.
† Values are the means of APEs for all section thicknesses.

Figu re 2. Graph shows the mean APE versus nodule size for five different section thicknesses. At each section thickness, APE increased consistently with a decrease in nodule size.

Figu re 3. Graph shows the mean APE versus nodule size for four different thresholds. At each threshold, except for the threshold value of −600 HU and 3.2-mm nodule, APE increased consistently with a decrease in nodule size.
values, and log transformations of the data were performed to fit the data in a linear regression model (21). The resulting regression of mean APE values on synthetic nodule sizes was highly significant ($R^2 = 0.99, P < .01$).

On scatter plots of percentage errors against nodule diameter, FOV, reconstruction interval, threshold, and section thickness, nodule volume tends to be overestimated with a threshold of $-600\, \text{HU}$ and underestimated with thresholds of $-300$ and $-400\, \text{HU}$. As the size of a nodule decreased, the range of error increased in both directions (Fig 4).

**DISCUSSION**

Because multi–detector row CT facilitates the acquisition and analysis of three-dimensional CT images, the use of volumetric analysis of lung nodules is likely to increase. The primary clinical application of nodule volumetric measurements is to monitor nodule growth (11,12). The three-dimensional measurements have obvious advantages over conventional binlinear measurements for accurately representing volumes. Authors (12) of a study about three-dimensional estimation of nodule volume reported that the volume of nodules as small as 3 mm in diameter could be measured with a high accuracy.

An image-based nodule volume measurement method, however, has many intrinsic and practical limitations and is subject to measurement errors. For correct interpretation of nodule growth, it is critical to understand how various parameters affect nodule volume measurements and become sources of measurement error. Ko et al (13) evaluated the effect of imaging variables on the volumetric measurement of lung nodules by using a realistic phantom and suggested that a partial volume method, a high-frequency reconstruction algorithm, and CT performed with a diagnostic dose (120 mAs) would improve the precision of the volume measurements of lung nodules. Winer-Muram et al (14) measured and analyzed the volumes of simulated nodules in phantoms and in patients with stage I lung cancer and reported that the lung nodule volumes in phantoms were overestimated; this overestimation was directly proportional to the image section thickness and inversely proportional to the nodule size.

In the current study, we analyzed the effect of four key parameters (threshold, section thickness, FOV, and reconstruction interval) on nodule volume measurements by using synthetic lung nodules. Our results demonstrated the relative significance of these parameters with regard to how they affect measurement error. We believe this information contributes to the optimization of imaging protocols for accurate and repeatable nodule volume measurements with serial CT studies.

It is well known that varying the window center alters the measured volume even for relatively large lesions or nodules (9,19,22). Our study findings also revealed that the segmentation threshold is a significant determinant of measurement errors. The most common approach for measuring the volume of lung nodules is based on gray-level thresholding, which allows a user to segment lung nodules from the background parenchyma and measure the volume of the segmented voxels (12,13,15,16). The selection of an appropriate threshold for lung nodule segmentation is crucial. Optimal threshold levels may be determined from valleys on the histogram of lung nodule and parenchyma attenuations or from levels that provide a statistical maximal separation between the regions to be distinguished (16). Previous study findings have demonstrated that a structure is most accurately represented when the display window center is set midway between the CT attenuation of the structure of interest and the background (22). In our study, thresholds of $-500$ and $-400\, \text{HU}$, which are closer to the mean of the attenuations of the nodule (130 HU) and the background ($-990\, \text{HU}$), resulted in smaller errors than thresholds of $-600$ and $-300\, \text{HU}$ (Fig 3). In practice, the application of a single fixed threshold is the simplest and most consistent approach for segmenting nodules. The midway attenuation between the nodules and parenchyma may not be fixed, however, because the attenuation values of nodules vary depending on their type. Furthermore, segmentation for nonsolid and part-solid nodules frequently requires a much lower threshold than that required for solid nodules. Even in the same nodule, small peripheral lung neoplasms may have a replacement growth pattern. As found in a study by Kakinuma et al (23), the ground-glass opacities of lung cancer nodules not only increase in size or attenuation but may also decrease rapidly or slowly with the appearance of solid components. This variable growth pattern could not be recognized with the use of a single fixed threshold.

The results of our study also demonstrated that section thickness is an important parameter in nodule volume measurement. With multi–detector row CT, thin-section volumetric imaging in the whole lung becomes routine, particularly for lung cancer screening studies (24,25). A drawback of this practice is that a thin-section CT scan of the whole thorax generates a large dataset (typically 250–350 images of 1.0-mm section thickness), which requires considerable time to interpret and analyze (26).

With the CT image matrix size fixed at $512 \times 512$, the voxel dimension is proportional to the FOV. Some studies used a FOV of 9.6 cm (11,12) in measuring nodule volumes, while another study used a 30-cm FOV (13). The effect of FOV on volume measurement has not, however, been studied. In the current study, FOV did not have a significant effect on the APE. This result is not surprising, in part, because the size of in-plane resolution differences (0.20–0.59 mm) due to the choices in FOV in our study was much lower than that of longitudinal resolution differences (0.75–5.0 mm) due to the choices in section thickness.

In three-dimensional imaging, the reconstruction interval is one of the parameters that affects the longitudinal resolution. With smaller reconstruction intervals, the number of overlapping

<table>
<thead>
<tr>
<th>Table 2: Mean APE at Threshold of $-600, \text{HU}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nodule Diameter (mm) and Section Thickness (mm)</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>3.2</td>
</tr>
<tr>
<td>0.75</td>
</tr>
<tr>
<td>1.0</td>
</tr>
<tr>
<td>2.0</td>
</tr>
<tr>
<td>4.8</td>
</tr>
<tr>
<td>0.75</td>
</tr>
<tr>
<td>1.0</td>
</tr>
<tr>
<td>2.0</td>
</tr>
<tr>
<td>3.0</td>
</tr>
<tr>
<td>5.0</td>
</tr>
<tr>
<td>6.4</td>
</tr>
<tr>
<td>0.75</td>
</tr>
<tr>
<td>1.0</td>
</tr>
<tr>
<td>2.0</td>
</tr>
<tr>
<td>3.0</td>
</tr>
<tr>
<td>5.0</td>
</tr>
<tr>
<td>12.7</td>
</tr>
<tr>
<td>0.75</td>
</tr>
<tr>
<td>1.0</td>
</tr>
<tr>
<td>2.0</td>
</tr>
<tr>
<td>3.0</td>
</tr>
<tr>
<td>5.0</td>
</tr>
</tbody>
</table>

* Values are the means of the pooled data for reconstruction interval and FOV.

† Values are the means of APEs for all section thicknesses.
transverse images increases and the longitudinal resolution improves. Brink et al. (27) suggested that single-detector row helical CT images should be reconstructed with at least 60% overlap relative to the effective section thickness to

Figure 4. Scatter plots show relationships between the percentage error and nodule diameter, FOV, reconstruction interval, threshold, and section thickness. The nodule volume was overestimated at a threshold of \(-600\) HU but was underestimated at thresholds of \(-300\) and \(-400\) HU. As the size of nodules decreases, errors increase in both positive and negative directions.
obtain maximal longitudinal resolution. It is possible that a further increase in the overlap percentage may not provide a measurable gain in the longitudinal resolution. Although reconstruction interval was not expressed as a percentage of section thickness in our study, reconstruction interval itself was not a significant parameter in determining volume measurement error.

For a 3.2-mm nodule imaged on 5.0- or 3.0-mm-thick sections, the attenuation of all of the pixels within the nodule was below –300 HU in 10 cases and –400 HU in three cases because of the partial volume effect. Therefore, the volumes of these nodules were measurable at low thresholds (–500 or –600 HU) but not at high thresholds (–300 or –400 HU). We could not measure the volume of nodules in 2% of the cases. For these cases, we considered our measurement error to be 100%. We believe that this was a reasonable procedure for dealing with these cases.

In plots of error measurement against nodule size for thresholds and section thicknesses, the error progressively increased with decreasing nodule size except for the –600 HU threshold for the 3.2-mm nodule size. This phenomenon may have been caused by an interaction between parameters. Our statistical analysis demonstrated significant interaction between section thickness and threshold. Tables 1 and 2 were used to interpret this exception. In general, the error increases as the section thickness increases, and the nodule volume tends to be overestimated as the threshold decreases (Figs 2, 4). For a 3.2-mm nodule at a threshold of –600 HU, however, the error decreases as the section thickness increases (Tables 1, 2). The attenuation of the surface voxels of a nodule decreases as the section thickness increases because of partial volume effect, but at the same time the volume of the voxel also increases as the section thickness increases. The volume measurement error, therefore, depends on both the threshold and the section thickness; with some combinations, both parameters act in the same direction, and with other combinations, the parameters act in opposition direction. Compared with large nodules, small nodules are sensitive to these effects because the proportion of surface voxels is larger in small nodules. For a 3.2-mm nodule, thick section thickness acted in the direction of underestimation, which counteracted the effect of overestimation by a threshold of −600 HU.

Our analysis also demonstrated a statistically significant decreasing relationship between measurement error and nodule size: Smaller nodules tended to have higher measurement errors. As nodules become smaller, the partial volume effect and measurement error increase (12, 14, 15). This effect is reduced with the use of smaller voxel or isotropic voxel size.

There were several limitations of our study. We used motion-free spherical homogeneous nodule phantoms, which do not accurately represent clinical reality. All of the synthetic nodules had the same attenuation value, and nodules with attenuation equivalent to ground-glass opacity of lesions were not included. The software package we used was not specifically designed for measuring lung nodule volume, and the measurement of nodule volumes was based on a threshold holding method. We, however, believe that this simplified phantom study design has allowed us to achieve the aim of our study, that is, the evaluation of the effect of reconstruction and technical parameters on nodule volume measurement error.

In summary, the results of our study demonstrated that nodule size, section thickness, and threshold were significant parameters in determining volume measurement errors. The FOV and reconstruction interval did not significantly affect volume measurement error. For an accurate volumetric measurement of a lung nodule, it is critical to select a section thickness and/or segmentation threshold appropriate for the size of the nodule.

Practical application: Accurate measurement of nodule volume is crucial in the current scheme of imaging-based assessment of malignancy in small nodules. Because the acquisition and reconstruction parameters for CT imaging can significantly affect the accuracy of nodule volume measurement, it is important to select appropriate parameters to achieve a desired accuracy. The results of the present study suggest that we should be more attentive to the selection of section thickness and segmentation threshold than to selection of FOV and reconstruction interval for an accurate measurement of nodule volume.

References
Digital Slot-Scan Charge-coupled Device Radiography versus AMBER and Bucky Screen-Film Radiography: Comparison of Image Quality in a Phantom Study

PURPOSE: To evaluate the image quality and performance of a chest digital radiography system and to compare this with the image quality and performance of advanced multiple-beam equalization radiography (AMBER) and Bucky screen-film radiography systems.

MATERIALS AND METHODS: The chest digital radiography system is a digital charge-coupled device (CCD) chest imaging unit that uses slot-scan technology. A contrast-detail test object was used in combination with a phantom that simulates the primary and scatter transmission for the lungs and mediastinum. Twelve phantom images were obtained with each modality (ie, CCD digital radiography and AMBER and Bucky screen-film radiography) and were judged by six observers. CCD digital radiography soft-copy reading was compared with AMBER hard-copy reading. To measure image quality, contrast-detail curves were constructed from the observer data. The Wilcoxon signed rank test was used for statistical analysis.

RESULTS: For the lung configuration, contrast-detail curves showed lower threshold depth for hard-copy images obtained with CCD digital radiography than for those obtained with Bucky screen-film radiography. For hard-copy images, the difference between contrast-detail curves for CCD digital radiography and those for Bucky screen-film radiography was statistically significant ($P < .006$). No significant difference was found between CCD digital radiography and AMBER for hard-copy images obtained in either the lung or mediastinum configuration. For the lung configuration, a lower threshold depth was observed for CCD digital radiography soft-copy reading than for AMBER hard-copy reading, with significantly different contrast-detail curves for CCD digital radiography soft copy and AMBER hard copy ($P < .006$). No significant difference was found between either system for the mediastinum configuration.

CONCLUSION: Contrast-detail curves indicate that image quality for the CCD chest system provides a digital alternative to AMBER and Bucky screen-film radiography.

© RSNA, 2005
capture and image processing are performed without cassette handling. Thin-film transistor–based flat-panel detector systems have better image quality than conventional screen-film radiography and photostimulable phosphor plate computed radiography systems (2). The wide dynamic exposure range and linear response are the major advantages of digital radiography compared with screen-film radiography. Because of the large difference in x-ray absorption between the lungs and mediastinum, image contrast, particularly in the chest, is expected to improve with digital radiography (2).

Recently, a new digital radiography system (ThoraScan; Nucletron-Oldelft, Veenendaal, the Netherlands) was introduced for chest imaging. The ThoraScan unit contains a charge-coupled device (CCD)–based detector that is built from an array of detector elements. CCD detectors offer low noise and high sensitivity (3,4). With the ThoraScan unit, a slot-scan technique is used. Advantages of the slot-scan concept include dose-efficient scatter rejection and the possibility of using small detectors to image large areas without demagnification.

Contrast-detail studies and assessment of physical imaging characteristics, such as the modulation transfer function (MTF), provide important indications of image quality (5,6). Such studies are recommended as a first step in the assessment of new medical imaging equipment (6). Thus, the purpose of our study was to evaluate the image quality and performance of digital radiography systems for chest imaging (ThoraScan) and to compare this with the image quality and performance of advanced multiple-beam equalization radiography (AMBER) and Bucky screen-film radiography systems.

MATERIALS AND METHODS

ThoraScan System for Chest Imaging

See the Appendix for a description of the system.

AMBER and Bucky Screen-Film Radiography

The CCD digital radiography system was compared with AMBER (Nuclertron-OldeIn, the Netherlands) and Bucky wall stand (Optimus S/S; Philips Medical Systems, Best, the Netherlands) screen-film radiography systems. The AMBER unit uses a scanning x-ray beam, with beam intensity modulated in response to measurements from a detector located behind the film cassette. The system adjusts beam intensity to compensate for reduced optical density in dense regions, such as the mediastinum, with improved visualization of mediastinal structures and retrocardiac and retrodiaphragmatic areas (7).

Contrast-Detail Test Object

A contrast-detail radiographic test object (University Medical Center Nijmegen, Nijmegen, the Netherlands) was used for the contrast-detail study. This test object consisted of a polymethyl methacrylate tablet (area, 26.5 × 26.5 cm; thickness, 10.0 mm) with cylindrical holes of varying diameter and depth in a 15 × 15 matrix, which yielded a total of 225 cells (8). In each cell, either one or two holes are present. Cells in the first three rows contain only one central hole, while cells in the other rows contain two identical holes—one in the middle and one in a randomly selected corner (Fig 1). In the horizontal direction, the hole depth increases logarithmically from 0.3 to 8.0 mm. In the vertical direction, the hole diameter increases logarithmically from 0.3 to 8.0 mm.

Image Acquisition

Two different phantom configurations were used—one for simulating lung characteristics, and one for simulating mediastinum transmission characteristics. The phantom used for simulating lung characteristics was the Lucite and aluminum phantom described by Conway et al (9). In the direction from focus to detector or film, the lung phantom is composed of 1.0 cm of polymethyl methacrylate, 0.25 cm of aluminum, 4.4 cm of polymethyl methacrylate, a 19-cm air gap, 0.15 cm of aluminum, and 1.0 cm of polymethyl methacrylate. The 1-cm-thick contrast-detail test object was placed halfway into the 19-cm air gap. To simulate mediastinum transmission characteristics, an additional 9-cm layer of polymethyl methacrylate was inserted into the phantom just in front of the test object. To evaluate both the lung and the mediastinum characteristics, six images were acquired.
in each configuration with each of the three chest systems—CCD digital radiography, AMBER, and Bucky screen-film radiography—resulting in 36 images. Image sets were acquired to reduce the inaccuracy of contrast-detail measurement owing to quantum noise and variations in the quality of hard-copy images. For each system, the images were recorded consecutively. No problems were encountered with respect to heat loading for scanning systems with 12 consecutive phantom images or use of the imaging systems in daily clinical practice.

Table 1 gives parameter settings for the three chest imaging systems. Antiscatter grids were used for the screen-film radiography systems. For AMBER, a moving grid with a ratio of 12:1 and 36 lines per centimeter was used. For Bucky screen-film radiography, a moving grid with a ratio of 12:1 and 36 lines per centimeter was used. All images were acquired with tube voltage and filtration set according to the clinical protocols. Tube voltage and charge settings for each chest system are given in Table 2. Protocols were established by radiologists, physicists, and manufacturers together. Because of the scanning properties of CCD digital radiography and AMBER, exposure times and, therefore, tube charges were much larger for these systems than for Bucky screen-film radiography. Local exposure times, however, were short to avoid movement artifacts. The settings were chosen automatically for both the lung and the mediastinum configurations. Window width and window level settings for hard-copy images obtained with CCD digital radiography corresponded to the default clinical settings for an average-sized patient. The multiscale image contrast amplification postprocessing settings, used for both soft-copy and hard-copy images, were as follows: crossover, −2.0; edge contrast, 0.0; latitude reduction, 0.0; multiscale contrast, 3.0; and noise reduction, 1.0.

### Image Reading

The quality assurance system in our department includes quality control of equipment, including monitors, printers, and light boxes. The monitor and the laser printer satisfied the requirements of the American Association of Physicists in Medicine guidelines (10) with respect to low contrast and spatial resolution. The monitor (Dome C3; Dome Imaging Systems, Waltham, Mass) consisted of a 1536 × 2048-pixel gray-scale flat panel. The maximum luminance was 690 candela per square meter. For printing, a laser imager (Dryview 8700; Eastman Kodak, Rochester, NY) was used. CCD settings were determined by using dedicated lookup tables with respect to the printer and monitor. A setting of 8 bits was used to print hard-copy images.

For image reading, three radiologists and three physicists participated as observers. Among the radiologists was one author (L.J.K.M.). The radiologists had a mean experience of 16 years in reading chest images (range, 8–20 years). Among the physicists were two authors (W.J.H.V., J.G.). The physicists had a mean experience of 13 years in medical physics (range, 8–15 years). All 36 images were used for reading. For the lung configuration, the following paradigm was used: Each set of six lung images (corresponding to CCD digital radiography, AMBER, or Bucky screen-film radiography) was equally distributed among the six observers; that is, each observer scored two different images per set, and each image was scored twice by two different observers. Because reading was performed for all possible six permutations of the three modalities, the order in which the observers had to judge images from the different modalities cancelled out any learning effect.

For the mediastinum configuration, the Bucky screen-film images were initially left out of the scheme owing to poor image quality (ie, underexposed films), which reflected image characteristics in the mediastinal region on Bucky screen-film images. Image quality in this region is poor because of the small dynamic range of Bucky screen-film radiography systems compared with the wider dynamic range of AMBER systems and the much wider dynamic range of CCD digital radiography systems. Each of the remaining two sets of six images (CCD digital radiography and AMBER images) was distributed among the six observers in the same way as for the lung configuration. The order of images from the two modalities in which reading was performed was evenly alternated among the observers.

A four-alternative forced-choice detectability experiment was performed. The first three rows on the phantom were not evaluated because these rows contained only a central hole. Observers were instructed to mark the dot-containing corner in each cell even if they were uncertain of the dot’s location. All observers were familiarized with the experimental procedure prior to initiating each test. All images were judged on the same regular light box (Rotolux Planilux; Gerätebau F. Schulte, Warstein, Germany), and a pen was used to mark the detected location of the dot on the hard-copy images. The observers were free to choose their viewing distance. They were not allowed to move or lift the film to improve reading. Mediastinum images were presented first, followed by lung images. With this sequence, performance bias owing to learning effect was minimized because the mediastinum images had an overall worse image quality.

After these sessions, three observers, including two authors (W.J.H.V., J.G.), additionally scored the mediastinum images obtained with the Bucky screen-film radiography system. The corresponding set of six images was evenly distributed among the observers, and each observer scored two different images.

Compared CCD digital radiography soft-copy reading with AMBER film hard-copy reading, the same three observers scored the CCD digital radiography me-

<table>
<thead>
<tr>
<th>Parameter Settings for the Three Chest Imaging Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter</strong></td>
</tr>
<tr>
<td>Screen type</td>
</tr>
<tr>
<td>Film speed</td>
</tr>
<tr>
<td>Film type</td>
</tr>
<tr>
<td>Maximum kVp</td>
</tr>
<tr>
<td>Nominal focal spot size (mm)</td>
</tr>
<tr>
<td>Detector-to-focus distance (cm)</td>
</tr>
<tr>
<td>External filtration (mm)</td>
</tr>
<tr>
<td>Anode angle (degrees)</td>
</tr>
<tr>
<td>Anode diameter (mm)</td>
</tr>
<tr>
<td>Tube permanent filtration (mm)</td>
</tr>
</tbody>
</table>

* Eastman Kodak, Rochester, NY.

Note.—Al = aluminum, Cu = copper.
diastium and lung images more than 2 months later from the monitor. These images were the same as those that were scored from hard copy by the observers. The observers indicated the dot-containing corners by using a pencil and a template paper. The images were fully displayed by using the approximate full size of the monitor. Observers were not allowed to use other soft-copy tools. For the lung setting, the window width and window level were the same as those used for printing hard-copy images. This setting was the default window width and level used in clinical practice for an average-sized patient. Also, for the mediastinum setting, a clinically relevant window width and window level were chosen (ie, the window width was kept unchanged, but the window level was adjusted so that it was equal to the mean pixel value for these images). Observers were not allowed to adjust the window width or level settings. For dot detection, it would have been advantageous to choose a very small window width because the image background is uniform. The use of such a small window width, however, is not clinically relevant, and the corresponding results would therefore be biased in comparison with screen-film radiography results.

MTF Measurements

Measurements of the presampled MTF were performed by two authors (W.J.H.V., J.G.) who used the method described by Samei et al (11). An attenuating, thin-edge test device was fabricated for measurements. A 0.2-mm-thick tungsten layer (Goodfellows, Cambridge, England) was used as the attenuating material. Tungsten was preferred instead of lead because tungsten has a higher attenuation coefficient. With tungsten, a thinner layer can be used, which reduces the influence of penumbra effects. The layer measured 5 \( \times \) 10 cm, with a purity of 99.95%. For such thickness, the attenuation was assessed to be 84% for a typical 133-kVp polychromatic x-ray beam by using a spectrum generator (Institute of Physics and Engineering in Medicine, York, England) (12). The edge was created with wire electrical discharge machining and resulted in an accurate straight edge. The edge was polished with light pressure and high precision by using a wet stone. The edge was placed with a small vertical and horizontal angle at the center of the detector. A 2.5-cm-thick aluminum filter was inserted just behind the x-ray tube. The same filter was used in the complete calibration procedure.

Contrast Measurements

The dynamic range and optical density were measured with image contrast, \( \Delta C \), as a function of hole depth. Image contrast was measured by one author (W.J.H.V.) for both the lung and mediastinum configurations by using a densitometer (X-Rite, Grand Rapids, Mich). Contrast was computed as \( \Delta C = 1 - \Omega^{-2\Delta OD} \), where \( \Delta OD \) is the differential optical density of the hole with the largest diameter (2). Mean contrast values were obtained from two AMBER and two Bucky screen-film radiography hard-copy images and from two CCD digital radiography hard-copy images.

Radiation Exposure

Dose measurements were obtained by two authors (W.J.H.V., J.G.). Entrance dose was calculated with a 15-cm\(^3\) ionization chamber (Keithley Instruments, Cleveland, Ohio). In one experiment, we derived the time-dependent entrance dose from the dose-area product with respect to the CCD digital radiography system when scanning an average-sized patient (length, 170 cm; weight, 70 kg) and including the prescan. The dose-area product was measured with a dosimeter (Diamentor M4; PTW-Freiburg, Freiburg, Germany). For AMBER, entrance doses for the lung and mediastinum configurations were measured separately. This was necessary because the AMBER system increases the radiation dose to the patient when scanning the mediastinum.

Data and Statistical Analysis

Image evaluation was performed by one author (W.J.H.V.) by placing a matching reference standard paper over the annotated images. The detection probability for each cell was determined from the six different scoring results for each system and for either simulated configuration (lung or mediastinum). We used a model-based interpolation scheme to fit a curve through the observer data, as described by Karssemeijer and Thijsen (13). With this model, the probability of detecting a hole with certain diameter as a function of its depth is described by means of a psychometric curve with probability range from 0.25 (chance) to 1.00. For hole diameter, a threshold depth was found that corresponded to 62.5% correct responses. This threshold was just halfway up the psychometric curve. Finally, a contrast-detail curve was fitted through the threshold depth values by using a second-order polynomial fit.

Differences between estimated threshold depths in the contrast-detail curves were calculated for each hole diameter. Pair-wise testing was performed for the lung and mediastinum configurations by using the Wilcoxon signed rank test to evaluate the performance of CCD digital radiography with respect to either AMBER or Bucky screen-film radiography and to assess contrast-detail observations and the differences in contrast range between hard-copy images obtained with CCD digital radiography and AMBER (14). Testing was performed by a statistician (B.J.A.M.). We evaluated eight tests in total. The Wilcoxon signed rank test is a nonparametric test. This means that a conservative approach is used. For each comparison, the \( P \) value, the Wilcoxon signed rank test statistic (V), and the number of paired samples \( (n) \) were determined. To correct for multiple testing, we suggest applying a Bonferroni correction to the results. Because we present eight tests, this implies a significance level of .006 (for each individual test) instead of .05 (for example) to maintain a global .05 significance level across all tests included. With respect to this adjustment, readers should note that the Bonferroni method is a rather conservative approach to adjust for the multiple testing problem. We also remind readers that the Bonferroni threshold (ie, .006) should, of course, not be used as a rigid cutoff value.

<table>
<thead>
<tr>
<th>Parameter Settings</th>
<th>CCD Digital Radiography</th>
<th>AMBER Screen-Film Radiography</th>
<th>Bucky Screen-Film Radiography</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tube voltage (kV)</td>
<td>133</td>
<td>133</td>
<td>125</td>
</tr>
<tr>
<td>Tube charge (mAs)</td>
<td>82</td>
<td>100</td>
<td>2</td>
</tr>
<tr>
<td>Entrance dose (( \mu \text{Gy} ))</td>
<td>89</td>
<td>59, 184*</td>
<td>64</td>
</tr>
</tbody>
</table>

* Data indicate entrance dose levels for lung and mediastinum, respectively.
just as the usual level (i.e., .05) should never be taken as fixed decision boundary in individual tests. Rather, the .006 value is presented here to give some guidance on the order of magnitude of adjustment needed in interpreting $P$ values and, thus, in helping readers with their own interpretation of the data analysis.

RESULTS

Contrast-Detail Experiment

The threshold depth for each hole was determined by using four-alternative forced-choice detectability experiment psychometric detection curves (Fig 2). The curves show CCD digital radiography, AMBER, and Bucky screen-film radiography observations for the lung configuration. As can be expected, the detection curves shift to the left with an increased hole diameter, indicating better detection probability with increased hole size (Fig 2b vs 2a). Note that the CCD digital radiography curve is to the left of the AMBER and Bucky screen-film radiography curves for similar hole sizes, indicating better detectability with CCD digital radiography. For all contrast-detail experiments in this study, the threshold depth values for all imaging modalities were determined (Table 3).

The matching detection curves from the four-alternative forced-choice detectability experiment were fitted through the threshold depth values (Fig 3). For the lung configuration, the CCD digital radiography hard-copy contrast-detail curve is under the AMBER hard-copy curve, but the difference between the two curves was not statistically significant ($P = .170, V = 0$, and $n = 11$) (Fig 3a). The

<table>
<thead>
<tr>
<th>Hole Diameter (mm)</th>
<th>Lung Configuration</th>
<th>Mediasitum Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CCD Digital Radiography, Hard Copy</td>
<td>CCD Digital Radiography, Soft Copy</td>
</tr>
<tr>
<td>0.4</td>
<td>7.07</td>
<td>3.11</td>
</tr>
<tr>
<td>0.5</td>
<td>3.57</td>
<td>2.69</td>
</tr>
<tr>
<td>0.6</td>
<td>1.79</td>
<td>1.76</td>
</tr>
<tr>
<td>0.8</td>
<td>1.71</td>
<td>1.44</td>
</tr>
<tr>
<td>1.0</td>
<td>1.21</td>
<td>1.06</td>
</tr>
<tr>
<td>1.3</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>1.6</td>
<td>0.92</td>
<td>0.90</td>
</tr>
<tr>
<td>2.0</td>
<td>0.58</td>
<td>0.49</td>
</tr>
<tr>
<td>2.5</td>
<td>0.67</td>
<td>0.41</td>
</tr>
<tr>
<td>3.2</td>
<td>0.52</td>
<td>0.28</td>
</tr>
<tr>
<td>4.0</td>
<td>0.45</td>
<td>0.41</td>
</tr>
</tbody>
</table>

Note.—Threshold hole depths for given hole diameter as calculated from the pool of observers. Lower threshold values indicate better contrast-detail detectability performance for that chest imaging system. Ellipses indicate that a threshold depth value could not be calculated because detection probabilities for the corresponding diameter were too low (i.e., too close to chance) to fit a psychometric curve through the data.

Figure 2. Graphs demonstrate four-alternative forced-choice detectability experiment psychometric detection curves in lung configuration for hard-copy images obtained with CCD digital radiography (○), AMBER (●), and Bucky screen-film radiography (□). Hole sizes are (a) 1.0 mm and (b) 1.3 mm. Individual data points are derived from scoring data of six observers. Threshold depth corresponds to 62.5% (arrow) correct responses. Threshold 1 corresponds to CCD digital radiography, threshold 2 to AMBER, and threshold 3 to Bucky screen-film radiography. Note that a curve shift to the left indicates better detectability.

TABLE 3
Threshold Depth Values for Hole Detection in the Contrast-Detail Phantom for CCD Digital Radiography, AMBER, and Bucky Screen-Film Radiography

<table>
<thead>
<tr>
<th>Hole Diameter (mm)</th>
<th>Lung Configuration</th>
<th>Mediasitum Configuration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CCD Digital Radiography, Hard Copy</td>
<td>CCD Digital Radiography, Soft Copy</td>
</tr>
<tr>
<td>0.4</td>
<td>7.07</td>
<td>3.11</td>
</tr>
<tr>
<td>0.5</td>
<td>3.57</td>
<td>2.69</td>
</tr>
<tr>
<td>0.6</td>
<td>1.79</td>
<td>1.76</td>
</tr>
<tr>
<td>0.8</td>
<td>1.71</td>
<td>1.44</td>
</tr>
<tr>
<td>1.0</td>
<td>1.21</td>
<td>1.06</td>
</tr>
<tr>
<td>1.3</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>1.6</td>
<td>0.92</td>
<td>0.90</td>
</tr>
<tr>
<td>2.0</td>
<td>0.58</td>
<td>0.49</td>
</tr>
<tr>
<td>2.5</td>
<td>0.67</td>
<td>0.41</td>
</tr>
<tr>
<td>3.2</td>
<td>0.52</td>
<td>0.28</td>
</tr>
<tr>
<td>4.0</td>
<td>0.45</td>
<td>0.41</td>
</tr>
</tbody>
</table>
CCD digital radiography hard-copy curve also shows lower threshold depth compared with the Bucky screen-film radiography hard-copy curve, with significantly different contrast-detail results for CCD digital radiography hard copy compared with Bucky screen-film radiography hard copy \((P < .006, V = 0, \text{ and } n = 11)\) (Fig 3a).

For the mediastinum configuration, the AMBER contrast-detail curve demonstrates lower threshold depth than the CCD digital radiography hard-copy curve; the difference between these two curves, however, was not statistically significant \((P = .10, V = 52, \text{ and } n = 10)\) (Fig 3b). The CCD digital radiography hard-copy curve was located under the Bucky screen-film radiography curve (Fig 3b). Although the difference in image quality was visually apparent, the difference between the two curves was not statistically significant \((P = .030, V = 0, \text{ and } n = 6)\). The power of the statistical test was poor in this instance because a smaller number of paired data points (six) was involved. The CCD digital radiography soft-copy curve showed lower threshold depth than the AMBER hard-copy curve for the lung configuration, with significantly different contrast-detail detectability between CCD digital radiography soft-copy reading and AMBER hard-copy reading \((P < .006, V = 0, \text{ and } n = 11)\) (Fig 3a). For the mediastinum configuration, no significant difference was found between CCD digital radiography soft-copy reading and AMBER hard-copy reading (Fig 3b). In addition, the CCD digital radiography soft-copy curve showed a lower threshold depth than the CCD digital radiography hard-copy curve for both the lung and mediastinum configurations. The difference between contrast-detail curves for CCD digital radiography soft-copy reading and CCD digital radiography hard-copy reading was significant for both the lung \((P < .006, V = 1, \text{ and } n = 11)\) and mediastinum \((P < .006, V = 0, \text{ and } n = 10)\) configurations.

MTF Measurements

The MTF results for images obtained with CCD digital radiography were measured with the thin-edge device (Fig 4). The Nyquist frequency, which is the highest frequency that a digital image can unambiguously represent, is depicted in Figure 4. The Nyquist frequency of the CCD digital radiography system is 3.1 cycles per millimeter for 162-μm pixel size. In both the scanning and slot direction, the MTF values differ only slightly. For the lowest frequency (0.0–1.0 cycles per millimeter), the slot direction MTF seems to be somewhat better, whereas, for the higher frequency (>2.5 cycles per millimeter), the scanning direction MTF seems slightly superior.
Contrast Measurements

Image contrast values for AMBER and CCD digital radiography were determined for hard-copy images (Table 4, Fig 5). Hole depth ranged from 2.0 to 8.0 mm. For lower-contrast dots, ΔC became too small to be measured properly. The ΔC values for hard-copy images obtained with CCD digital radiography were lower than those for hard-copy images obtained with conventional AMBER for both the lung and mediastinum configurations (P = .020 and P = .060, respectively).

For the lung configuration, the ΔC values of hard-copy images obtained with CCD digital radiography were lower than the ΔC values of hard-copy images obtained with conventional Bucky screen-film radiography (P = .020). For the mediastinum configuration, the ΔC values of hard-copy images obtained with a small CCD detector reduces the costs.

MTF values for the CCD digital radiography system show good resolution capability in the scanning direction, as well as in the slot direction. Values of 0.70 at 1 cycle per millimeter and 0.35 at 2 cycles

Radiation Exposure

The entrance dose for CCD digital radiography is about 40% higher than the entrance dose for Bucky screen-film radiography (0.089 vs 0.064 mGy) (Table 2). The entrance doses for AMBER and Bucky screen-film radiography were similar in the lung configuration; entrance doses for AMBER, however, were higher in the mediastinum configuration. The time-dependent entrance dose measurement for the CCD digital radiography system was also determined (Fig 6). The entrance dose of the prescan is 7% (2.77 mGy/cm²) of the total entrance dose (40.10 mGy/cm²).

DISCUSSION

In this study, we describe the image quality and dose measurements of a new digital radiography technique for chest imaging. The size and shape of the CCD detector of the ThoraScan unit have some valuable advantages. First, the wide full-chest dimension provides for one-on-one fiberoptic coupling without demagnification. Demagnification has been the main cause of low optical coupling efficiency, with consequent insufficient image quality in previous (prototypic) generations of CCD–based systems for digital chest radiography (15,16). Second, the oblong detector shape limits the number of electronic components needed to operate the device. As a result, the power consumption is low enough to allow fan-cooled operation of the detector system. Moreover, the application of a small CCD detector reduces the costs.

MTF values for the CCD digital radiography system show good resolution capability in the scanning direction, as well as in the slot direction. Values of 0.70 at 1 cycle per millimeter and 0.35 at 2 cycles.
per millimeter for CCD digital radiography are greater than MTF values for conventional screen-film radiography. To compare, MTF values for images obtained with a 400-speed screen-film combination (Lanex regular screens and T-mat-G 65500 film; Eastman Kodak) is roughly 0.60 at 1 cycle per millimeter and 0.30 at 2 cycles per millimeter (4). Thus, the resolution capability of the chest CCD digital radiography system is at least comparable to that of screen-film radiography systems.

Findings from previous studies on the performance of digital flat-panel chest imaging systems versus conventional chest radiography systems have shown equal or better results with the digital detector systems (2,17). Aufrichtig (2) demonstrated that, compared with standard chest hard-copy images, unprocessed digital hard-copy images had improved contrast-detail detectability at a similar radiation dose. Chotas and Ravin (17) found better contrast-detail results with postprocessed digital hard-copy images than with conventional film images, with small signals of low inherent subject contrast levels detected more often on the digital images, even at reduced exposure levels. In the present study, CCD digital radiography results were also better than the conventional Bucky screen-film radiography results and were at least comparable to AMBER results. Results of studies in which digital radiography systems have been compared with AMBER systems for image quality have, as far as we know, not been published before.

In the present study, the specific attenuation and scatter characteristics of lungs and mediastinum were taken into account separately. This study design was chosen for optimal comparison between the imaging modalities. The new CCD digital radiography unit was compared not only with a standard conventional screen-film radiography (Bucky) unit but also with the AMBER system. This was done because the AMBER system was specially designed for improved visualization of the mediastinal area in the chest (18). Good performance of the CCD digital radiography system was also expected in the mediastinal area because of the relatively large dynamic range of the digital system. Accordingly, contrast-detail curves showed a lower threshold depth for CCD digital radiography than for Bucky screen-film radiography in the lung and mediastinum configuration.

For the mediastinum configuration, the difference was found not to be statistically significant, but this was probably because of the smaller number of paired data points (ie, six) that could be derived from the contrast-detail test object for Bucky screen-film radiography compared with CCD digital radiography for this configuration. This limited the statistical power of the Wilcoxon signed rank test. The result for AMBER hard-copy reading was better than the result for CCD digital radiography hard-copy reading in the mediastinum configuration. This can be explained by adjustment of the radiation dose during scanning with the AMBER system, which was doubled compared with the radiation dose of the CCD digital radiography system. This higher dose causes an improved signal-to-noise ratio for AMBER hard-copy images, with consequently better contrast-detail performance. The superior performance for AMBER hard-copy reading versus CCD digital radiography hard-copy reading, however, disappeared with CCD digital radiography soft-copy reading. CCD digital radiography soft-copy reading, which is the preferred medical reporting method in clinical practice, was found to be better than CCD digital radiography hard-copy reading. Also, CCD digital radiography soft-copy reading was found to be superior to AMBER hard-copy reading for the lung configuration. Moreover, no difference was found between the CCD digital radiography soft-copy reading and AMBER hard-copy reading techniques for the mediastinum configuration, despite the double entrance dose for AMBER with this configuration. This effect was expected because ΔC was significantly lower for hard-copy images obtained with CCD digital radiography than for hard-copy images obtained with AMBER.

The limitation of reduced contrast in digital hard copy disappears when reading digital soft copy. Accordingly, we have shown that the diagnostic potential improves with CCD digital radiography soft-copy reading.

In this study, the clinical protocols of the systems, which were established by radiologists, physicists, and manufacturers together, reflect common clinical practice. The use of matched entrance doses (and equal beam quality) instead of clinical protocols is not feasible because AMBER modulates the beam intensity in response to measurements from a detector located behind the film cassette. Moreover, the effective dose for CCD digital radiography (0.014 mSv) has been shown to be comparable to that of AMBER (0.016 mSv) in previous studies (19). Both modalities have a higher radiation exposure compared with Bucky screen-film radiography (0.009 mSv) (19). The purpose of AMBER is to provide optimal recording of the chest on hard-copy images without under- or overexposed areas. Establishing Bucky screen-film radiography with the same effective dose as AMBER would cause suboptimal imaging with overexposed areas.

During the implementation of CCD digital radiography in our hospital, the effective dose level that was chosen for CCD digital radiography was comparable to that of AMBER. In this way, we attempted to achieve appropriate image
quality that was comparable to that of AMBER. In later stages, dose optimization with CCD digital radiography will be investigated.

**Study Limitation**

The Bucky screen-film images in the mediastinum configuration were initially left out of the evaluation because of the inherently poor image quality; these images were later scored for comparison with CCD digital radiography and AMBER images. This may have resulted in a performance bias owing to learning effect, which would have been to the advantage of the Bucky screen-film images. The results for these images, however, remained poor compared with the results for CCD digital radiography and AMBER. Thus, the delayed evaluation of Bucky screen-film images in mediastinum configuration had no effect on the overall study outcome.

In this study, a number of systems were evaluated on a pair-wise basis to investigate the performance of CCD digital radiography compared with that of other systems. An omnibus test could have been used in this study to look at all comparisons. It is crucial to note, however, that the problem discussed in this article is for the specific comparison of CCD systems and other machines. Thus, a precise set of specific hypotheses tests is known and specified in advance. In this context, an omnibus test would be a distraction of the approach in this study. To correct for multiple testing in this study, a Bonferroni correction was applied.

We found lower contrast values for hard-copy images obtained with CCD digital radiography than for those obtained with AMBER and Bucky screen-film radiography. We used a clinically relevant window width and window level setting, as determined by the CCD digital radiography system. The window width, however, was rather wide (ie, moderately overestimated by the system) to ensure that all information would be inside the window. This resulted in moderately lower contrast. In daily clinical practice, the window width and window level setting will eventually be optimized by hand. Furthermore, the fact that 8-bit printing was used may have also limited optimal use of the lookup table for printing.

**Practical application:** To the best of our knowledge, the dedicated ThoraScan chest system used in the present study is the first digital full-field chest imaging system that is based on a CCD detector without optical demagnification. The contrast-detail study showed good imaging quality for the CCD digital radiography system when hard-copy images were used. The performance of the CCD digital radiography system was found to be superior to the performance of the conventional Bucky screen-film radiography system. For soft-copy reading, CCD digital radiography performance was comparable to that of state-of-the-art AMBER. Compared with conventional techniques, the CCD digital radiography system has potential for improved chest imaging.

**APPENDIX**

The ThoraScan chest imaging system is based on a CCD detector array and slot-scan technology. The x-ray beam is generated in a tungsten anode at a tube voltage of 60–150 kV. Because of the scanning properties of the system, the x-ray tube is water cooled to achieve the necessary heat dissipation. Images are acquired by scanning the patient’s chest. Scanning is performed by using a 1-cm-thick collimated fan-shaped x-ray beam in combination with a multilinear solid state scanning CCD detector. The detector moves synchronously with the fan-shaped x-ray beam. Because the combination of a collimated fan beam and a matching detector has an inherently high efficiency in rejecting scattered radiation, there is no need for an antiscatter grid. During the downward movement, a preliminary lower dose scan of the upper thorax is obtained to measure the optimum tube current at the preselected tube voltage. During the upward movement, the actual scanning is performed with the fixed tube current and tube voltage as set for the prescan.

The physical properties of the CCD detector system are presented in Table A1. The detector is an assembly of eight CCD chips that are implemented side by side. The CCD detector is sensitive at a width of 44.0 cm and at a height of 10.8 mm. The basic detector elements in the CCD measure 27 × 27 μm, but downsampling decreases the final image resolution. The standard resolution mode features a pixel size of 162 × 162 μm. For future implementation, the advantages of a high-resolution mode of 81 × 81 μm will be evaluated. The CCD chips are covered with a fiberoptic plate that protects the CCDs against the x-rays and optically couples the phosphor (scintillator) with the CCD. The scintillator layer that converts the x-rays into visible light is deposited on the fiberoptic plate and consists of thallium-doped cesium iodide.

The typical readout frequency is 500 μsec per line, resulting in image acquisition at 320 mm sec⁻¹. This results in a scanning time of 1.3 seconds in the standard resolution mode, with an effective (local) exposure time of 20 msec. In other words, local unsharpness is not expected because of the short local exposure time.

**Time-Delay Integration**

The CCD readout operates in time-delay integration mode. With this technique, a charge “packet” is transferred down a CCD column from cell to cell. At each CCD cell, the charge packet accumulates additional signal during a short time, τ. This process of accumulation and transfer is repeated until the charge packet is transferred down n detector rows (eg, with n equals 400, there are 400 27-μm detector elements along the 10.8-mm detector width). The CCD is also moved at a velocity equal to the average transfer rate of the charge packet but in the opposite direction. As a result, the charge packet accumulates signal originating from...
an approximately fixed point. Time-delay integration acquisition allows shorter scanning time and, thus, lower tube load than simple linear array acquisition. For time-delay integration, the total integration time is \( \tau \). To achieve the same signal level for a pixel with a linear array, the scanning time would have to be a factor \( n \) longer than would be required for a multilinear detector operating in time-delay integration mode (22).

**Image Processing**

Three basic types of corrections are applied to the raw image data, namely, direct current voltage offset, dark current correction, gain correction, and edge correction. With direct current voltage offset and dark current correction, a correction matrix is constructed that captures the characteristics of the detector columns with respect to voltage offset and thermally generated dark current in the CCDs. This compensation needs to be performed before every scan because the correction matrix depends on the internal temperature of the detector. After correction for dark current and direct current voltage offset, gain correction is needed to equalize the output from all columns. Gain correction is performed by taking samples from the detector output when exposed to a uniform level of radiation during the daily calibration procedure. Finally, edge correction is needed to compensate for the fact that the detector elements of the eight-tiled CCDs are not in perfect physical contact at the edges of the CCDs. The resulting seven gaps are filled by interpolation with signal values from adjacent pixels. For final postprocessing, the ThoraScan system uses multiscale image contrast amplification (MUSICA; Agfa-Gevaert, Mortsel, Belgium) (20).

**MTF Curve**

The MTF curve describes the resolution capability of an imaging system. It is the ratio of output contrast to input contrast of an imaging system. Each step in the imaging process can contribute to a loss of contrast (21). Because the MTF of a scanning system might not be symmetric, a separate analysis has to be carried out in the scanning and slot (row) directions. The MTF in the slot direction and the MTF in the scanning direction are both determined by using the pixel aperture, the thallium-doped cesium iodide scintillator, and the fiberoptic plate. The MTF in the slot direction may also be influenced by imperfect colinearity of the detector columns in the scanning direction and by the charge transfer process of the serial register. The MTF in the scanning direction degrades further as a result of cesium iodide decay characteristics, imperfect column transfer efficiency, and asynchronism between scanning velocity and charge transfer timing (3,4).

**Acknowledgments:** The authors gratefully acknowledge the participation of the following persons in the panel of observers: Harmine M. Zonderland, MD, PhD, Albert de Roos, MD, PhD, and Wouter Teeuwisse, BSc (Leiden University Medical Center, Department of Radiology). Herbert Barkhuysen, PhD (Nueltron-Oldelft), is gratefully acknowledged for providing technical data on the CCD digital radiography system.

**References**

Bile Leakage during Transjugular Intrahepatic Portosystemic Shunt Creation: In Vitro Effect of Bile on Growth and Function of Human Umbilical Vein Endothelium

**PURPOSE:** To evaluate the effect of bile on growth and proliferation of human umbilical vein endothelium cultured in vitro, with a view toward clarifying the effect of bile leakage during transjugular intrahepatic portosystemic shunt creation.

**MATERIALS AND METHODS:** This study was approved by the ethical review committee, and written informed consent was obtained from all mothers. Endothelial cells (ECs) were collected from human umbilical veins and cultured in vitro. After 24–48 hours in culture, ECs were distributed into groups supplemented with the following concentrations of bile in the culture medium: 0%, 5.0%, 10.0%, 15.0%, 20.0%, and 25.0%. The cells were harvested 5 days after supplementation with bile. The morphologic features, von Willebrand factor (vWF) level, tetrazolium salt (MTT) assay value of light absorption, total protein level, and nitric oxide synthase (NOS) activity of the ECs were evaluated.

**RESULTS:** All explanted cells were identified as ECs by using the vWF test. Compared with ECs in the control group without bile, ECs in culture medium with a bile concentration of 5.0%, 10.0%, or 15.0% showed no marked morphologic changes, whereas ECs in culture medium with a bile concentration of 20.0% or 25.0% were reduced greatly in number and looked markedly immature. The MTT value of light absorption, total protein level, and vWF secretion were significantly decreased (P < .05 for all) in ECs in culture medium with 25.0% bile compared with these parameters in ECs in culture medium without bile, although these parameters did not significantly differ between the ECs in culture medium of 5.0% or 10.0% bile and the ECs in culture medium without bile. Compared with NOS activity in ECs when no bile was present in the culture medium, NOS activity in ECs was significantly decreased at all bile concentrations (P < .05).

**CONCLUSION:** Low concentrations of bile do not markedly inhibit cell growth; the inhibiting effect of bile on ECs progresses with an increase in bile concentration.

Transjugular intrahepatic portosystemic shunt (TIPS) creation has been widely used for management of variceal bleeding and ascites secondary to portal hypertension. However, restenosis occurs frequently. Although several procedures, including thrombolysis, balloon redilation, or placement of another stent, can be performed, there are no truly reliable methods for preventing or treating restenosis. It seems impossible to solve this problem currently because we know little about the mechanisms of the restenotic process of TIPS. However, several important findings related to these mechanisms have been reported in
past years. One of the findings is that bile duct injury and bile leakage during TIPS creation may be associated with restenosis (1–5).

The histopathologic findings of TIPS restenosis in both humans and animals are similar to the histopathologic findings of the intravascular restenosis seen after stent placement, which primarily consist of pseudointimal proliferation (1–8). Pseudointimal proliferation tissue consists mainly of smooth muscle cells and extracellular collagenous matrix (1–8). Bile leakage and its effect on pseudointimal proliferation during TIPS creation have been the focus of study for many investigators (1,3,4). Bile duct injury and bile leakage—at least on a microscopic level—seem inevitable during a TIPS procedure owing to the disintegrated nature of hepatic tissue. Several investigators (1–3) have demonstrated in both humans and pigs that bile leakage during TIPS creation may promote pseudointimal proliferation; this idea is also supported by the improved TIPS patency observed when a covered stent is used (2,9–11). The mechanism by which use of the covered stent discourages restenosis seems to be the exclusion of bile leakage and, thereby, a decrease in pseudointimal proliferation. However, this mechanism is inconsistent with results of an in vitro and in vivo investigation (4), which revealed that bile seems to inhibit smooth muscle cell proliferation instead of promoting it. Thus, the purpose of our study was to evaluate the effect of bile on the growth and proliferation of human umbilical vein endothelial cells cultured in vitro, with a view toward clarifying the effect of bile leakage during TIPS creation.

MATERIALS AND METHODS

Culture of Endothelium

Isolation of endothelial cells.—This study, which involved the collection and use of umbilical cords, was approved by the ethical review committee of our institution, and written informed consent was obtained from all mothers. The isolation procedures for the endothelial cells used in this study were performed by one author (Q.L.) alone.

Twenty 20-cm-long umbilical cords were obtained within 6 hours after delivery of healthy babies by healthy mothers. The umbilical vein was cannulated and flushed with 1000 mL of phosphate-buffered saline; the cord was then filled with 0.1% Type IV collagenase (Worthington Biochemical, Lakewood, NJ) and clamped at the ends. Following incubation for 10 minutes at 37°C, the cord was unclamped and the fluid was collected by gently massaging the cord. The resulting fluid suspension containing endothelial cells was centrifuged at 1000 rotations per minute for 10 minutes, and the cell pellet was suspended in medium 1640 containing 0.2 mmol/mL L-glutamine (GIBCO BRL, Carlsbad, Calif), 100 U penicillin, 100 g/mL streptomycin (Nanjing Jinling Pharmaceutical, Nanjing, China), and 20% fetal bovine serum (Worthington Biochemical).

On the basis of different concentrations of bile administered to the culture medium, five experimental groups of endothelial cells in culture media with bile concentrations of 5.0%, 10.0%, 15.0%, 20.0%, and 25.0% and a control group of endothelial cells in culture medium without bile were created and subjected to morphologic observation, total protein assays, von Willebrand factor (vWF) measurement, and nitric oxide synthase (NOS) activity estimation; endothelial cells in an additional five bile concentrations of 2.5%, 5.0%, 7.5%, 12.5%, 17.5%, 20.0%, and 22.5% were used for tetrazolium salt (MTT) assays. Endothelial cell cultures were established by plating cells at a density of 1 × 10^4 in each well of a six-well plate (Denmark NUNC, Roskilde, Denmark). Twenty cultures were created for MTT assays, 10 were created for NOS studies, and 15 each were created for total protein assays and vWF measurement. Twenty-four to 48 hours after the endothelial cells were plated, the culture medium was replaced with an appropriate experimental growth medium containing different concentrations of bile as the experimental protocol. The medium was changed every 48 hours until the cells were harvested.

Preparation of bile.—Bile was obtained from a T tube after a cholecystectomy had been performed in a 77-year-old man with chronic cholecystitis secondary to cholelithiasis. The collection and use of bile were approved by the ethical review committee of our institution, and written informed consent was obtained from the patient. The collection of bile was performed in aseptic conditions 4 days after cholecystectomy while the patient had a normal body temperature. The bile was frozen immediately after collection and was kept frozen until use.

Medium preparation.—On the basis of results (not shown) from a pilot study, we had determined that bile at a concentration of 25.0% or higher is lethal to endothelial cells. For the experimental groups, endothelial cells were cultured in medium 1640 that was supplemented with 0.2 mmol/mL L-glutamine, 100 U penicillin, 100 g/mL streptomycin, and 20% fetal bovine serum and contained bile at different concentrations. The same culture conditions and culture medium—but no bile—were used for the control group cells.

Cell Assays

Morphologic observation.—The cultured cells were evaluated every day by using an inverted microscope (Olympus CK2; Olympus Opto-Technology, Nagano, Japan). The evaluation was performed by a pathologist with 30 years of experience who was blinded as to which specimens contained bile and which did not. Morphologic evaluation of endothelial cells included assessment of the following: their size and shape, the visibility of the nucleolus and organelles, and the clarity of cellular rims.

MTT assays.—After endothelial cells were seeded in a 96-well plate and after 24 hours in culture, various concentrations of bile (2.5%, 5.0%, 7.5%, 10.0%, 12.5%, 15.0%, 17.5%, 20.0%, 22.5%, and 25.0%) were added to the culture medium. Twenty cultures were created for each bile concentration level and for the control group without bile. After continuous culture for 5 days, 20 μL of MTT liquor (5 mg of MTT per milliliter of phosphate-buffered saline, Fluka, Buchs, Switzerland) was added to every well and the cells were incubated for another 4 hours. After the upper layer of the culture medium was removed, 20 μL of dimethylsulfoxide (Shanghai Bioengineering, Shanghai, China) was added to the medium and oscillated for 10 minutes. The 490-nm wavelength of light absorption was used for measuring the light absorption value of every well with a spectrophotometer (Spectra MAX 250; Molecular Devices, Sunnyvale, Calif).

Total protein assays.—After endothelial cells were seeded in a six-well plate and after 24 hours in culture, various concentrations of bile (5.0%, 10.0%, 15.0%, 20.0%, and 25.0%) were added to the culture medium. Fifteen cultures were created for each bile concentration level and for the control group without bile. After continuous culture for 5 days, the endothelial cells were harvested for protein assays. Total protein in the cultured endothelial cells was measured by using a bicinchoninic acid spectrophotometer assay (BCA Protein Assay Reagent; Pierce, Rockford, Ill). Sample results were compared with a standard curve generated from purified bovine serum albumin (Pierce).

Measurement of vWF.—After endothelial cells were seeded in a six-well plate and after 24 hours in culture, various concentrations of bile (5.0%, 10.0%, 15.0%, 20.0%, and 25.0%) were added to the culture medium. Twenty cultures were created for each bile concentration level and for the control group without bile. After continuous culture for 5 days, the endothelial cells were harvested for protein assays. Total protein in the cultured endothelial cells was measured by using a bicinchoninic acid spectrophotometer assay (BCA Protein Assay Reagent; Pierce, Rockford, Ill). Sample results were compared with a standard curve generated from purified bovine serum albumin (Pierce).
20.0%, and 25.0%) were added to the culture medium. Fifteen cultures were created for each bile concentration level and for the control group without bile. The cells were harvested after 5 days in culture. vWF was measured so that we could verify the presence and function of the endothelial cells according to their secretion of vWF. A vWF kit (Thrombosis Institute, Suzhou University, Suzhou, China) and a standard technique (12) were used.

Estimation of NOS activity.—After endothelial cells were seeded in a six-well plate and after 24 hours in culture, various concentrations of bile (5.0%, 10.0%, 15.0%, 20.0%, and 25.0%) were added to the culture medium. Ten cultures were created for each bile concentration level and for the control group without bile. The cells were harvested after another 5 days in culture, and NOS activity was then estimated by using a standard technique (13).

Statistical Analysis

Kruskal-Wallis tests were used for analysis of all the data, including the MTT value of light absorption, total protein content, vWF level, and NOS activity, and Bonferroni-Dunn t tests were employed for comparison between groups. A P value of .05 or less was considered to indicate a significant difference. All statistical analyses were performed by using SAS version 8.02 (SAS Institute, Cary, NC).

RESULTS

Morphologic Observation

Under an inverted microscope, normal attached endothelial cells had a flat spin-
intima in a 6½-month-old TIPS. These findings are similar to those of Saxon et al (3), who found that large biliary fistulas formed in seven of nine TIPS in a porcine model and in seven of eight human TIPS specimens in which parenchymal tract stenosis and occlusion were documented. Other experiments revealed the occurrence of bile leakage in TIPS porcine models, and there was a much higher incidence of stenosis or occlusion of the shunt in cases with bile leakage than in cases without bile leakage (3,4). Therefore, it seems a reasonable speculation that biliary tract injury and bile leakage during TIPS creation may be one of the factors involved in TIPS restenosis.

In our previous investigation involving a porcine model (4), we were unable to demonstrate significant differences in histologic quantity of pseudointimal proliferation between the group with bile leakage and the group without bile leakage. However, there were fewer cellular components in the group with bile leakage than in the group without bile leakage. Few endothelial cells covered the TIPS stents with bile leakage, while a good layer of endothelial cells covered the TIPS stents without bile leakage. Additionally, the results of our previous in vitro investigation (4) indicated that bile inhibits smooth muscle cell proliferation. Therefore, we suggest that the mechanism by which bile leakage causes TIPS restenosis might be associated with the promotion of thrombosis instead of smooth muscle cell proliferation. It is known that a defect in endothelial cell linings is associated with coagulation and thrombosis (16), and thus it seems logical that bile leakage may promote the thrombosis of TIPS stents by inhibiting endothelialization of the stent.

The results of this in vitro experiment indicate that bile inhibits endothelial cell growth in a concentration-dependent manner. Low concentrations of bile inhibit endothelial cell growth slightly, while the effect of bile on endothelial cell growth increases with an increase in bile concentration, until a level of bile (≥40%) that is lethal to the endothelial cells is reached. This morphologic result is in agreement with our finding that bile inhibits the NOS activity and vWF secretion of endothelial cells.

Injury of biliary tracts and bile leakage seem inevitable during TIPS creation. However, bile is supposed to be diluted immediately owing to high blood flow in the TIPS, and the bile concentration should be low. Bile leakage is verified mainly by the recognition of bile staining in areas of pseudointimal proliferation in the TIPS (1–5). In a case reported by LaBerge et al (14), a biliary tract was transected, but bile staining in the areas of proliferation in the TIPS and endothelial cell layer involvement were not mentioned. Therefore, it is possible that bile leakage may be invisible if the bile leak is too small to be seen.

Ducoin et al (17) found that more inflammatory cells and less collagen tissue were demonstrated in thicker pseudointimal proliferation tissue than in thinner pseudointimal tissue and that more inflammatory macrophages were found in pseudointimal tissue that showed bile staining. Therefore, it has been suggested that bile may play an important role in an inflammatory reaction that leads to pseudointimal proliferation in TIPS. Ducoin et al (17) agreed with the theory that bile promotes thrombosis owing to a persistent inflammatory reaction. Similar opinions have been voiced by Laberge and colleagues (14) and Sanyal et al (18). Our results suggest that more extensive bile leakage with visible bile staining may cause a severe inflammatory reaction that leads to marked thrombosis and that less extensive bile leakage without visible bile staining may cause slight inflammation that leads to late restenosis caused by pseudointimal proliferation.

vWF is a cytokine that plays an important role in platelet adhesion and accumulation. It can be used for identifying endothelial cells and estimating endothelial cell function. The results of this in vitro study demonstrated that a certain concentration of bile might inhibit endothelial cell growth and vWF secretion. These results agree with our previous experimental results and speculations in suggesting that bile may inhibit stent endothelialization and then promote thrombosis and restenosis of TIPS.

NOS, as an important message factor, has been widely mentioned in reports of restenosis after percutaneous transluminal angioplasty and stent placement (19). It is well known that NOS plays important roles in vessel healing after injury that include vessel dilatation, regulation of smooth muscle cell proliferation, and inhibition of platelet adhesion and thrombosis. TIPS restenosis is pathologically similar to vascular restenosis, but the
pseudointimal proliferation in TIPS restenosis seems much faster and more severe than that in vascular restenosis (3,4,6). Bile leakage is one of the major factors that is present in TIPS creation but not vascular stent placement. The results of NOS activity assessment in this in vitro experiment indicate that bile, by inhibiting NOS synthesis through a suppression of endothelial NOS activity and thus contributing to thrombosis of the shunt, may be involved in TIPS restenosis.

There were several limitations to this study: First, because the origination of endothelial cells in TIPS is uncertain, the endothelial cells cultured in this study, which were obtained by using umbilical veins instead of a TIPS restenosis model, may not exactly resemble the endothelial cells in restenotic TIPS. A further study involving the use of restenotic TIPS in a swine model is being pursued in our institution. Second, only a single individual’s bile was used in this study, and results may have varied if different bile compositions had been used. A further study involving multiple samples of bile from different patients is important for clarifying our results.

Another limitation was that it is almost impossible to know the true bile concentration in a TIPS environment. The concentrations of bile studied in this experiment seem much higher than those in a TIPS environment because bile is easily diluted by the blood flow in a TIPS. However, it is realistic to think that high concentrations of bile can occur momentarily in the TIPS environment, especially when the TIPS is dysfunctional. In addition, the results of colorimetric MTT assay may have been interfered with by the color of the bile, especially at higher concentrations.

In conclusion, results of this in vitro study show that low concentrations of bile do not markedly inhibit cell growth; the inhibiting effect of bile on endothelial cells progresses with an increase in bile concentration. Clinical application: The restenosis of TIPS seems to correlate with bile leakage during TIPS creation in both pigs and humans (1–5,14,15). However, in our previous in vitro and in vivo investigations (4,6), we observed that the proliferation of smooth muscle cells, the dominant cellular component of the restenotic tissue in TIPS, was not promoted by bile. The combination of data from the present study and our previous results suggests that bile leakage is an important factor in TIPS restenosis and that bile leakage contributes to restenosis apparently by inhibiting the growth and activity of endothelial cells rather than by promoting the proliferation of smooth muscle cells. Therefore, it is practical to reduce a TIPS restenosis by isolating a bile leak into the shunt by using a stent-graft.

Acknowledgments: The authors thank Yicheng Ni, MD, PhD, for his efforts in draft preparation and manuscript revision and Ping Song, MD, for her technical assistance with cell preparations and assays.

References

### Table 2

<table>
<thead>
<tr>
<th>Bile Concentration (%)</th>
<th>vWF Level (%)</th>
<th>Total Protein (mg/mL)</th>
<th>NOS Activity (u/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10 ± 2</td>
<td>0.32 ± 0.01</td>
<td>1.00 ± 0.21</td>
</tr>
<tr>
<td>5.0</td>
<td>12 ± 3</td>
<td>0.34 ± 0.01</td>
<td>0.57 ± 0.19</td>
</tr>
<tr>
<td>10.0</td>
<td>14 ± 3</td>
<td>0.36 ± 0.01</td>
<td>0.46 ± 0.19</td>
</tr>
<tr>
<td>15.0</td>
<td>11 ± 4</td>
<td>0.30 ± 0.01</td>
<td>0.35 ± 0.12</td>
</tr>
<tr>
<td>20.0</td>
<td>4 ± 5</td>
<td>0.20 ± 0.01</td>
<td>0.23 ± 0.06</td>
</tr>
<tr>
<td>25.0</td>
<td>1 ± 5</td>
<td>0.15 ± 0.01</td>
<td>ND (T)</td>
</tr>
</tbody>
</table>

Note.—Data are mean values ± standard deviations.
* Significantly different among all bile concentration groups ($t^2 = 27.213, P < .001$ [Kruskal-Wallis test]) and significantly lower in the group with 20.0% or 25.0% bile than in the group with no bile ($P < .05$ [Bonferroni-Dunn t test]).
† Significantly different among all bile concentration groups ($t^2 = 18.857, P = .002$ [Kruskal-Wallis test]) and significantly lower in the group with 25.0% bile than in the group with no bile ($P < .05$ [Bonferroni-Dunn t test]).
‡ Significantly different among all bile concentration groups ($t^2 = 11.667, P < .009$ [Kruskal-Wallis test]) and significantly lower in the groups with 5.0%, 10.0%, 15.0%, or 20.0% bile than in the group with no bile ($P < .05$ [Bonferroni-Dunn t test]).
§ ND = no data. This value was too low to detect.
Hepatic Lesions Deemed Too Small to Characterize at CT: Prevalence and Importance in Women with Breast Cancer

**PURPOSE:** To retrospectively evaluate the prevalence and clinical importance of hepatic lesions considered too small to characterize (TSTC) at initial computed tomography (CT) in women with breast cancer.

**MATERIALS AND METHODS:** Approval for this retrospective study was obtained from the institutional review board, which waived the requirement for informed consent. For each woman who received a diagnosis of breast cancer between 1998 and 2002, the authors reviewed the report of the first contrast material–enhanced CT examination that included assessment of the liver. For women with no definite liver metastasis and at least one hepatic lesion considered TSTC, reports of follow-up imaging examinations were reviewed for a change in lesion size; medical records and images were reviewed if there was a change in lesion size. The 95% confidence intervals (CIs) were calculated for best- and worst-case analyses of cases in which different assumptions were used to classify a lesion as benign.

**RESULTS:** Of 7692 women, 1012 (13.2%) underwent contrast-enhanced CT including liver assessment. The mean age of the 1012 women was 54.6 years (range, 20.7–89.1 years). The median time from diagnosis of breast cancer to initial CT examination was 14.1 weeks (range, −3.7 to 296 weeks). The presence of at least one hepatic lesion deemed TSTC was reported in 277 of 941 women (29.4%) in whom no definite hepatic metastasis was reported. Subsequent imaging examinations were performed in 191 of the 277 women (69.0%) (median time from initial CT to last follow-up imaging examination, 54 weeks; range, 0.3–302 weeks). Those examinations revealed the lesions were unchanged in 175 (91.6%) women, no longer visible in eight (4.2%), and larger in six (3.1%). In two women (1.0%), change could not be determined. The enlarging hepatic lesions deemed TSTC represented metastatic breast cancer (three patients), metastatic pancreatic cancer (one patient), or cysts (one patient); in one patient, the etiology was not known. Results of best- and worst-case analyses showed that the lesions were benign in 96.9% (95% CI: 93%, 99%) and 92.7% (95% CI: 88%, 96%) of women, respectively.

**CONCLUSION:** In 92.7%–96.9% of women with breast cancer and hepatic lesions deemed TSTC but no definite liver metastases at initial CT, the lesions represented a benign finding.

Both the prognosis and the treatment of patients with breast cancer are highly dependent on the presence or absence of visceral metastasis. Metastasis to the liver is observed infrequently in patients with breast cancer at initial presentation but portends a worse clinical outcome (1–3). Metastatic breast cancer to the liver contributes to a shorter median survival rate, which has been reported to be in the range of 3–14 months (4). Although the liver is the site of metastasis in 5%–20% of women with metastatic breast cancer (4), liver metastasis rarely occurs in women with ductal carcinoma in situ. For staging purposes, the National Comprehensive Cancer Network practice guidelines for patients with breast cancer recommend...

© RSNA, 2005
cancer (5) recommend performing abdominal computed tomography (CT), ultrasonography (US), or magnetic resonance (MR) imaging if results of liver function tests are abnormal, if the serum alkaline phosphatase level is elevated, and/or in patients with T3N1M0 disease; in stage IIa or IIb disease, these imaging tests are optional.

Hepatic lesions 1.5 cm in diameter or smaller are frequently difficult to characterize at CT and are often reported as being “too small to characterize” (TSTC) by the interpreting radiologist. With the use of multi-detector row helical CT and its attendant thinner collimation, increasingly smaller hepatic lesions can be detected, and, thus, a larger number of small lesions can be identified in patients. The clinical importance of these lesions often remains unknown until biopsy or other imaging examinations are performed for further characterization or until follow-up imaging is performed months later. Results of prior studies (6,7) have shown that most of these lesions are benign, even in patients known to have malignancy; nevertheless, it can be difficult to assume that a small hepatic lesion is benign in an individual patient at presentation. Given that breast cancer is one of the most common cancers, the number of women with hepatic lesions deemed TSTC is of substantial magnitude, and it would be useful to have some guidance in the management of those lesions. Thus, the purpose of our study was to retrospectively evaluate the prevalence and clinical importance of hepatic lesions considered TSTC at initial CT in women with breast cancer.

MATERIALS AND METHODS

Patients and CT Scanning

Approval for this retrospective study was obtained from our institutional review board, which waived the requirement for informed consent. This study complied with the Health Insurance Portability and Accountability Act.

The tumor registry at Memorial Sloan-Kettering Cancer Center, a dedicated tertiary cancer center, was searched electronically by personnel of the clinical information center to identify all women with breast cancer diagnosed between 1998 and 2002 who also underwent contrast material–enhanced CT, including examination of the liver, after the initial diagnosis of breast cancer. CT examinations performed within a month before the initial pathologic diagnosis of breast cancer were also included because such examinations are sometimes ordered and performed before breast biopsy when the mammographic appearance is highly suggestive of breast cancer. A member of the radiology department’s information technology group then excluded from this study those CT examinations that were performed outside our institution owing to inconsistent availability of the official reports for those examinations.

The initial and follow-up CT examinations were performed at our institution with various scanners (GE Medical Systems, Milwaukee, Wis) between 1998 and 2004, with our standard protocols tailored to each particular scanner. Before October 2000, some images were obtained with a conventional nonhelical scanner; most scans, however, were obtained with a helical scanner with one to 16 detector rows. Standard collimation was 10 mm for the nonhelical scanner, 7 mm for single–detector row scanners, 3.75 mm for four–detector row scanners, 2.5 mm for eight–detector row scanners, and 1.25 mm for 16–detector row scanners. The pitch used on helical scanners varied during the study period and with different scanners but ranged from 0.75 to 1.5. The standard section thickness for image viewing was 10 mm for nonhelical scanners, 7 mm for single–detector row scanners, and 7.5 mm for four–, eight–, and 16–detector row scanners. A dynamic power injection of 150 mL of nonionic intravenous contrast material was given at 2.5 mL/sec (or slower if mandated by suboptimal venous access). Time delay to scanning varied with the type of scanner used but was determined on the basis of the typical time to portal venous phase imaging; automated bolus tracking to initiate scanning was not employed. All scans were sent to our enterprise-wide PACS (picture archiving and communication system) for interpretation on PACS workstations.

Data Collection

A member of the radiology department’s information technology group determined the number of initial CT examinations that were performed with the nonhelical scanner by determining which had been performed before the scanner’s replacement date.

The electronic medical records were searched by personnel of the clinical information center for patient age, date of breast cancer diagnosis, type of breast cancer (as determined with pathologic examination), and stage at diagnosis. A research assistant in the radiology department calculated the mean patient age at diagnosis, as well as the median time from the diagnosis to the initial CT examination and from the initial CT examination to the last available follow-up imaging examination.

The indication for performing each initial CT, as stated in the radiology report, was recorded by one of two body imaging fellows (H.I.K., S.A.P.). Each CT report was categorized by one of those fellows as mentioning no liver metastasis, at least one liver metastasis, or findings indeterminate for liver metastasis. The report of each CT examination categorized as mentioning either no liver metastasis or findings indeterminate for liver metastasis was then classified according to the presence and number of small hepatic lesions considered TSTC. A hepatic lesion deemed TSTC is typically 1.5 cm in diameter or smaller (6–8) or is determined by the interpreting radiologist at our institution to be too small to allow accurate characterization on the available images. Although the majority of hepatic lesions deemed TSTC are hypodense or slightly hyperattenuating relative to surrounding liver parenchyma at CT, no distinction was made in this study between hypodense and hyperattenuating lesions. For each patient whose initial CT report mentioned at least one hepatic lesion deemed TSTC but no definitive liver metastasis, the official reports of subsequent CT, MR imaging, and US examinations were reviewed by one of two body imaging fellows (H.I.K., S.A.P.). Specifically, if the report of the last available follow-up imaging examination stated that the initial hepatic lesion(s) was unchanged in size, the report of at least one additional intervening follow-up imaging examination (if performed) was reviewed to assess for intervening change. If the report of the last available follow-up imaging examination indicated that the initial hepatic lesion(s) had changed in size or was no longer visible, the reports of all available intervening imaging examinations were reviewed.

On the basis of this review, each hepatic lesion deemed TSTC was categorized as showing no change, a decrease in size, or an increase in size. Indeterminate change was noted if the follow-up examination demonstrated at least one new hepatic lesion that obscured the initial lesion but that did not clearly arise from growth of the original TSTC lesion. For lesions that increased in size, one of two body imaging fellows (H.I.K., S.A.P.) and an attending radiologist (D.M.P.) deter-
Radiology

women with breast cancer and CT including liver (n = 1012)

Liver Metastasis Reported at Initial CT?

<table>
<thead>
<tr>
<th>Liver Metastasis Reported at Initial CT?</th>
<th># of Hepatic TSTC Lesions Reported</th>
<th># of Hepatic TSTC Lesions Reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes (n = 71)</td>
<td>0 (n = 628)</td>
<td>0 (n = 36)</td>
</tr>
<tr>
<td>no (n = 885)</td>
<td>1 (n = 131)</td>
<td>1 (n = 6)</td>
</tr>
<tr>
<td>indeterminate (n = 56)</td>
<td>&gt;1 (n = 126)</td>
<td>&gt;1 (n = 14)</td>
</tr>
</tbody>
</table>

Follow-up Imaging?

<table>
<thead>
<tr>
<th>Follow-up Imaging?</th>
<th>Lesion Size?</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes (n = 178)</td>
<td>not assessable (n = 1)</td>
</tr>
<tr>
<td>no (n = 79)</td>
<td>larger (n = 6)</td>
</tr>
<tr>
<td></td>
<td>smaller or not visible (n = 7)</td>
</tr>
<tr>
<td></td>
<td>unchanged (n = 164)</td>
</tr>
</tbody>
</table>

Biopsy?

<table>
<thead>
<tr>
<th>Biopsy?</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes (n = 1)</td>
</tr>
<tr>
<td>no (n = 5)</td>
</tr>
</tbody>
</table>

Figure 1. Flow diagram shows results of follow-up of hepatic lesions deemed TSTC in women with breast cancer.

Data Analysis

Best- and worst-case scenario analyses were performed by one author (D.M.P.) because of the lack of a reference standard for some of the lesions that grew slightly, were no longer visible at follow-up imaging, or could not be assessed at follow-up imaging due to interval appearance of multiple hepatic metastases.

For the best-case analysis, a hepatic lesion deemed TSTC was assumed to be benign if it (a) was unchanged in size at follow-up imaging, (b) was not visible at follow-up imaging but was considered benign on the basis of the clinical context by the treating physicians, (c) was shown to represent a benign lesion or no lesion at MR imaging, (d) enlarged slightly but remained of indeterminate etiology at follow-up, or (e) was indeterminate for change in size owing to the development of metastases that obscured the region of the original lesion.

For the worst-case analysis, a hepatic lesion deemed TSTC was assumed to be benign only if it was unchanged in size at follow-up imaging or was shown to represent a benign lesion or no lesion at MR imaging.

Statistical Analysis

A 95% confidence interval (CI) was calculated by a faculty statistician for the best- and worst-case analyses to help extrapolate the results from this study sample to the general population of women with breast cancer.

RESULTS

Of 7692 women who received a diagnosis of breast cancer at our institution between 1998 and 2002, 1012 (13.2%) underwent contrast-enhanced CT that included assessment of the liver at our institution. One hundred eighty-three CT examinations were performed before October 2000 with nonhelical scanners; 829 were obtained with helical scanners with one, four, eight, or 16 detector rows. The median time from diagnosis of breast cancer to initial CT examination was 14.1 weeks (range, −3.7 to 296 weeks). The indications for each CT examination, as stated in the “clinical statement” section of the official CT report, included breast cancer staging (n = 894), local recurrence of breast cancer (n = 1), symptoms (n = 61), and other reasons (n = 56). Of the 1012 women who underwent CT, the breast cancer at diagnosis (as determined with pathologic examination) was stage 0 in 33 women (3.3%), stage I in 238 (23.5%), stage II in 558 (55.1%), stage III in 118 (11.7%), and stage IV in 65 (6.4%). The mean age of the 1012 women was 54.6 years (range, 20.7–89.1 years). At histologic examination, breast cancer was classified as infiltrating ductal carcinoma in 772 women (76.3%), lobular carcinoma in 172 (17.0%), and non-invasive intraductal carcinoma in 27 (2.7%); 41 patients (4.1%) had other types of cancer.

Liver Lesions

Of the 1012 women with breast cancer who underwent CT including liver examination, 885 (87.5%) had no liver metastasis reported at initial CT, 71 (7.0%) had at least one liver metastasis, and 56 (5.5%) had CT findings reported as indeterminate for liver metastasis (Fig 1). Of the 941 women with no liver metastasis or findings indeterminate for liver metastasis, 664 (70.6%) had no hepatic lesions deemed TSTC, 137 (14.6%) had one lesion deemed TSTC, and 140 (14.9%) had more than one lesion deemed TSTC.

Of the 277 women with at least one hepatic lesion deemed TSTC but no definite liver metastasis reported at CT, 191 (69.0%) underwent at least one subsequent CT, MR imaging, or US examination that included liver assessment (median time from initial CT to last follow-up imaging examination, 54 weeks; range, 0.3–302 weeks). For those 191 women who underwent follow-up imaging, the hepatic lesions deemed TSTC were unchanged in 175 (91.6%), no longer visible in eight (4.2%), and larger in six patients (3.1%). In two other women (1.0%), the lesions could not be further categorized owing to the interval appearance of new hepatic...
lesions that obscured the initial lesions; it was not clear whether the original hepatic lesions grew or were simply obscured by new hepatic metastases.

**Enlarging Lesions**

In one of the six women with enlarging lesions, the initial CT examination showed one lesion measuring $6 \times 5$ mm; follow-up CT performed 9 months later revealed interval enlargement of this lesion to $15 \times 15$ mm and the appearance of new hepatic lesions compatible with metastases. CT-guided fine-needle aspiration of one of the hepatic lesions revealed metastasis from the patient’s known pancreatic cancer.

In a second patient, the initial CT examination showed one lesion measuring $6 \times 3$ mm; follow-up CT performed 6 weeks later revealed interval enlargement of this lesion to $16 \times 13$ mm and the appearance of multiple new liver lesions of variable size (Fig 2); all were presumed to represent progression of metastatic breast cancer. Two other patients each had two lesions deemed TSTC at initial CT; one of the lesions in one patient and both lesions in the other patient were enlarged at follow-up CT and showed typical features of metastasis.

In another patient, two hepatic lesions deemed TSTC at initial CT enlarged slightly at follow-up CT performed 11 months later (Fig 3); these lesions were shown to represent hepatic cysts at gadolinium-enhanced MR imaging of the liver.

In the final patient, the initial CT examination revealed two hepatic lesions deemed TSTC measuring $8 \times 6$ mm and $7 \times 6$ mm (Fig 4); the former increased to $11 \times 8$ mm and the latter remained stable at follow-up CT performed after 13.5 months. Both lesions remained of unclear etiology at subsequent follow-up imaging.

**Lesions That Were No Longer Visible**

Four of eight patients with hepatic lesions deemed TSTC at initial CT that were no longer visible at follow-up CT underwent chemotherapy during that interval. Three of these four women had no known extrahepatic metastases and were receiving adjuvant chemotherapy; the hepatic lesions in these three women were not considered to represent metastases by their treating physicians. The fourth patient had known osseous and nodal metastases; the hepatic lesion was considered of indeterminate etiology.

In a fifth patient, one of the hepatic lesions deemed TSTC at initial CT showed, in retrospect, a definite target sign appearance, which is a finding compatible with a metastasis. The patient also had pulmonary and pleural metastases at that time, for which she received chemotherapy.

In a sixth patient, the only available follow-up CT examination was performed without intravenous contrast material, and this limited the evaluation of the liver. In a seventh patient, the lesion deemed TSTC at the initial CT examination could not be identified in retrospect. The lesion deemed TSTC in an eighth patient was reported as a possible artifact, and no liver lesion was seen at MR imaging performed 1 week later.

**Analysis of Best- and Worst-Case Scenarios**

For the best-case analysis, hepatic lesions deemed TSTC in the 191 women...
who underwent follow-up imaging were assumed to be benign because they were unchanged in size at follow-up imaging (175 women), were shown to represent cysts (one woman with enlarging lesion) or no lesion (one woman) at subsequent MR imaging, were not visible at follow-up imaging (one woman), or were indeterminate for change in size due to the development of hepatic metastases that obscured the region of the original lesion (two women). In this best-case analysis, the hepatic lesions in 185 of the 191 women (96.9%; 95% CI: 93%, 99%) would be considered benign.

For the worst-case analysis, hepatic lesions deemed TSTC in the 191 women who underwent follow-up imaging were assumed to be benign because they remained unchanged in size at follow-up imaging (175 women) or were shown to represent cysts (one woman with enlarging lesions) or no lesion (one woman) at subsequent MR imaging. In this worst-case analysis, the hepatic lesions in 177 of the 191 women (92.7%; 95% CI: 88%, 96%) would be considered benign.

Extrahepatic Metastases

Of the 277 women who had at least one hepatic lesion deemed TSTC but no definite liver metastasis reported at initial CT, 146 had extrahepatic metastases. Lymph node metastasis was found at presentation in 130 of the 277 women (46.9%). The nodal metastases were in axillary lymph nodes in 122 of the 130 women (93.8%) and in nonaxillary (predominantly mediastinal and supraclavicular) lymph nodes in eight (6.2%). Extrahepatic nonnodal metastases reported at presentation were located in bone (n = 18), lung (n = 9), pleura (n = 5), chest wall (n = 2), and other (n = 2) sites; metastases involved more than one site in several women.

DISCUSSION

Small liver lesions are frequently seen at CT in women with breast cancer, and more will be seen as CT collimation gets increasingly thinner with dissemination of more advanced multi-detector row technology. In our study, 29.4% of women with breast cancer and no liver metastasis reported at initial CT were found to have at least one hepatic lesion deemed TSTC.

Jones et al (6) reported that 82% of 1454 outpatients in their study were referred for abdominal CT because of known extrahepatic malignant disease (from a wide range of primary tumors) or previous hepatic metastasis. Fifty-one percent of the patients known to have cancer had small (15 mm or smaller in diameter) hepatic lesions at CT that were benign, and 26% had malignant small lesions. In a study of 2978 patients with various types of cancers, Schwartz et al (7) found small hepatic lesions in 12.7% of patients; these lesions were benign (according to stability at follow-up CT) in 80.2% of patients and metastatic in 11.6%. In 92.7%–96.9% of the women without gross liver metastasis at presentation in our study, the hepatic lesions deemed TSTC at initial CT represented a benign finding; these worst- and best-case estimates were needed owing to the lack of a reference standard for some of the lesions, such as those that were not visible at follow-up imaging.

The percentage of hepatic lesions deemed TSTC that were benign was higher in our study than in other studies, probably because of the different criteria used for selecting patients and lesions; our study was restricted to women with breast cancer and to hepatic lesions deemed TSTC that were discovered only at initial presentation in women without definite liver metastasis. Given the relatively high prevalence of hepatic lesions deemed TSTC in patients with breast cancer and the relatively low reported prevalence of liver metastasis in such patients at initial presentation (1–3), it is not surprising that most hepatic lesions deemed TSTC are found to be benign.

Although metastatic disease is the greatest clinical concern when at least one hepatic lesion deemed TSTC is demonstrated at CT, those lesions also could be due to various other entities, such as cysts, biliary hamartomas, and hemangiomas. Hepatic cysts and hemangiomas are well-known, common benign entities that often can be readily characterized at imaging when they are of sufficient size and show typical imaging features. Biliary hamartomas are less well known to many radiologists but are common, benign malformations that consist of focal, disorderly collections of bile ducts. The bile ducts within these collections can be of variable size and subcapsular or parenchymal in location. Biliary hamartomas are more often multiple than single at imaging and typically are less than 5 mm in diameter. In a study of patients with various primary malignancies, all biliary hamartomas had lower attenuation than the surrounding normal liver parenchyma at contrast-enhanced CT (9). The spectrum of radiologic manifestations of biliary hamartomas in that study included one or two circumscribed lesions measuring 5–10 mm; multiple lesions measuring approximately 5 mm each; innumerable tiny, nearly uniform 2–5-mm lesions; and innumerable lesions measuring 2–15 mm. Biliary hamartoma is one benign etiology of liver lesions deemed...
TSTC that can mimic liver metastasis and should be considered in the differential diagnosis of such lesions.

Robinson et al (10), in a recent study of small hepatic lesions in patients known to have or suspected of having malignant disease, used CT to assess various lesion features, including size, shape, margin sharpness, homogeneity, and attenuation. These authors found that most isolated small liver lesions were benign, even in patients known to have or suspected of having cancer; with use of a model derived by using multiple logistic regression analysis, lesions smaller than 5 mm and showing a sharp margin were found to have only a 6% probability of being malignant.

Note that our study was designed to assess only those liver lesions that the interpreting radiologist deemed TSTC at initial CT; other small hepatic lesions that the radiologist could characterize would not have been included in this study. Subjective features of a lesion, such as margin sharpness, internal homogeneity, and relative attenuation, can influence the assessment of a specific lesion. Given that a lesion can be difficult to characterize accurately unless section thickness is less than half the lesion diameter, we believe that lesions up to 1.5 cm in diameter can be TSTC owing to partial volume effects with 7.5-mm-thick sections. Despite this relatively lenient size criterion for deeming a hepatic lesion to be TSTC, nearly all such lesions in our study turned out to be benign. The judgment as to whether a small hepatic lesion can be characterized is necessarily subjective.

In addition to immediate biopsy, several imaging options exist for the evaluation of hepatic lesions deemed TSTC at initial CT. Eberhardt et al (11) found that targeted liver US can help improve the detection and characterization of hepatic TSTC lesions when such lesions are specifically sought and larger than 0.5 cm. In that study, the authors assessed the utility of US evaluation of small hepatic lesions found at CT in patients with cancer and found that 66% of lesions were seen at US when they were specifically being sought and the specific CT finding was being referred to; only 32% of lesions were found when the CT finding was not specifically referenced. Forty-eight percent of the indeterminate lesions could be identified at US, and 93% of the US-detected small hepatic lesions could be further characterized as cysts, solid or metastatic lesions, or hemangiomas. The diagnosis made at US was confirmed in 93% of the lesions at follow-up hepatic US.

Similarly, MR imaging can help definitively characterize many small hepatic lesions as cysts or hemangiomas (12), particularly if heavily T2-weighted sequences (such as single shot or turbo fast spin echo) and gadolinium-based contrast material are used. Positron emission tomography is of limited value for characterizing lesions smaller than 1–1.5 cm (13,14). Even when a hepatic lesion deemed TSTC cannot be confidently characterized at imaging, interval follow-up CT, MR imaging, or US can be useful for assessing a change in lesion size; for all practical purposes, those small hepatic lesions that remain stable over time are assumed by treating physicians to be benign. Obviously, any new hepatic lesion that appears at follow-up imaging must be assumed to represent a metastasis until proved otherwise.

Clinical considerations in an individual patient help determine the pretest probability that a hepatic lesion deemed TSTC represents a metastasis. In a woman with ductal carcinoma in situ, for example, it is virtually certain that a hepatic lesion deemed TSTC does not represent metastasis from that breast lesion. The decision of imaging follow-up of a hepatic lesion deemed TSTC versus immediate further work-up must be tempered by the particular clinical scenario, including histologic type and initial stage of breast cancer, results of liver function tests, and the presence of symptoms or signs that might be due to hepatic metastasis.

Small liver lesions also often present a diagnostic challenge in patients with types of cancers other than breast cancer. For example, in patients with colorectal cancer, preoperative characterization of liver lesions is essential yet can be difficult as more hepatic lesions deemed TSTC do not represent metastasis from that breast lesion. The decision of imaging follow-up of a hepatic lesion deemed TSTC versus immediate further work-up must be tempered by the particular clinical scenario, including histologic type and initial stage of breast cancer, results of liver function tests, and the presence of symptoms or signs that might be due to hepatic metastasis.

Small liver lesions also often present a diagnostic challenge in patients with types of cancers other than breast cancer. For example, in patients with colorectal cancer, preoperative characterization of liver lesions is essential yet can be difficult as more hepatic lesions deemed TSTC do not represent metastasis from that breast lesion. The decision of imaging follow-up of a hepatic lesion deemed TSTC versus immediate further work-up must be tempered by the particular clinical scenario, including histologic type and initial stage of breast cancer, results of liver function tests, and the presence of symptoms or signs that might be due to hepatic metastasis.

In this retrospective study, we reviewed the official CT reports, not the actual CT scans, to determine the presence of hepatic lesions deemed TSTC. This study, however, was not designed to assess the ability of CT or the readers to depict, detect, or characterize small liver lesions; instead, it was undertaken to assess the prevalence and importance of hepatic lesions deemed TSTC in a large number of women with breast cancer in an actual clinical practice at a large, dedicated oncology center. In addition, our
Radiologists routinely report all hepatic lesions deemed TSTC, except those seen with widespread hepatic metastatic disease—particularly at the first CT examination that a patient undergoes at our institution. Because virtually all CT scans at our institution are obtained in patients with cancer, our radiologists are quite focused on, and adept at, detecting liver lesions, even if they are TSTC. As with all imaging findings, interobserver and intraobserver variability will exist in the detection and reporting of these hepatic lesions at CT; this study was not designed to assess that variability.

We believe that it was important to include in this study the hepatic lesion deemed TSTC in the one patient whose lesion was shown to represent metastasis from pancreatic cancer, even though it decreased the percentage of these lesions that were deemed benign. Our study was designed to determine the clinical importance of hepatic lesions deemed TSTC in women with breast cancer. Breast cancer does not preclude one from having other cancers, and this patient is a reminder of that important fact.

Although the CT scans were not obtained with a standardized technique because the available technology and protocols evolved rapidly during the study period, all were obtained in a single institution; most scans were obtained with multissection helical scanners. This study was not designed to assess the effect of using various types of CT scanner technology on the radiologist’s ability to detect or characterize hepatic lesions deemed TSTC. We used relatively thick sections (7.0–7.5 mm) that were not optimized for characterization of liver lesions; however, this is the range of section thicknesses used in our standard protocols designed to survey the body for metastasis. If thinner sections were used, it is likely that some of the lesions seen would have been characterized as cysts or hemangiomas and that additional, smaller hepatic lesions deemed TSTC would have been discovered. In addition, some of the hepatic lesions that were not visible at follow-up CT might have been detected had thinner sections been used. A recent study of multi–detector row helical CT revealed that image reconstruction at section thicknesses of 5.00, 3.75, and 2.50 mm yielded statistically significant, progressively higher sensitivity in the detection of liver lesions in general; however, a section thickness of less than 5 mm did not improve sensitivity in the detection of liver metastases 1.5 cm or smaller. The results of that study cannot be compared directly with our results because we used thicker sections to view our scans.

Although hepatic lesions deemed TSTC are common at initial CT performed in women with breast cancer, in our study these lesions were of benign etiology in 92.7%–96.9% of the women in whom no definite liver metastasis was evident. As increasingly thin sections are obtained at multi–detector row helical CT, a larger number of benign hepatic lesions deemed TSTC will be discovered. On the basis of the results of this study, at our institution we now recommend that interval follow-up imaging, rather than immediate imaging work-up or biopsy, be performed when at least one hepatic lesion deemed TSTC is discovered at CT in a woman with newly diagnosed breast cancer—particularly if the woman has stage 0 breast cancer and does not have other clinical findings suspicious for liver metastasis.

Acknowledgments: We thank Chinuyere Onyebuchi, MPH, and Jay Brijlal, BS, for assistance in data handling and analysis and Chaya Moskovitz, PhD, for statistical advice and calculations.

References
Acute Appendicitis: Added Diagnostic Value of Coronal Reformations from Isotropic Voxels at Multi–Detector Row CT

**PURPOSE:** To assess retrospectively the added value of coronal reformations from isotropic voxels obtained with 16-section multi–detector row computed tomography (CT) of the abdomen and pelvis in patients with suspected acute appendicitis.

**MATERIALS AND METHODS:** This study was approved by the institutional review board, and informed consent was waived. One hundred consecutive patients (21 men, 79 women; mean age, 38 years) with suspected appendicitis underwent 16-section multi–detector row CT (section thickness, 0.625 mm; pitch, 1.75; table speed, 35 mm/sec [17.5 mm per rotation, two rotations]; and gantry speed, 0.5 second per rotation), with coronal reformations. Twenty-four patients had appendicitis; 76 did not. Protocol included 150 mL oral iopamidol administered at 3 mL/sec. Transverse scans were reconstructed with 5-mm-thick sections at 5-mm intervals and 0.625-mm-thick sections at 0.625-mm intervals. The second data set was reformatted coronally, with 3-mm-thick sections at 5-mm intervals. Three independent blinded readers interpreted transverse scans alone and then coronal scans; confidence in visualization of any portion of appendix, entire appendix, wall thickening, distention, inflammation, fluid, and appendicitis was scored with 1–5 scale. Sensitivity and specificity were determined for each reader and compared by means of signed rank test. Agreement between readers was determined with \( \kappa \) statistic. Differences in mean confidence ratings for each finding were determined with Wilcoxon signed rank test.

**RESULTS:** Mean sensitivity and specificity for all three readers together were 96% and 95% for transverse reformations alone and 95% and 94% for combined transverse and coronal reformations (not significant), respectively. Visualization rates for portion or all of appendix were higher for combined transverse and coronal reformations than for transverse reformations alone (higher mean confidence scores: 0.23 higher \( \left[ P < .009 \right] \) and 0.51 higher \( \left[ P < .001 \right] \), respectively). In patients without appendicitis, transverse and coronal reformations together enhanced confidence in exclusion of wall thickening, distention, and fluid (lower confidence scores: 0.21 lower \( \left[ P < .001 \right] \), 0.17 lower \( \left[ P < .01 \right] \), 1.00 lower \( \left[ P < .001 \right] \), respectively). Combined transverse and coronal reformations enhanced confidence in identification of appendix in mean of 57 patients. Combined transverse and coronal scans helped exclude appendicitis in mean of 38 patients and aided diagnosis of it in 15.

**CONCLUSION:** Sixteen-section multi–detector row CT transverse and coronal reformations are equally sensitive and specific for diagnosis of appendicitis. Coronal reformations improve confidence in visualization of appendix (whether diseased or normal) and in diagnosis or exclusion of appendicitis.

Computed tomography (CT) is increasingly used in the evaluation of patients who are suspected of having acute appendicitis (1–5). With single– or multi–detector row technology, CT has reported sensitivity values of 80%–100%, specificity values of 91%–99%,
positive predictive values of 95%–97%, and accuracy values of 94%–100% in the evaluation of patients with suspected appendicitis (2–10). However, CT is less accurate in the examination of patients with equivocal clinical and physical findings of acute appendicitis. Further, the CT diagnosis of acute appendicitis may be uncertain or difficult in certain subsets of patients. For example, it may be difficult to identify the appendix in patients with scant intraperitoneal fat tissue, suboptimal opacification of the terminal ileum, a retrocecal appendix, or an appendix located adjacent to the adnexa. In such patients, viewing of the scans in a plane other than the transverse plane, such as the coronal or a curved plane, may enhance confidence in the diagnosis or exclusion of acute appendicitis.

With 16-section multi–detector row CT, it is now possible to scan the entire abdomen and pelvis within a single and comfortable breath hold at a resolution less than 1 mm in the x-axis, y-axis, and z-axis (11,12). These data sets result in voxels that are both less than 1 mm and nearly isotropic, which suggests that reformations in any desired plane will be similar in spatial resolution to those in the transverse plane (13). Thus, the purpose of our study was to assess retrospectively the added value of coronal reformations from isotropic voxels obtained with 16-section multi–detector row CT of the abdomen and pelvis in patients suspected of having acute appendicitis.

MATERIALS AND METHODS

Patients and Diagnosis

This study was approved by the institutional review board of our medical center, which waived informed consent. The study was compliant with the Health Insurance Portability and Accountability Act. One author (R.C.N.) is a consultant for GE Healthcare.

From June 1, 2003, to February 1, 2004, a total of 100 consecutive patients who were suspected of having acute appendicitis underwent 16-section multi–detector row CT to rule out acute appendicitis. The patients included 21 men and 79 women, and the mean age was 38 years (range, 19–80 years). In 100 patients, acute appendicitis was confirmed in 24, and exclusion was confirmed in 76.

The medical records, surgical reports, and pathologic reports were reviewed by one author (I.P.H.) to determine the diagnosis. Patients were considered to have acute appendicitis if they underwent appendectomy and the pathologic specimens confirmed acute appendicitis (n = 24). Patients were considered not to have acute appendicitis if an alternative diagnosis was established (n = 33) and treated or if a normal appendix was removed at appendectomy. The alternative diagnoses in these patients included pelvic inflammatory disease or tubo-ovarian abscess (n = 11), colitis (n = 8), cholecystitis (n = 2), ureteral stone (n = 2), enteritis (n = 2), ileus (n = 2), pylonephritis (n = 2), renal infarct (n = 1), mesenteric adenitis (n = 1), cystitis (n = 1), and pancreatitis (n = 1). One patient with enteritis had a normal appendix that was removed at surgery. Patients were also considered not to have acute appendicitis if they did not undergo appendectomy and had no evidence of persistent pain, abscess, or unexplained fever during follow-up (n = 43). In the patients without acute appendicitis, the mean follow-up interval was 9 months (range, 4–11 months).

Scanning

Scanning was performed from the dome of the diaphragm through the pubic symphysis with a CT scanner (Lightspeed 16; GE Medical Systems, Milwaukee, Wis). Patients ingested 450 mL of a 2% barium sulfate suspension (Readi-Cat 2; E-Z-Em, Westbury, NY) 1–2 hours before scanning. Iopamidol (Isovue; Bracco Diagnostics, Princeton, NJ) was injected (Empower CT; E-Z-Em) at a dose of 150 mL (300 mg of iodine per milliliter) and a rate of 3 mL/sec. Scanning was performed during the portal venous phase as determined with bolus tracking and automated triggering technology. The protocol was as follows: 140 kVp; 350 mA; sections, 16; section thickness, 0.625 mm; pitch, 1.75; table speed, 35 mm/sec (17.5 mm per rotation with two rotations); and gantry speed, 0.5 second per rotation. The transverse section data were reconstructed twice: first with 5-mm-thick sections at 5-mm intervals in the transverse plane and then with 0.625-mm-thick sections at 0.625-mm intervals. The second set of reconstructed transverse scans were then reformatted in the coronal plane with 3-mm sections at 5-mm intervals.

Reconstruction was performed with a commercially available console system devoted to rapid reconstruction (Xtream; GE Medical Systems); the system consists of dual 2.66-GHz processors (Xenon; Intel, Santa Clara, Calif) with a CT scan generator capable of reconstructing six to 10 scans per second. The scan generator required approximately 2 minutes to reconstruct both transverse and coronal scans. The entire process was performed by the technologist at the operator’s console. The 5-mm transverse and 3-mm coronal scans were transferred to a picture archiving and communication system workstation (Centricity 1.0; GE Medical Systems) as a separate series of scans for interpretation.

Scan Evaluation

After patient identifiers were removed from the CT scans, they were loaded onto a workstation (Advantage Windows; GE Medical Systems) for review. Three readers with subspecialty training in abdominal imaging (P.A.H., T.A.J., E.K.P.) served as independent readers who were blinded to the diagnosis; they had 1, 3, and 14 years of experience dedicated to abdominal imaging, respectively. Readers first assigned confidence scores to the transverse scans alone and then immediately assigned scores to the coronal scans obtained in the same patient. Impressions from the transverse scans were therefore fresh in the readers’ minds. Scores of 4 or 5 were considered affirmative. For both the transverse and coronal data sets, readers rated scans for findings indicative of acute appendicitis, and these findings included identification of a portion of the appendix on at least one scan, identification of the entire length of the appendix, distention (diameter, ≥8 mm), wall thickening, periappendiceal inflammatory change, and periappendiceal fluid.

A confidence score was obtained for each finding, with a scale of 1–5 (1, definitely absent; 2, probably absent; 3, cannot determine; 4, probably present; and 5, definitely present). In addition, on the basis of all the findings from each set of scans, readers judged whether acute appendicitis was present (with the same scale of 1–5 as was used for the confidence score), and thus, confidence scores were provided for acute appendicitis. The judgment regarding acute appendicitis was based on the individual reader’s assessment. After assigning scores to the combined transverse and coronal scans, readers judged whether the coronal scans added value to the transverse scans for identification of the appendix (whether diseased or normal), the exclusion of appendicitis, or the diagnosis of acute appendicitis. This judgment was subjective.

Statistical Analysis

The sensitivity and specificity values of each reader were determined for both the transverse scans alone and the transverse and coronal scans combined. P values for
TABLE 1
Mean Sensitivity and Specificity Values for Diagnosis of Acute Appendicitis

<table>
<thead>
<tr>
<th>Reader</th>
<th>Transverse Scans Alone</th>
<th>Combined Transverse and Coronal Scans</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity (%)</td>
<td>Specificity (%)</td>
</tr>
<tr>
<td>1</td>
<td>96 (23/24)</td>
<td>96 (73/76)</td>
</tr>
<tr>
<td>2</td>
<td>100 (24/24)</td>
<td>92 (70/76)</td>
</tr>
<tr>
<td>3</td>
<td>92 (22/24)</td>
<td>96 (73/76)</td>
</tr>
</tbody>
</table>

Note.—There were no significant differences in sensitivity and specificity between transverse scans alone and combined transverse and coronal scans for any reader. Data in parentheses were used to calculate percentages.

TABLE 2
Agreement among Readers for Diagnosis of Acute Appendicitis

<table>
<thead>
<tr>
<th>Reader Combinations</th>
<th>Transverse Scans Alone</th>
<th>Combined Transverse and Coronal Scans</th>
</tr>
</thead>
<tbody>
<tr>
<td>Readers 1 and 2</td>
<td>0.69</td>
<td>0.84</td>
</tr>
<tr>
<td>Readers 2 and 3</td>
<td>0.65</td>
<td>0.81</td>
</tr>
<tr>
<td>Readers 1 and 3</td>
<td>0.73</td>
<td>0.84</td>
</tr>
</tbody>
</table>

Note.—Numbers are the mean weighted $\kappa$ statistic. The differences between image sets were statistically significant ($P < .001$).

comparisons of the mean sensitivity and specificity for all three readers between transverse and combined transverse and coronal scans were computed with the signed rank test. Differences with $P$ values of .01 or less were considered statistically significant.

Agreement between readers for the diagnosis of acute appendicitis on the basis of transverse scans alone and on the basis of combined transverse and coronal scans was determined with the weighted $\kappa$ statistic. Differences in the reader mean weighted $\kappa$ values between the transverse scans alone and the combined transverse and coronal scans were assessed with the jackknife method. The mean confidence rating for each finding of acute appendicitis was classified according to reader, diagnosis, and scan set. Differences in ratings were determined with the Wilcoxon signed rank test. Statistical software (SAS, version 8.2; SAS Institute, Cary, NC) was used.

RESULTS

Diagnosis of Acute Appendicitis

For the diagnosis of acute appendicitis, there was no significant difference in sensitivity or specificity between transverse scans alone and combined transverse and coronal scans for any reader (Fig 1) (Table 1).

Reader Agreement

For the transverse scans alone, there was good agreement ($\kappa$ range, 0.69–0.73) among all three readers (Table 2). For the combined transverse and coronal scans, the $\kappa$ statistic (range, 0.81–0.84) was even higher, and this finding indicated a significantly higher level of agreement for the combined transverse and coronal scans than for the transverse scans alone ($P < .001$).

Confidence Scores for Imaging Findings and Diagnosis of Appendicitis

In patients with acute appendicitis, mean confidence scores for the individual imaging findings associated with acute appendicitis and the overall diagnosis of acute appendicitis are shown in Table 3. Confidence scores were significantly higher for the combined transverse and coronal scans than they were for the transverse scans alone for identification of a portion of the appendix or of the entire appendix (Fig 2), as well as for wall thickening, periappendiceal inflammation, and the overall diagnosis of acute appendicitis. Such differences in confidence scores were not observed, however, for the presence of distention and periappendiceal fluid.

For patients without acute appendicitis, Table 4 shows mean confidence scores for the individual findings and the diagnosis of acute appendicitis for transverse scans alone and combined transverse and coronal scans. For identification of a portion of or the entire length of the appendix, mean confidence scores were higher for combined transverse and coronal scans than they were for transverse scans, and this finding indicated that, even in patients without appendicitis, coronal scans add confidence to the identification of the appendix (Fig 3). For the combined transverse and coronal scans, confidence scores for distention, wall thickening, inflammation, fluid, and overall diagnosis of acute appendicitis were lower than they were for the transverse scans alone. This indicates that the readers were more confident in the absence of findings of acute appendicitis on the combined transverse and coronal scans than they were in the absence of those on the transverse scans alone.

Added Value of Coronal Scans

Table 5 shows the readers’ subjective impressions of whether the coronal scans added value to the transverse scans alone for the identification of the appendix (diseased or normal) and the diagnosis or exclusion of appendicitis. This table reports the values for each reader and the means for all three readers. The coronal scans added value to the diagnosis of acute appendicitis compared with the diagnosis with transverse scans alone in approximately half the patients with that diagnosis, and they added value to the exclusion of acute appendicitis compared with the exclusion with transverse scans alone in approximately half the patients without the diagnosis (Fig 4).

DISCUSSION

Although radiologists are skilled in the interpretation of CT scans in the transverse plane, the development of 16-section multi-detector row CT isotropic data sets coupled with fast reconstruction hardware and software has stimulated interest in viewing the abdomen in planes other than the transverse plane. For example, Wong et al (14) used four-section multi-detector row CT with multiplanar reformations to evaluate the liver and focused on the relationship of masses to portal veins, hepatic veins, bile ducts, and the vena cava. Caoli and Paulson (15) and Furukawa et al (16) found multiplanar reformations helpful in the evaluation of small-bowel obstruction, particularly for identification of the point of transition from dilated to decompressed bowel. Similarly, coronal reformations from scans obtained along the length of the pancreatic duct and common bile duct have been shown to be a useful ad-
We explored whether coronal reformations would serve as a useful adjunct to transverse scans in the diagnosis of acute appendicitis. To our knowledge, the role of coronal reformations obtained with multi-detector row CT in the evaluation of acute appendicitis has not been previously reported. Our results indicate that a scan obtained in the coronal plane serves as a useful adjunct to a scan obtained in the transverse plane in patients who are suspected of having acute appendicitis. In this study, the coronal scans were interpreted immediately after the transverse scans were interpreted.

The sensitivity and specificity for the combined transverse and coronal scans were similar to those for the transverse scans alone and ranged from 92% to 96%...
for sensitivity and from 93% to 95% for specificity. The sensitivity and specificity reported in recent reports that indicate that multi-detector row CT performed with both intravenous and oral contrast material is an accurate examination in patients who are suspected of having acute appendicitis (2–10,19).

The value of the coronal reformations is apparent in measures of agreement and diagnostic confidence among independent observers. Specifically, the agreement for the diagnosis of acute appendicitis among the three readers was consistently higher for the interpretations of combined transverse and coronal scans than it was for those of transverse scans alone.

In the patients with acute appendicitis, the confidence scores of the three readers for the diagnosis of acute appendicitis were higher for the combined transverse and coronal scans than they were for the transverse scans alone. Similarly, in the patients without acute appendicitis, confidence scores for the findings indicative of acute appendicitis were lower for combined transverse and coronal scans; these results indicated that the readers were more confident in the absence of findings on the coronal scans.

In patients with or without appendicitis, confidence in identification of a portion of or the entire length of the appendix was higher with the combined transverse and coronal scans than it was with the transverse scans alone. Indeed, one of the primary tasks in the evaluation of the patient with suspected acute appendicitis is the identification of the appendix. Once it is found, the next step is to search for evidence of acute appendicitis or to confirm a normal appendix. Our study findings indicate that the combination of transverse and coronal scans enhances confidence in identification of the appendix, whether diseased or normal. Enhanced confidence in appendiceal identification has implications for the reporting of the findings of studies. Specifically, a report with findings that indicate that the appendix has been identified and is normal is more definitive than one that indicates that, although there was no evidence of acute appendicitis, the appendix itself was not clearly identified. The subjective impression of all three readers was that the coronal scans added value to the transverse scans for the identification of the appendix (whether diseased or normal), the diagnosis of acute appendicitis, and the exclusion of acute appendicitis.

Although multiplanar reformations of scans of the abdomen have been described for diseases in the liver, the urinary tract, the small bowel, and the pancreas, little has been written in regard to the role of multiplanar reformations in the setting of acute appendicitis (14–17,19). We restricted our reformations to the coronal plane because the processing was seamless, was semiautomatic, and required only a few strokes of the keyboard at the operator’s console by the CT technologist. In addition, the coronal view, which is analogous to a frontal view of an abdominal radiograph, may be more intuitive for surgeons and radiologists. Our protocol required no imaging manipulation by radiologists, either at the operator’s console or at a dedicated three-dimensional workstation. In our clinical practice, the reformations are sent directly to the picture archiving and communication system and appear as a separate series for interpretation.

### TABLE 3
Mean Confidence Scores for Three Readers in Patients with Acute Appendicitis

<table>
<thead>
<tr>
<th>Portion of Appendix</th>
<th>Entire Appendix</th>
<th>Distention</th>
<th>Wall Thickening</th>
<th>Inflammation</th>
<th>Fluid</th>
<th>Appendicitis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transverse alone</td>
<td>4.71</td>
<td>4.25</td>
<td>4.65</td>
<td>4.57</td>
<td>4.47</td>
<td>2.81</td>
</tr>
<tr>
<td>Combined transverse and coronal</td>
<td>4.94</td>
<td>4.76</td>
<td>4.56</td>
<td>4.61</td>
<td>4.75</td>
<td>2.69</td>
</tr>
<tr>
<td>( P ) value*</td>
<td>&lt;.008</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

* Wilcoxon signed rank test.

### TABLE 4
Mean Confidence Scores for Three Readers in Patients without Acute Appendicitis

<table>
<thead>
<tr>
<th>Portion of Appendix</th>
<th>Entire Appendix</th>
<th>Distention</th>
<th>Wall Thickening</th>
<th>Inflammation</th>
<th>Fluid</th>
<th>Appendicitis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transverse alone</td>
<td>4.25</td>
<td>3.28</td>
<td>1.45</td>
<td>1.59</td>
<td>1.61</td>
<td>1.32</td>
</tr>
<tr>
<td>Combined transverse and coronal</td>
<td>4.50</td>
<td>3.99</td>
<td>1.28</td>
<td>1.38</td>
<td>1.42</td>
<td>1.22</td>
</tr>
<tr>
<td>( P ) value*</td>
<td>.01</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

* Wilcoxon signed rank test.

### TABLE 5
Added Value of Coronal Scans

<table>
<thead>
<tr>
<th>Patients with Appendicitis (n = 24)</th>
<th>Patients without Appendicitis (n = 76)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification</td>
<td>Diagnosis</td>
</tr>
<tr>
<td>Reader</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>8 (33)</td>
</tr>
<tr>
<td>2</td>
<td>12 (50)</td>
</tr>
<tr>
<td>3</td>
<td>10 (42)</td>
</tr>
<tr>
<td>All*</td>
<td>10 (42)</td>
</tr>
</tbody>
</table>

Note.—Data are numbers of patients, with percentages in parentheses. Confidence scores of 4 or 5 (on a scale of 1–5) were considered affirmative. Values were based on a response of 4 or 5 for each query.

* Numbers are means, with percentages in parentheses.
though our study demonstrates the value of coronal reformations as an adjunct to transverse scans, reformations in the sagittal plane may be similarly helpful.

Even with coronal reformations, in some cases it is difficult to identify the entire length of the appendix. In this situation, a curved reformation along the length of the entire appendix may prove useful. Such a reformation, which is analogous to a curved reformation from a scan obtained along the pancreatic duct or ureter, would "straighten out" the tortuous appendix and allow visualization of it in entirety (17,18). Our rationale for not performing such reformations routinely is that radiologists would be required to spend additional time at the workstation.

Prior to the development of 16-section multi-detector row CT, it was difficult to scan the entire abdomen and pelvis at a section thickness of less than 1 mm during a single comfortable breath hold.

Figure 3. CT scans in 29-year-old woman with sharp, crampy, abdominal pain and no fever. (a) Transverse CT scan with intravenous and oral contrast material shows tubular structure (arrows) opacified with oral contrast medium medial to cecum. One reader was uncertain whether that structure on this scan represented the appendix. (b) CT coronal reformation shows portion of appendix (arrows) arising from medial aspect of cecum. Appendix is filled with oral contrast material, with no distention, wall thickening, or evidence of surrounding inflammation. Coronal scan enhanced confidence in identification of appendix. This patient was treated for presumed pelvic inflammatory disease and recovered uneventfully.

Figure 4. CT scans obtained 1 day after hysterosalpingography in 32-year-old woman with pain in right lower quadrant of the abdomen, fever, and leukocytosis. (a) Transverse CT scan with oral and intravenous contrast material shows tiny contrast medium-filled structure (arrow) medial to cecum, which may represent appendix. Fluid-filled tubular structure (arrowheads) adjacent to right pelvic side wall, which likely represents adnexal structure rather than appendix. (b) Coronal scan shows normal appendix (arrow) arising medial to cecum. Inferior to normal appendix is a fluid-filled tubular structure (arrowheads) in pelvis with adjacent inflammatory changes. Reader confidence in appendiceal identification and exclusion of appendicitis was higher for the combined transverse and coronal scans than it was for the transverse scan alone. At surgery, a pyosalpinx and tubo-ovarian abscess were drained, and appendix was considered grossly normal.
This study had some limitations. Our intention was to demonstrate the value of coronal reformations as an adjunct to transverse scans alone. Accordingly, the coronal scans were interpreted after interpretation of the transverse scans, and impressions from the transverse scans were fresh in the minds of the interpreters. Such a design may have biased the results in favor of the subsequent interpretation, which in this case was based on the findings on the coronal scans. It was not our intent, however, to assess the value of the coronal reformations as stand-alone images, as we believe that in most clinical settings radiologists would be unlikely to abandon transverse scans in favor of coronal scans alone.

Another potential limitation was that we used a transverse diameter of 8 mm or more to identify appendiceal distention. While an 8-mm diameter is an optimal threshold measurement for distinguishing a normal from an abnormal appendix, a normal appendix filled with gas or contrast material may have a diameter as large as 10 mm (21,22). In this study, specific luminal contents were not considered. Nevertheless, readers identified diseased and normal appendices with a high level of sensitivity and specificity.

In conclusion, we found that combined transverse and coronal reformations from isotropic voxels obtained with 16-section multi–detector row CT rather than transverse scans alone are useful in the evaluation of patients who are suspected of having acute appendicitis. The addition of axial reformations adds confidence in the identification of the appendix and in the diagnosis or exclusion of acute appendicitis.

References
12. Horgan KM, Fishman EK. The current status of multidetector row CT imaging and communication system as a series for practice, the 5-mm transverse and coronal plane, with 3-mm intervals, which reconstructed transverse scans were reformatted in the coronal plane, with 3-mm sections at 5-mm intervals. For a typical patient, this series resulted in approximately 70–100 scans. For the second set of reconstructions, the scans with 0.625-mm thickness were reconstructed at 0.625-mm intervals. This series resulted in approximately 560–800 scans (13), but the scans in this series were not used for diagnostic purposes because of the large number of scans and because noise caused problems that were due to the section thickness of only 0.625 mm. Rather, the second set of reconstructed transverse scans were reformatted in the coronal plane, with 3-mm sections at 5-mm intervals, which resulted in 50–75 scans per patient. In our practice, the 5-mm transverse and coronal scans are sent to the picture archiving and communication system as a series for interpretation. The 0.625-mm transverse scans are archived but are not interpreted or sent to the picture archiving and communication system.

The CT dose index for the 16-section multi–detector row CT protocol detailed previously in this study is equivalent to that of an analogous four-section multi–detector row CT (LightSpeed; GE Medical Systems) protocol with the following technical parameters: 140 kVp; 220 mA; section thickness, 2.5-mm; pitch, 1.75; table speed, 15 mm per rotation; gantry speed, 0.8 second per rotation; and reconstruction thickness, 5 mm (20).
Achalasia with Complete Relaxation of Lower Esophageal Sphincter: Radiographic-Manometric Correlation

PURPOSE: To evaluate retrospectively the presence of complete relaxation of the lower esophageal sphincter (LES) at manometry in patients with achalasia depicted on barium esophagograms.

MATERIALS AND METHODS: The institutional review board approved this retrospective study and did not require informed consent. A search of radiology and manometry records identified 21 patients (12 men, nine women; mean age, 52.4 years) with achalasia depicted on barium esophagograms who had undergone manometric examinations and met the inclusion criteria. Radiologic reports and images were reviewed for presence or absence of primary peristalsis, impaired LES opening, esophageal dilatation, delayed emptying of barium, and nonperistaltic contractions. Manometry reports were reviewed for presence or absence of peristalsis or simultaneous esophageal contractions. Resting and residual LES pressures were recorded to determine whether LES relaxation was complete or incomplete. Medical records were reviewed to determine clinical presentation and follow-up (treatment and patient course), and radiographic files were reviewed to determine radiographic findings at follow-up examinations. Clinical characteristics (eg, age, dysphagia, and weight loss) were correlated with LES relaxation at manometry. Data were analyzed statistically with Fisher exact and Wilcoxon rank sum testing.

RESULTS: All 21 patients with radiographic findings of achalasia had aperistalsis at manometry. Fourteen patients (67%) had incomplete LES relaxation at manometry during swallowing, and seven (33%) had complete LES relaxation. There were no significant differences between patients with complete LES relaxation and those with incomplete LES relaxation in mean age ($P = .59$), duration of dysphagia ($P = .18$), or weight loss ($P > .99$). Clinical follow-up findings were available for six patients with complete LES relaxation at manometry and 10 with incomplete relaxation. Symptoms resolved after treatment in all six patients with complete LES relaxation. Six (60%) of 10 patients with incomplete LES relaxation had resolution of symptoms after treatment, and four (40%) had substantial improvement.

CONCLUSION: These data suggest that in patients with typical radiographic findings of achalasia, the barium study can be used to guide treatment without a need for manometry. If radiographic findings are equivocal, however, manometry may be required for a more certain diagnosis.

Primary achalasia is an uncommon esophageal motility disorder. Affected individuals typically present with long-standing dysphagia, sometimes accompanied by chest pain, regurgitation of undigested food, and recurrent aspiration pneumonia (1–3). The classic manometric criteria for achalasia include an absence of primary peristalsis in the body of the esophagus and incomplete relaxation of the lower esophageal sphincter (LES) (3,4).
This condition is also characterized at barium esophagography by esophageal aperistalsis with impaired opening of the LES, manifested by tapered, beaklike narrowing of the distal esophagus at or adjacent to the gastroesophageal junction (3,5,6). As the disease progresses, barium studies may also reveal progressive esophageal dilatation with a standing column of barium that empties slowly from the esophagus.

Despite these classic radiographic and manometric findings in achalasia, several previous investigators have reported a subset of patients with findings of achalasia at barium examination in whom manometry revealed complete LES relaxation at manometry during swallowing (7–10). The authors of these studies concluded that normal LES relaxation at manometry does not preclude a diagnosis of achalasia when patients present with characteristic clinical and radiographic features of this condition. Such data raise questions about the respective roles of barium studies and manometry in the diagnosis of achalasia. Thus, the purpose of our study was to evaluate retrospectively the presence of complete LES relaxation at manometry in patients with achalasia depicted on barium esophagograms.

**MATERIALS AND METHODS**

**Patients**

Our institutional review board approved all aspects of the retrospective study and did not require informed consent from patients whose radiographic images or medical records were included in our study. Our study was compliant with requirements of the Health Insurance Portability and Accountability Act.

A computerized search of the radiology database at our university hospital for a 6-year period from January 1998 through December 2003 revealed data for 190 patients who had undergone barium examinations and for whom the final diagnosis in the radiologic reports was achalasia. Achalasia is defined radiographically as the absence of primary peristalsis in the body of the esophagus during all swallows, with tapered, beaklike narrowing of the distal esophagus at or adjacent to the gastroesophageal junction (3,5,6). Fifty-six (29%) of the 190 patients also underwent esophageal manometry within 1 year of the barium studies. Thirty-five (62%) of these 56 patients were excluded from analysis for the following reasons: (a) previous surgical procedures (eg, Heller myotomy) had been performed for treatment of achalasia in six patients; (b) endoscopic procedures (eg, endoscopic balloon dilation or botulinum toxin injection) had been performed for treatment of achalasia between the barium studies and manometry in two patients; (c) the radiographic studies had been performed with water-soluble contrast agents to rule out perforation immediately after any form of treatment for achalasia in eight patients; (d) medical records revealed secondary rather than primary achalasia in three patients; and (e) no medical records were available in 16 patients.

The remaining 21 patients comprised our study group. Twelve patients (57%) were men, and nine (43%) were women. The mean age of the 21 patients was 52.4 years (range, 19–85 years), the mean interval between the barium studies and manometry was 2.4 months (range, 1 day to 11 months), and the median interval was 1.5 months. The barium studies were performed before manometry in 14 patients and after manometry in seven.

**Imaging**

Sixteen (76%) of the 21 patients underwent esophagography at our institution. The studies included left posterior oblique double-contrast views obtained with the patient upright and a 250% wt/vol high-density barium suspension (E-Z-HD; E-Z-Em, Westbury, NY) and/or right anterior oblique single-contrast views obtained with the patient prone and a 50% wt/vol barium suspension (Entrobar; Lafayette Pharmaceuticals, Lafayette, Ind). As part of the examination, patients were asked to take multiple discrete swallows in the prone, right anterior oblique position so that esophageal motility could be evaluated, unless markedly delayed emptying of barium from the esophagus in the upright position precluded such repositioning. All of these studies were performed by residents or fellows supervised by gastrointestinal radiologists or by one of three attending gastrointestinal radiologists, and all were interpreted by the three gastrointestinal radiologists (M.S.L., S.E.R., and I.L., with 22, 20, and 29 years of experience, respectively). The remaining five of the 21 patients underwent barium studies at outside hospitals and had detailed radiographic reports in their medical records. The radiographic images from these five barium studies and two of the 16 performed at our hospital were not available for review.

**Image and Report Review**

The original radiologic reports from all 21 barium studies and the images from 14 of them were reviewed simultaneously by one author (M.S.L.) without knowledge of the specific clinical or manometric findings. In all 21 cases, he recorded whether or not primary peristalsis was absent in the esophagus intermittently or for all swallows and whether or not there was tapered, beaklike narrowing of the distal esophagus, a well-known radiographic sign of impaired LES opening on barium esophagograms in patients with achalasia (3,5,6). Esophageal dilatation, delayed emptying of barium, and non-peristaltic contractions were also noted. Although the examinations generally were not recorded on videotape, the radiologic reports provided a relatively detailed assessment of esophageal motility, specifically stating whether primary peristalsis and features of LES dysfunction were present or absent. However, the reports mentioned non-peristaltic contractions, esophageal dilatation, or delayed emptying of barium only when these abnormalities were present on barium esophagograms, so they were assumed to be absent unless depicted on the images or described in the reports.

Thirteen patients (62%) underwent follow-up radiographic examinations with water-soluble contrast material (diatrizoate meglumine and diatrizoate sodium [Gastroview; Mallinckrodt, St Louis, Mo]) (nine patients) or barium (four patients) after endoscopic or surgical treatment of their achalasia. The studies with water-soluble contrast agents were performed within 3 days of treatment, and the barium studies were performed a mean interval of 7.6 months after treatment (range, 2 months to 6.5 years). The radiologic reports from these studies were reviewed to assess the response to treatment.

**Manometry and Report Review**

In all 21 patients, esophageal manometry was performed at our hospital by a gastroenterologist with 21 years of experience (D.A.K.) who used a solid-state intraluminal transducer assembly (Konigsberg Instruments, Pasadena, Calif). The Konigsberg catheter was placed in the stomach with a transnasal approach. Gastric (ie, intraabdominal) positioning of the pressure transducers was confirmed by a rise in pressure during the inspiratory phase of respiration. The patient was then placed in a sitting posi-
Radiology

Considered high if they exceeded 45 mm Hg, and LES relaxation was considered incomplete if residual pressures were 8 mm Hg or higher during swallowing.

Medical Records Review

Medical records subsequently were reviewed by one author (R.A.) to determine the clinical presentation and follow-up (ie, treatment and patient course). Resolution of symptoms after treatment was considered the reference standard for the presence of achalasia. Clinical characteristics of these 21 patients (age, frequency and mean duration of dysphagia, frequency and mean amount of weight loss, and frequency of other presenting symptoms) were also correlated with LES relaxation at manometry.

Statistical Analysis

A statistical analysis comparing patients with complete and those with incomplete LES relaxation was performed by means of a Fisher exact test for categorical variables (frequency of dysphagia, chest pain, heartburn, regurgitation, respiratory symptoms, and weight loss; number of patients with a hypertensive LES; number of patients with prior treatment; and average number of treatments) and a Wilcoxon rank sum test for continuous variables (mean age, mean duration of dysphagia, mean weight loss, mean LES resting pressure, and mean LES residual pressure) (S Plus 4.0; Insightful, Seattle, Wash). P values less than .05 were considered to indicate a statistically significant difference.

RESULTS

Clinical Findings

Twenty patients (95%) had dysphagia, primarily for solid foods. The mean duration of dysphagia was 3.9 years (range, 5 months to 20 years). Other symptoms included regurgitation in 17 patients (81%); chest pain in 11 (52%); heartburn in 10 (48%); respiratory symptoms such as coughing, choking, or pneumonia in six (29%); and weight loss in six (29%) (mean weight loss, 7 lb [3.15 kg]; range, 0–50 lb [0–22.5 kg]). Eight patients (38%) had undergone previous endoscopic gastroesophageal junction procedures (eg, botulinum toxin injections or balloon dilation) an average of 2.2 years (range, 1 month to 8 years) before the index barium study or manometry.

Radiographic Findings

In all 21 patients (100%), primary peristalsis was absent in the body of the esophagus during all swallows of barium. All patients also had an impaired LES opening, manifested by tapered, beaklike narrowing of the distal esophagus at or adjacent to the gastroesophageal junction, and all had varying degrees of esophageal dilatation with delayed emptying of barium from the esophagus (Figs 1, 2). Eight patients (38%) had occasional weak nonperistaltic contractions.

Manometric Findings

Peristalsis was absent in the body of the esophagus during all swallows in 19 patients (90%) and during most swallows in the remaining two (10%). In addition, 16 patients (76%) had abnormal simultaneous contractions during some swallows. Fourteen patients (67%) had incomplete LES relaxation at manometry during swallowing, and the remaining seven (33%) had complete LES relaxation (Figs 1, 2).

The mean LES residual pressure was 16.3 mm Hg (range, 8.6–31.0 mm Hg) in the 14 patients with incomplete LES relaxation during swallowing and 3.4 mm Hg (range, −2.2 to 5.5 mm Hg) in the seven patients with complete LES relaxation (P = .001). The mean LES resting pressure was 45.5 mm Hg (range, 12.3–88.0 mm Hg) in the 14 patients with incomplete LES relaxation during swallowing and 25 mm Hg (range, 12.0–45.0 mm Hg) in the seven with complete LES relaxation—a significant difference (P = .01). Seven (50%) of the 14 patients with incomplete LES relaxation had a hypertensive sphincter, and seven (50%) had a normotensive sphincter, whereas all seven patients (100%) with complete LES relaxation had a normotensive sphincter (P = .047). Thus, seven (33%) of the 21 patients with achalasia seen on barium esophagograms had normal LES function at manometry (Figs 1, 2). The manometric findings in patients with complete and patients with incomplete LES relaxation during swallowing are summarized in the Table.

Clinical and Manometric Correlation between Patients with Complete and Those with Incomplete LES Relaxation

The clinical characteristics of the 21 patients are correlated with LES relaxation at manometry in the Table. Patients with complete LES relaxation during swallowing were slightly younger than those with incomplete LES relaxation (mean age, 48.3 vs 53.7 years), but

Figure 1. Esophagogram in 23-year-old man with 6-year history of dysphagia. Left posterior oblique view obtained with the patient upright shows typical findings of achalasia: dilated aperistaltic esophagus, beaklike narrowing (arrow) near gastroesophageal junction, and slow emptying of barium into the stomach. Manometry revealed complete LES relaxation at rest. Despite the manometric findings, this patient’s dysphagia resolved after treatment for achalasia with Heller myotomy.

Medical Records Review

In all 21 patients (100%), primary peristalsis was absent in the body of the esophagus during all swallows of barium. All patients also had an impaired LES opening, manifested by tapered, beaklike narrowing of the distal esophagus at or adjacent to the gastroesophageal junction, and all had varying degrees of esophageal dilatation with delayed emptying of barium from the esophagus.
this difference was not significant \((P = .59)\). There were also no significant differences between the two groups in the frequency \((P = .33)\) or mean duration \((P = .18)\) of dysphagia, the frequency \(>(P > .99)\) or mean amount \((P = .52)\) of weight loss, or the frequency of other presenting symptoms. One patient with complete LES relaxation had a much longer duration of dysphagia than the other patients in this group. Even when this outlier was excluded from analysis, there was no significant difference between groups in the mean duration of dysphagia \((P = .46)\).

### Clinical Follow-up

Twenty (95%) of the 21 patients underwent endoscopic or surgical treatment for achalasia after the barium and manometric studies. All seven patients with complete LES relaxation at manometry underwent treatment, which involved botulinum toxin injections \((n = 1)\), Heller myotomy \((n = 2)\), botulinum toxin injections and Heller myotomy \((n = 3)\), or botulinum toxin injections and balloon dilation followed by Heller myotomy \((n = 1)\). Thirteen (93%) of the 14 patients with incomplete LES relaxation at manometry underwent treatment, which involved botulinum toxin injections \((n = 3)\), balloon dilation \((n = 3)\), Heller myotomy \((n = 2)\), botulinum toxin injections and Heller myotomy \((n = 3)\), balloon dilation and Heller myotomy \((n = 1)\), or botulinum toxin injections, balloon dilation, and Heller myotomy \((n = 1)\).

Clinical follow-up findings were available for six (86%) of the seven patients with complete LES relaxation at manometry and for 10 (71%) of the 14 patients with incomplete LES relaxation. Symptoms resolved after treatment in all six patients with complete LES relaxation and in six (60%) of the 10 patients with incomplete LES relaxation. The remaining four patients with incomplete LES relaxation had substantial improvement in symptoms. Symptoms resolved in all 13 patients in both groups who underwent Heller myotomy. The average number of treatments needed for resolution or substantial improvement of symptoms was 3.0 (range, 1–9) in the patients with complete LES relaxation versus 2.2 (range, 1–4) in those with incomplete LES relaxation \((P = .38)\).

### Radiographic Follow-up

All nine patients who underwent follow-up radiographic examinations with water-soluble contrast media within days after endoscopic or surgical treatment of achalasia were found to have decreased narrowing at the gastroesophageal junction; and slow emptying of barium into the stomach. Manometry revealed complete LES relaxation during swallowing with a normotive sphincter at rest. Despite the manometric findings, this patient’s dysphagia resolved after treatment for achalasia with botulinum toxin injection and Heller myotomy.
DISCUSSION

In our study, seven (33%) of 21 patients with typical radiographic findings of achalasia had complete LES relaxation at manometry during swallowing. An argument could be made that one or more of our patients with achalasia diagnosed at radiography but normal manometric LES relaxation had false-positive barium study results and did not have achalasia. However, all seven patients in this group underwent treatment for achalasia with botulinum toxin injections, endoscopic balloon dilation, and/or Heller myotomy, and follow-up data in six revealed complete resolution of dysphagia after treatment in all cases. Our findings suggest that a substantial number of patients with achalasia depicted on barium studies have complete LES relaxation at manometry during swallowing. Earlier studies (7,8) also revealed that 20%–30% of patients with typical findings of achalasia at radiographic examinations had normal manometric LES relaxation. Thus, the degree of LES relaxation at manometry is not a reliable criterion for the diagnosis of achalasia.

In two of the earlier studies that revealed complete LES relaxation in patients with achalasia, patients with normal LES function at manometry were younger and had less weight loss, a shorter duration of dysphagia, and less esophageal dilatation on barium esophagograms than those with incomplete LES relaxation (7,8). The authors therefore concluded that such patients had a less advanced form of achalasia. In our study, however, there were no significant differences between these groups in mean age, mean weight loss, or mean duration of dysphagia. Another study also revealed no significant differences between the two groups in any of these clinical parameters (9). Such findings therefore do not corroborate the hypothesis that complete LES relaxation at manometry is a characteristic of early achalasia.

The previous studies did not reveal significant differences in LES resting pressures between patients with achalasia and complete LES relaxation at manometry and those with achalasia and incomplete LES relaxation (7,8). In our study, however, patients with complete LES relaxation during swallowing had a significantly lower mean LES resting pressure than those with incomplete LES relaxation (P = .01) and were significantly more likely to have a normotensive sphincter (P = .047). Our findings indicate that a substantial number of patients with achalasia have complete LES relaxation during swallowing and normal LES resting pressures with no evidence of LES dysfunction at manometry. The patients with complete and those with incomplete LES relaxation at manometry underwent comparable numbers of treatment procedures (suggesting similar severity of disease), and both groups showed excellent responses to treatment, with resolution or improvement of symptoms, regardless of the degree of LES dysfunction at manometry.

These data raise questions about the role of manometry in patients with long-standing dysphagia who are suspected of having achalasia. If the barium study reveals typical findings of achalasia, the authors therefore established a diagnosis of achalasia. However, we studied an inability to identify patients with achalasia diagnosed at manometry who had false-negative esophagograms. As a result, manometry still may be required to establish the diagnosis in patients clinically suspected of having achalasia who have equivocal or even negative radiographic results.

Other conditions occasionally can mimic the findings of primary achalasia on barium esophagograms. The most important condition is secondary achalasia caused by malignant tumor at the gastroesophageal junction or, less commonly, by benign conditions such as Chagas disease (12). In patients with secondary achalasia, however, the narrowed segment tends to be longer than that in primary achalasia (extending more than 3.5 cm above the gastroesophageal junction) and is frequently associated with nodularity, ulceration, or asymmetry (12,13). A smooth, tapered peptic stricture in the distal part of the esophagus may also resemble the beaklike narrowing of achalasia, but most patients with peptic strictures have hiatal hernias and normal esophageal peristalsis. Thus, it usually is possible to differentiate these conditions from primary achalasia on the basis of the radiographic findings.

It is important to recognize the inherent limitations of our retrospective study, which included selection bias and possible inaccurate reporting of symptoms in the medical records. Because of the relatively small number of patients who underwent manometry, manometric correlation was not possible for most patients with achalasia diagnosed at radiography, creating an additional selection bias. We also had to rely on the original radiographic reports for characterization of the motility findings because the fluoroscopic findings generally were not recorded on videotape at the time of these examinations. Unfortunately, some patients had a relatively long interval between the barium studies and manometry, but achalasia is a chronic disease that gradually evolves over time, minimizing the effect of this bias. Finally, our failure to detect significant differences for some parameters between patients with complete LES relaxation and those with incomplete relaxation at manometry could have been related to inadequate sample sizes, a common methodologic problem in studies of this type (14). Because of our study limitations, a prospective investigation of a large series of patients suspected of having achalasia is needed, with barium studies and manometry performed in all cases to elucidate further the respective roles of these examinations.

In conclusion, we found that one-third of patients with achalasia depicted on barium studies had normal LES relaxation at manometry, but there were no significant differences in the clinical presentation between those with complete and those with incomplete LES relaxation during swallowing. All patients with achalasia on barium esophagograms had an excellent response to treatment, regardless of the degree of LES relaxation at manometry. Our experience suggests that in patients with typical radiographic findings of achalasia, the barium study can be used to guide treatment without a need for manometry. If the radiographic findings are equivocal, however, manometry may be required for a more certain diagnosis.

References


Gastrointestinal Stromal Tumor: New Nodule-within-a-Mass Pattern of Recurrence after Partial Response to Imatinib Mesylate

**Purpose:** To investigate a new pattern of tumor recurrence observed at imaging in patients with metastatic gastrointestinal stromal tumor (GIST) after initial partial response to imatinib mesylate.

**Materials and Methods:** Ninety-two patients with metastatic GIST who underwent treatment in a clinical trial with oral imatinib mesylate were followed up for 29 months. An institutional review board–approved protocol was used. The study complied with the Health Insurance Portability and Accountability Act, and written informed consent was obtained from all patients. Images of the chest, abdomen, and pelvis, acquired with computed tomography (CT), positron emission tomography (PET), and, in some cases, magnetic resonance imaging, were evaluated for treatment response and disease recurrence. Thirty-nine patients (29 men, 10 women; age range, 18–84 years; mean, 49.2 years) had recurrent disease after an initial variable period of response (range, 2–24 months; median, 14.4 months). Initial response was determined with findings of decreased uptake of fluorine 18 fluorodeoxyglucose at PET, shrinkage of tumor, and decreased attenuation at CT. Images were evaluated for disease recurrence by two experienced radiologists who were blinded to each other’s interpretation but not to clinical details. Final reading was performed by consensus.

**Results:** A nodule within a mass was seen in 21 of 39 patients (in intrahepatic tumor [n = 8], extrahepatic tumor [n = 10], or both intra- and extrahepatic tumors [n = 3]) and was the first sign of disease progression in 17 of 21 patients. Other patterns of recurrence included new site of disease (n = 7), regrowth of preexistent lesion (n = 20), and mixed (more than one) pattern (n = 9). Disease progression was verified at needle biopsy (n = 16), follow-up imaging (n = 14), and/or surgical resection (n = 9).

**Conclusion:** A nodule within a mass is an important sign of recurrent GIST, but measurements of overall tumor size may not enable detection of such nodules.

---

Gastrointestinal stromal tumor (GIST), while relatively rare, is the most common of the mesenchymal malignancies that may affect the gastrointestinal tract and represents approximately 1%–3% of all malignant gastrointestinal tumors (1). While the pathologic features of GIST were first described approximately 2 decades ago, a clear understanding of the pathophysiologic and molecular basis of this disease was achieved only in the past few years. GISTs express the cell-surface transmembrane receptor KIT, which has tyrosine kinase activity and is the protein product of the KIT proto-oncogene. Gain-of-function mutations of KIT are frequent in GIST. These mutations result in the constitutive activation of KIT signaling, which leads to uncontrolled cell proliferation and resistance to apoptosis. It has been reported that KIT activation occurs in all cases of GIST, regardless of the mutational status of KIT (1,2).
Unresectable or metastatic GIST is a fatal disease that resists conventional cytotoxic chemotherapy. In one patient series, the response rate to doxorubicin was less than 5% (3). The effectiveness of radiation therapy for this disease has not been proved. The median duration of survival for patients with a metastatic GIST is approximately 20 months, and that for patients with local recurrence is 9–12 months (2,4).

Technologic advances recently led to the development of a targeted molecular therapeutic agent, the receptor tyrosine kinase inhibitor imatinib mesylate (Gleevec; Novartis, New York, NY), formerly known as STI-571, which has been used as a systemic treatment for GIST (1,2).

In solid tumors, the progression of metastatic disease typically is heralded by the appearance of tumors in other sites and/or an increase in the size of preexisting lesions (5–10). Conversely and most commonly, a positive response to conventional chemotherapeutic agents is signaled by shrinkage of the tumor. When the effectiveness of chemotherapy decreases after an initial positive response, the tumor usually increases in size, and new sites of disease may appear. Very little information is available in the literature about the appearance of GIST after treatment with imatinib mesylate. The purpose of this study was to investigate a new pattern of tumor recurrence observed at imaging in patients with metastatic GIST after initial partial response to imatinib mesylate.

MATERIALS AND METHODS

Study Group

Ninety-two consecutive patients (56 men, 36 women; age range, 18–86 years; median age, 47 years) with biopsy-proven metastatic GIST who were treated at our institution (Dana-Farber Cancer Institute, Harvard Medical School, Boston, Mass) with imatinib mesylate (clinical trial) were followed up prospectively over a period of 29 months. The study was performed by using an institutional review board–approved protocol and complied with the Health Insurance Portability and Accountability Act. Written informed consent was obtained from all patients. Follow-up was performed with clinical examination, routine blood analysis (complete and differential blood cell counts, electrolyte measurements, and liver function tests), and imaging. One of two clinical oncologists (J.D., G.D.D.) and a nurse practitioner performed clinical follow-up. G.D.D. is a consultant with and an occasional invited lecturer for Novartis and received research support from that company.

A total of 39 (42%) of 92 patients showed evidence of disease progression after an initial variable period of response to imatinib mesylate (range, 2–24 months; mean, 13.5 months; median, 14.4 months). This group included 29 men and 10 women with a mean age of 49.2 years (range, 18–84 years; median, 47 years). Of the 39 patients who experienced disease progression, all had purely abdominal disease at baseline (prior to initiation of imatinib mesylate therapy). Five of these patients had only extrahepatic disease, four had only intrahepatic disease, and 30 had both intra- and extrahepatic disease.

Image Acquisition

All patients underwent contrast material–enhanced computed tomography (CT) (Volume Zoom; Siemens, Erlangen, Germany) of the chest, abdomen, and pelvis. Sections with a thickness of 7 mm were acquired with a delay of 70 seconds after intravenous administration of 100 mL of ioxilan (Oxiplan 300; Guerbet, Blooming- ton, Ind). Unenhanced scans were not routinely obtained, but follow-up in six patients included CT without intravenous contrast material for evaluation of intractable bleeding or renal stone or when the creatinine level was elevated. All six patients underwent contrast-enhanced CT either immediately after unenhanced CT (n = 2) or at the next follow-up visit, which was within 6 weeks in all cases. CT scans were obtained approximately every 6 weeks, per the clinical drug trial protocol.

Positron emission tomography (PET) was performed by using a PET imager (ECAT Exact HR+; Siemens/CTI, Knoxville, Tenn). After patients fasted for 4 hours, they received an intravenous injection of 740 MBq (20 mCi) of fluorine 18 fluorodeoxyglucose (FDG). With a 45–60-minute delay after the injection, PET was performed from the base of the skull to the upper thighs by using a multipledetector whole-body protocol. Standardized values for uptake were not used in this study. PET scans were obtained at 6–12-week intervals.

Magnetic resonance (MR) imaging (Signa; GE Medical Systems, Waukesha, Wis) of the liver was performed in 15 of 39 patients in whom percutaneous interventional therapy (radiofrequency ablation or cryoablation) was contemplated. If the interventional therapy was administered, the patients underwent follow-up MR imaging every 3 months afterward, in addition to the standard follow-up examination for monitoring of response. MR imaging included transverse T1-weighted spin-echo (repetition time msec/echo time msec, 600/14; section thickness, 4 mm; field of view, 34 cm), transverse T2-weighted fast spin-echo (5100/100; echo train length, 12; section thickness, 4 mm; field of view, 30 cm), and transverse fatsuppressed fast multplanar spoiled gradient-echo (285/1.6; flip angle, 75°; section thickness, 5 mm; field of view, 34 cm) sequences applied before and after intravenous injection of 20 mL gadoxetate dimeglumine (Magnevist; Berlex Laboratories, Wayne, N.J).

Image Review and Evaluation

Two experienced radiologists (S.S. and E.V.S., with 10 and 28 years of experience, respectively, in cross-sectional imaging, and 3 and more than 10 years of experience, respectively, in cancer imaging) performed the measurements on CT scans by using electronic calipers at a picture archiving and communication system workstation (IMPAX; Agfa, Ridgefield Park, NJ). We ensured that the same lesion was evaluated on pretreatment and follow-up scans and that the measurement technique was uniform with respect to section selection and anatomic level. The radiologists were blinded to each other’s reading but were aware of both the underlying disease and the therapy. The final data compilation was performed by consensus. The radiologists specifically looked for enlargement or decrease in the size of lesions, changes in lesion attenuation, the presence of new lesions and/or new sites of involvement, and the appearance of nodules within preexistent lesions (a new pattern that we had observed).

Criteria.—A change of more than 25% in the bidimensional measurement (sum of length and width) of the tumor was classified as an increase or a decrease in size. A change of more than 20 HU in attenuation values, with assessment based on a mean of three different matched measurements, was considered substantial enough to be classified as decreased or increased attenuation, similar to observations by Chen et al (11). The presence of a lesion in a location where none was present on the previous CT scan was considered to indicate a new site of disease.

The presence of a new nodule within
patterns of progression, including enlargement of preexistent lesions and presence of new lesions. For each CT examination, the images were reviewed independently by at least two of three radiologists (S.S., E.V.S., P.J.D.), and the final results were documented by consensus. Findings on PET images were correlated with CT findings by the nuclear medicine physician (A.V.D.A.) in collaboration with at least one of the three radiologists.

**Documentation of nodules.**—Verification of recurrent tumor in patients who exhibited the imaging pattern of a nodule within a mass was accomplished with percutaneous CT or ultrasonographic (US) image-guided fine-needle and/or core-needle biopsy, surgical resection, and/or concomitant or follow-up imaging with CT, MR, or PET.

**Image comparisons.**—The images from the six unenhanced CT studies were retrospectively compared (S.S., E.V.S., A.V.D.A.) with the contrast-enhanced scans obtained either at the same time or at the next follow-up session, for completeness. FDG PET images were subsequently compared with the CT scans by using a side-by-side display at the picture archiving and communication system workstation.

**RESULTS**

**Patterns of Recurrence**

The pattern of a nodule within a mass was present at imaging in 21 (54%) of 39 patients and was verified by means of fine-needle and/or core-needle biopsy in eight of 21, surgical excision in six of 21, and follow-up imaging in seven of 21. The appearance of such a nodule on CT and/or PET scans was the first sign of disease progression in 18 (81%) of 21 patients in whom such nodules were seen (17 [44%] of 39 patients with disease recurrence). Other patterns of recurrence included a new site of disease (seven [18%] of 39), regrowth of a preexistent lesion (20 [51%] of 39), and mixed pattern (more than one pattern of disease progression) (nine [23%] of 39) (Table).

**Detection of a Nodule within a Mass**

With regard to the pattern of a nodule within a mass, both radiologists identified the largest nodule on the CT scans, with a difference of opinion occurring only with regard to the presence of additional nodules. The nodules identified on CT scans were also identified on PET scans in 18 of 21 patients and on MR images in nine of 21 patients.

**Distribution of Nodules**

Among the 21 patients, a nodule within a mass was seen in intrahepatic tumors in eight (38%), in extrahepatic tumors in 10 (48%), and in both extra- and intrahepatic tumors in three (14%) patients (Fig 1). Mural nodules were seen in 12 of these 21 patients, and both intramatrix and mural nodules were seen in the other nine of 21 patients (Fig 2). No patient had clear evidence of an intramatrix nodule alone (without an accompanying mural nodule).

**Verification of Tumor Recurrence**

When the nodules were initially identified on CT or PET scans, they were 2 mm to 1.4 cm in diameter. Tissue specimens obtained with percutaneous CT- or US-guided fine-needle and/or core-needle biopsy in eight of 21 patients and surgical excision in six of 21 patients showed evidence of tumor recurrence at histopathologic examination. Follow-up imaging with CT, MR, or PET demonstrated either interval enlargement of nodules within preexistent masses that were presumed to be dormant (on the basis of absence of metabolic activity on prior FDG PET scans or stability in size) and/or increased metabolic activity evidenced by new abnormal foci of uptake on FDG PET scans.

**Size of Nodules**

When nodules were initially identified, they were 2 mm to 1.4 cm in diameter. During the period of data collection, all nodules demonstrated growth, and some even enlarged to the size of the preexistent mass or engulfed it. All patients in whom initial images showed a nodule within a mass developed multiple nodules that increased in size over the period of data collection.

**Imaging Appearance of Nodules**

Nodules were not well visualized on unenhanced CT scans but were seen to enhance on the immediate postcontrast scan (n = 2) or the next follow-up postcontrast scan (n = 4), on which they demonstrated attenuation equal to that of the surrounding normal liver parenchyma. The enhancement of nodules was similar to that of the original tumor on the baseline scan (subjective observation) prior to the initiation of imatinib mesylate therapy.

MR images (obtained in nine of 21 patients) showed isointense to mildly hy-
perintense signal in the lesions, compared with signal intensity in normal liver parenchyma on T2-weighted images. The masses within which the nodules originated appeared hyperintense, with signal intensity similar to that of fluid. The signal in the nodules appeared isointense compared with signal intensity in liver parenchyma on unenhanced T1-weighted images, and the nodules demonstrated enhancement similar to that seen on postcontrast CT scans (Fig 3). The FDG PET images depicted new foci of increased FDG uptake in 18 of 21 patients; all of these lesions corresponded to nodules seen at CT. One of the patients did not undergo PET, and two of 21 patients had a negative PET scan. The diameters of the latter two lesions were 7 and 9 mm, respectively, on CT scans when the lesions were first seen.

Figure 1. Patient 35. Metastatic GIST in 53-year-old man. (a) Transverse contrast-enhanced CT scan obtained prior to initiation of imatinib mesylate therapy shows three lesions (arrows) in the liver and a large mass in the left upper abdominal quadrant intimately associated with the stomach, with small areas of internal calcification. All lesions are heterogeneously enhanced. (b) PET projection image obtained prior to therapy shows large FDG-avid mass (straight arrow) in left upper quadrant of abdomen and smaller masses (curved arrow) in liver that correspond to lesions in a. (c) Transverse contrast-enhanced CT scan obtained 3 months after start of therapy shows shrinkage of tumors and decreased CT attenuation values in all lesions. (d) PET projection image obtained 2 months after start of therapy shows no discernible uptake of FDG in the left upper quadrant, a finding that signifies decreased metabolic activity. Metabolic activity in the hepatic masses also appears to have ceased. Normal uptake of FDG is seen in both renal collecting systems and the urinary bladder. (e) Transverse contrast-enhanced CT scan obtained 17 months after start of therapy demonstrates recurrent nodule (arrow) in extrahepatic mass after initial good response. (f) PET projection image obtained at the same time as e demonstrates new focal area of abnormal FDG uptake (straight arrow) that corresponds to nodule in e. A few smaller nodules (curved arrows) also are visible.
There were subjectively noted morphologic differences on CT and MR images between nodules that originated in intrahepatic lesions and nodules that originated in extrahepatic lesions. The former appeared solid, with relatively homogeneous enhancement, while the latter tended to have a more ringlike enhancement pattern (Fig 4).

**DISCUSSION**

The pattern of tumor recurrence described here, the nodule within a mass, is a new imaging appearance that, to our knowledge, has not been reported previously. Our study results demonstrate that this finding indicates progression of disease. According to commonly accepted criteria, masses that are stable in size indicate no progression of disease. However, this new finding of enhancing nodules within stable masses was the first sign of disease progression in 17 of 21 patients with GIST that was treated successfully, at least initially, with imatinib mesylate. It is important to note that although the clinical protocol did not specify that the presence of a nodule indicated tumor progression, Southwest Oncology Group criteria were used in the study design, and the radiologists and clinicians therefore considered this finding a sign of recurrence.

Tumor measurement protocols used to assess response to therapy typically are based on the measurement of one dimension (as in the protocol developed by the Response Evaluation Criteria in Solid Tumors Group [12]) or two dimensions (as in the standard Southwest Oncology Group criteria [13]) on transverse CT images or involve the calculation of volume by using computer-assisted modeling (6,8,9,14,15). Because these techniques are predicated on external measurements of the lesion, the presence of a nodule within the mass would go undetected.

<table>
<thead>
<tr>
<th>Patient No./Age (y)/Sex</th>
<th>Nodule</th>
<th>New Site</th>
<th>Regrowth*</th>
<th>Time to Progression (mo)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/65/M</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>5</td>
</tr>
<tr>
<td>2/60/M</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>17</td>
</tr>
<tr>
<td>3/47/F</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>12</td>
</tr>
<tr>
<td>4/42/M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>18</td>
</tr>
<tr>
<td>5/80/M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>16</td>
</tr>
<tr>
<td>6/57/M</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>16</td>
</tr>
<tr>
<td>7/47/M</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>12</td>
</tr>
<tr>
<td>8/36/M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>4</td>
</tr>
<tr>
<td>9/78/F</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>9</td>
</tr>
<tr>
<td>10/43/M</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>17</td>
</tr>
<tr>
<td>11/38/F</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>6</td>
</tr>
<tr>
<td>12/58/F</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>20</td>
</tr>
<tr>
<td>13/39/M</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>13</td>
</tr>
<tr>
<td>14/45/M</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>6</td>
</tr>
<tr>
<td>15/36/M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>8</td>
</tr>
<tr>
<td>16/64/M</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>4</td>
</tr>
<tr>
<td>17/62/M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>8</td>
</tr>
<tr>
<td>18/84/M</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>13</td>
</tr>
<tr>
<td>19/47/M</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>10</td>
</tr>
<tr>
<td>20/54/M</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>21</td>
</tr>
<tr>
<td>21/54/F</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td>22/58/M</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>13</td>
</tr>
<tr>
<td>23/40/M</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>12</td>
</tr>
<tr>
<td>24/18/M</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>24</td>
</tr>
<tr>
<td>25/45/M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>19</td>
</tr>
<tr>
<td>26/47/M</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>17</td>
</tr>
<tr>
<td>27/40/M</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>19</td>
</tr>
<tr>
<td>28/63/M</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>18</td>
</tr>
<tr>
<td>29/46/M</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>5</td>
</tr>
<tr>
<td>30/36/F</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>19</td>
</tr>
<tr>
<td>31/28/M</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>10</td>
</tr>
<tr>
<td>32/36/F</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>21</td>
</tr>
<tr>
<td>33/46/F</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>20</td>
</tr>
<tr>
<td>34/36/M</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>17</td>
</tr>
<tr>
<td>35/53/M</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>17</td>
</tr>
<tr>
<td>36/45/F</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>13</td>
</tr>
<tr>
<td>37/32/M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>16</td>
</tr>
<tr>
<td>38/48/F</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>14</td>
</tr>
<tr>
<td>39/47/M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>15</td>
</tr>
</tbody>
</table>

Note.—The numbers of patients with each pattern of recurrent disease were as follows: nodule (within a preexistent mass), 21 patients; new site, seven patients; regrowth, 20 patients.

* Regrowth was defined as increased size of the preexistent tumor.
The development of imatinib mesylate represents one of the early and important innovations in molecular targeting of cancer with chemotherapy (1). Imatinib mesylate appears to selectively target cells that produce the mutant form of c-KIT. Experience with the imaging characteristics of tumor response to this class of drugs, many of which are in early stages of clinical development, is limited (1). The criteria for response to therapy have largely been modeled on previous experience with conventional cytotoxic chemotherapeutic agents. According to the results of our study, indicators of progression of disease during imatinib mesylate therapy may be tumor enlargement, new sites, and/or a new nodule within the mass.

It is interesting to compare findings at follow-up imaging after radiofrequency ablation of hepatic tumors with the findings after chemotherapy in GIST. After successful treatment with radiofrequency ablation, tumors do not shrink rapidly or disappear, as they do after treatment with cytotoxic chemotherapeutic agents; rather, they demonstrate a well-defined area that lacks contrast enhancement on CT scans or MR images (10), an appearance reminiscent of that in GIST after treatment with imatinib mesylate. Patterns of tumor recurrence after treatment have been extensively studied in the radiofrequency ablation literature (10,16–20); according to accepted criteria, a recurrent tumor is defined as a new focal area with contrast enhancement on CT or MR images and/or hyperintense signal on T2-weighted MR images and with nodularity or irregular thickening of the boundaries of the treated area (10). Recurrence after radiofrequency ablation is found only at the margins of the treated area and increases the overall size of the treated lesion (10,17,18), while the pattern of a nodule within a mass that we have seen in GIST does not increase the overall size of the mass because the recurrence lies completely within the measured area or volume.

Twenty patients (21%) with progressive disease in our series demonstrated an apparent global increase in size and enhancement of preexistent nonenhanced or hypoenhanced lesions without development of intratumoral nodules. While it is possible that this pattern may differ from that of the nodule within a mass, it is also possible that this global increase might represent a particularly aggressive variant of the nodule within a mass, one that grows rapidly, fills, and then expands the original tumor. Further studies with sequential and detailed imaging may better evaluate this speculation.

The appearance of a nodule within a mass on CT scans or of a new focal area of increased FDG uptake on PET scans is suggestive of disease recurrence even after a relatively long period of apparent quiescence characterized by no growth and low attenuation at CT and by lack of FDG uptake in the treated lesions after initial response to imatinib mesylate. New focal FDG uptake was seen in the majority (86%) of our patients and correlated with new nodules in masses on CT scans. The two patients who had negative PET scans may have had lesions with a size below the threshold for detection, as the nodules on the corresponding CT scans were smaller than 1 cm.

We speculate that the origin of this pattern of a nodule within a mass is due to the localized cloning of mutant tumor cells, followed by the development of resistance to c-KIT inhibition with imatinib mesylate. This impression is further strengthened by the appearance of some nodules as enlargements of a clone of mutated cells within the nonenhancing (and PET-negative) tumor matrix, in contradistinction to the appearance of mural nodules at the edge of the lesion, where, arguably, residual tumor cells might be expected to exist. Imatinib mesylate is a selective inhibitor of certain tyrosine kinases, among them the constitutively active tyrosine kinase in the cells of GISTs (1,2,21). Our speculation has been supported by experiments in human tumor-cell lines that are dependent on the KIT pathway. Exposure of these cells to imatinib mesylate blocked the kinase activity of KIT, arrested proliferation, and caused apoptotic cell death (22,23). Further stud-
ies to assess the mutations in these nodules are required to prove this point.

In several other studies (1,2,21) in which biopsy was performed in low-at-
tenation masses that appeared quiescent on images, reduced numbers of tu-
mor cells were found at the baseline biopsy, and a hypocellular myxohyaline stroma with small numbers of scattered atypical nuclei and, frequently, prominent stromal hemorrhage were demonstrated. Frank necrosis of the tumor rarely was seen (2) even though the im-
aging appearance was suggestive of necrosis, with a marked decrease in contrast enhancement on CT scans and MR im-
ages and decreased uptake of the radio-
tracer on PET scans. The appearance of clonal nodules, therefore, is in keeping with the apparent mechanism of action of imatinib mesylate, in that the thera-
peutic agent does not induce substantial necrosis in the tumor.

Limitations of the study include the fact that pathologic proof was not avail-
able for every patient or for every lesion. In addition, the time intervals between imaging examinations slightly differed for some of the patients, and MR images were available only in a subset of patients with the finding of a nodule within a mass. Since nodules were detected initially only on CT scans, we do not believe that any bias was introduced because of the analysis of MR images.

In conclusion, several patterns of dis-
ease recurrence can be identified at imaging in patients with GIST after an initial positive response to imatinib mesylate treatment, and we report the pattern of a nodule within a mass, a finding that is important for the diagnosis of progres-
sive disease because commonly accepted unidimensional, bidimensional, or vol-
ume measurements of tumor mass do not enable the detection of nodular recur-
rence.

Acknowledgments: The authors are grateful to Amy Potter and Kristin Rancourt for assistance with the manuscript.

References
1. George S, Desai J. Management of gastro-
intestinal stromal tumors in the era of tyrosine kinase inhibitors. Curr Treat Op-
tions Oncol 2002; 3(6):489–496.
2. Demetri GD, von Mehren M, Blanke CD, et al. Efficacy and safety of imatinib me-
480.
ical characteristics of gastrointestinal stromal tumors (GIST) (abstr). In:
4. DeMatteo RP, Lewis JJ, Leung D, Mudan SS, Woodruff JM, Brennan MF. Two hun-
dred gastrointestinal stromal tumors: re-
6. Prasad SR, Jhaveri KS, Saini S, Hahn PF, Halpern EF, Sumner JE. CT tumor measure-
ment for therapeutic response assessment: comparison of unidimensional, bidimen-
sional, and volumetric techniques—initial observations. Radiology 2002; 225:416–
419.
8. Sohaib SA, Turner B, Hanson JA, Farquhar-
son M, Oliver RT, Reznek RH. CT assess-
10. Dromain C, de Baere T, Elias D, et al. Hepatic tumors treated with percutane-
11. Chen MY, Rechtold RE, Savage PD. Cystic changes in hepatic metastases from gastro-
intestinal stromal tumors (GISTs) treated with Gleevec (imatinib mesylate). AJR Am J
Roentgenol 2002; 179:933–946.
12. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the re-
sponse to treatment in solid tumors. Eu-
ropean Organization for Research and Treatment of Cancer, National Cancer In-
situte of the United States, National Can-
13. Green S, Weiss GR. Southwest Oncology Group standard response criteria, end-
point definitions and toxicity criteria. In:
Pharmacology, New York: Marcel Dekker, 1997; 260–263.
14. Dachman AH, MacEneaney PM, Adedipe A, Carlin M, Schumm LP. Tumor size on computed tomography scans: is one mea-
surement enough? Cancer 2001; 91:555–
560.
frequency ablation with cooled-tip elec-
ment and follow-up in 16 patients. Radi-
18. Goldberg SN, Gazelle GS, Solbiati L, et al. Ablation of liver tumors using percutane-
19. Gazelle GS, Goldberg SN, Solbiati L, Livraghi T. Tumor ablation with radio-
20. McGahan JP, Dodd GD 3rd. Radiofre-
21. George S, Desai J. Management of gastro-
intestinal stromal tumors in the era of tyrosine kinase inhibitors. Curr Treat Op-
tions Oncol 2002; 3:489–496.
22. Heinrich MC, Griffith DJ, Druker BJ, Wait
CL, Ott KA, Zigler AJ. Inhibition of c-kit receptor tyrosine kinase activity by STI
23. Tuvesson DA, Willis NA, Jacks T, et al. STI571 inactivation of the gastrointestinal
stromal tumor c-KIT oncprotein: bi-
ological and clinical implications. Onco-
Living Donor Candidates for Right Hepatic Lobe Transplantation: Evaluation at CT Cholangiography—Initial Experience

POURPOSE: To retrospectively evaluate computed tomographic (CT) cholangiography in the depiction of second-order biliary tract anatomy in living donor candidates for right hepatic lobe transplantation.

MATERIALS AND METHODS: Human research committee approval was obtained, informed consent was not required, and the study was compliant with the Health Insurance Portability and Accountability Act. The authors identified all living right-lobe liver donor candidates who underwent CT cholangiography at their institution between October 2001 (when CT cholangiography was introduced at the institution) and March 2003 (n = 62). There were 41 men (mean age, 36 years; range, 18–55 years) and 21 women (mean age, 40 years; range, 22–55 years). Two readers in consensus rated quality of second-order bile duct visualization at CT cholangiography on a four-point scale (0, not seen; 3, excellent visualization) and noted the presence of variant second-order biliary tract branching anatomy. CT cholangiography findings were compared with those at surgery in subjects who underwent right hepatic lobe retrieval (n = 24). In addition, adult donors who underwent right hepatic lobe retrieval between January 2000 and March 2003 (29 men, mean age, 35 years [range, 20–52 years]; 18 women, mean age, 38 years [range, 23–54 years]) were identified. Numbers of donors who underwent intraoperative cholangiography before and after the introduction of CT cholangiography were compared by using the Fisher exact test.

RESULTS: The mean second-order bile duct score at CT cholangiography was 2.9 (range, 2–3). Of 24 subjects who underwent right lobe retrieval, biliary tract anatomy determined at CT cholangiography was concordant with findings at surgery in 23 (96%). Variant second-order branching anatomy was seen in 13 subjects (54%) at surgery; one variant branch was missed at CT cholangiography. Of 47 subjects who underwent right hepatic lobe retrieval, significantly fewer subjects required conventional intraoperative cholangiography after the introduction of CT cholangiography (three of 24 subjects [12%]) than before (23 of 23 subjects; P < .0001).

CONCLUSION: CT cholangiography accurately depicts biliary tract anatomy in living donor candidates for right hepatic lobe transplantation, and donors who undergo preoperative CT cholangiography are unlikely to need conventional intraoperative cholangiography.
biliary tract anatomy affects the surgical approach and biliary anastomotic technique and may preclude liver donation (5,6). Although endoscopic retrograde cholangiography is the standard test for defining biliary anatomy, it is invasive and has a major complication rate of 1.4%–3.2% (7,8). The noninvasive evaluation of the biliary tract with computed tomographic (CT) cholangiography has shown promising results in small series (9,10). Therefore, we undertook this study to retrospectively evaluate CT cholangiography in the depiction of second-order biliary tract anatomy in living donor candidates for right hepatic lobe transplantation.

MATERIALS AND METHODS

Subjects

This was a retrospective single-institution study approved by our Committee on Human Research. Patient informed consent was not required. Our study was compliant with the Health Insurance Portability and Accountability Act. CT cholangiography has been described in reports of studies in the United States (11–13) and has been used extensively in Asia (14–19) and Europe (10,20–22). In addition, intravenous iodipamide meglumine (Cholografin; Bracco Diagnostics, Princeton, NJ) is approved by the U.S. Food and Drug Administration for imaging of the biliary tract. For these reasons, our Committee on Human Research did not require its approval for use of this agent in clinical examinations of the biliary tract.

One radiologist (Z.J.W.) retrospectively identified all consecutive CT cholangiograms obtained at our institution in living donor candidates for right-hepatic lobe transplantation between October 2001 (when CT cholangiography was introduced at our institution) and March 2003. There were 62 donor candidates: 41 men with a mean age of 36 years (range, 18–55 years) and 21 women with a mean age of 40 years (range, 22–55 years). There was no statistically significant difference in mean age between men and women (P = .15, unpaired two-sample t test). Twenty-four of these 62 patients subsequently underwent right hepatic lobe retrieval. The surgeons performing transplantation were aware of all findings at CT cholangiography before surgery. In this series, no patients were excluded from liver donation on the basis of biliary duct anatomy as seen at CT cholangiography. Intraoperative cholangiography was performed at the discretion of the surgeons when confirmation of biliary anatomy was believed to be necessary.

One radiologist (Z.J.W.) also identified all adult donors at our institution who underwent right hepatic lobe retrieval between January 2000 (the start of adult living related liver transplantation) and March 2003 (n = 47). Subjects included 29 men with a mean age of 35 years (range, 20–52 years) and 18 women with a mean age of 38 years (range, 23–54 years). There was no statistically significant difference in mean age between men and women (P = .41, unpaired two-sample t test).

The first author (Z.J.W.) recorded the need for conventional intraoperative cholangiography before and after the introduction of CT cholangiography to see whether the existence of CT cholangiography had any effect on the frequency of conventional intraoperative cholangiography. Before the introduction of CT cholangiography at our institution, intraoperative cholangiography was routinely performed to define biliary anatomy before right hepatic lobe retrieval. The medical records of all 62 patients who underwent CT cholangiography were reviewed by one radiologist (Z.J.W.) to identify any potential donors who were excluded from liver donation because of a biliary anatomic variant demonstrated at CT cholangiography. Surgeons were aware of the results of CT cholangiography before surgery, and intraoperative cholangiography was performed at the discretion of the surgeon.

CT Cholangiographic Technique

All 62 patients underwent multi-detector row CT evaluation of the hepatic vascular anatomy, liver volume, and liver parenchyma immediately before undergoing CT cholangiography. Our CT cholangiographic technique was heterogeneous because it was evolved over time. CT scans obtained before February 2002 (n = 8) were obtained with a four-detector row unit (high-speed mode, Lightspeed LX/i; GE Medical Systems, Milwauk ee, Wis), and scans obtained during or after February 2002 (n = 54) were obtained with a 16-detector row unit (Lightspeed; GE Medical Systems). The first author estimated the average dose of CT cholangiography by multiplying the dose length product value recorded from the scanner console by multipliers specific to the abdomen according to the method outlined by Jessen et al (23). The estimated dose for a CT cholangiogram is about 7.6 mSv. The average dose for an intraoperative cholangiogram is similar to that for an abdominal radiograph, which has been reported in the literature to have a dose equivalent of 0.6–1.7 mSv (24).

No oral contrast material was administered for CT cholangiography. Before the administration of cholangiographic contrast material, each subject received 25 mg of intravenous diphenhydramine (Benadryl; Pfizer, New York, NY). In subjects examined before January 2003 (n = 41), we also administered intravenous morphine sulfate (Abbott Laboratories, Chicago, Ill; 0.04 mg per kilogram of body weight) to contract the sphincter of Oddi (25–27) and possibly improve biliary distention. However, because the image quality of the CT cholangiograms did not improve with intravenous morphine administration (B.M.Y., unpublished data, November 2001), the remaining subjects (n = 21) were not given morphine. Twenty milliliters of iodipamide meglumine 52% (Cholografin; Bracco Diagnostics) diluted in 80 mL of normal saline was infused over 30 minutes.

The liver was imaged during a single breath hold 15 minutes after the completion of infusion. Section thickness, gantry rotation time, and table-top speed were 1.25 mm, 0.8 per second, and 13.5 mm/sec for the 16–detector row scanner and 2.5 mm, 0.8 per second, and 15 mm/sec for the four–detector row scanner, respectively. All subjects were observed for reactions to contrast material from the start of contrast material injection until the end of the CT examination (about 55 minutes). Two minor reactions occurred: One patient developed mild wheezing and one developed scattered urticaria. These reactions resolved without specific treatment within 15 minutes of onset.

Image Processing and Interpretation

The field of view of the source images was 360–440 mm, with a pixel size of 0.7–0.9 mm. CT scans were reconstructed at a reduced field of view (250–300 mm) and at intervals of 1.25 mm (n = 8, four–detector row CT scanner) or 0.625 mm (n = 54, 16–detector row CT scanner). Volumetric images were then reconstructed on a dedicated three-dimensional workstation (Advantage for Windows 4.0 or 3.1; GE Medical Systems)
Second-Order Biliary Branch Anatomy at CT Cholangiography

<table>
<thead>
<tr>
<th>Second-Order Biliary Branch Anatomy</th>
<th>No. of Subjects (n = 62)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional</td>
<td>35 (56)</td>
</tr>
<tr>
<td>Variant</td>
<td></td>
</tr>
<tr>
<td>Low insertion of right posterior duct onto main duct</td>
<td>8 (13)</td>
</tr>
<tr>
<td>Right posterior duct joining left duct</td>
<td>11 (18)</td>
</tr>
<tr>
<td>Trifurcation</td>
<td>7 (11)</td>
</tr>
<tr>
<td>Right anterior duct joining common hepatic duct</td>
<td>1 (2)</td>
</tr>
</tbody>
</table>

Note.—Numbers in parentheses are percentages.

Figure 1. Oblique coronal volume-rendered CT cholangiogram (image acquired with 1.25-mm section thickness) in a 41-year-old man examined for potential right hepatic lobe donation. Conventional branching anatomy of the bile ducts is seen. In particular, the right posterior branch (second-order branch) is fully visible (arrow).

by using maximum intensity projection and volume-rendered techniques. Two radiologists (B.M.Y. and F.V.C., with 2 and 7 years, respectively, of subspecialty experience with CT of the liver), who were unaware of findings at clinical and other imaging examinations, reviewed by consensus all source and reconstructed CT images on a picture archiving and communication system workstation (Impax; Agfa, Mortsel, Belgium). Although these two radiologists were involved in the interpretation of some of the CT cholangiograms at the time they were obtained, they were unaware of the surgical findings when they were reviewing images for this study. Second-order biliary branches were scored by using a four-point scale, as follows: 0, not seen; 1, faintly seen; 2, well seen but the confluence or a portion of the biliary branch segment is not seen; and 3, excellent visualization from proximal commencement to distal confluence.

Second-order biliary tract anatomy was classified as conventional or variant. In conventional anatomy, the right posterior duct (which drains Couinaud segments VI and VII) joins the right anterior duct (which drains Couinaud segments V and VIII) to form the right hepatic duct, which then joins the left hepatic duct (which is formed by ducts draining Couinaud segments II, III, and IV) (3). All other anatomic configurations were considered variant.

Comparisons

In the 24 subjects who underwent CT cholangiography and right hepatic lobe retrieval, the biliary anatomy shown at CT cholangiography was compared with surgical findings. The comparison was done by one radiologist (Z.J.W.). The number of subjects in whom intraoperative cholangiography was necessary was compared for the 47 patients undergoing right hepatic lobe retrieval before and after the introduction of CT cholangiography.

Statistical Analysis

The proportion of subjects who underwent intraoperative cholangiographic assessment of second-order biliary anatomy was compared for the donors who undergoing right hepatic lobe retrieval before and after the introduction of CT cholangiography by using the Fisher exact test. A P value of less than .05 was considered indicative of a statistically significant difference. Statistical analysis was performed by using Stata software package version 7.0 (Stata, College Station, Tex).

RESULTS

The mean second-order biliary branch visualization score was 2.9 (range, 2–3). Second-order biliary tract branching anatomy determined at CT cholangiography is summarized in the Table. Examples of CT cholangiograms are shown in Figures 1 and 2.

Variant second-order biliary tract anatomy was seen at CT cholangiography in 27 (44%) of the 62 subjects (Table). Of the 24 subjects who underwent right hepatic lobe retrieval, second-order biliary branch anatomy as determined at CT cholangiography was concordant with findings at surgery in 23 (96%) subjects. Twelve (50%) of these 24 subjects had variant second-order biliary branch anatomy seen at CT and confirmed with surgery, and the variants consisted of aberrant insertion of the right posterior duct into the common hepatic duct (n = 6), aberrant insertion of the right posterior duct into the left main duct (n = 3), trifurcation of the common hepatic duct (n = 2), or insertion of an accessory right anterior duct into the common hepatic duct (n = 1).

In one subject, an aberrant right duct insertion was missed at CT cholangiography (Fig 3). Two right-sided second-order biliary branches separated by less than 1 mm of tissue were found in this subject at surgery. The CT cholangiogram had been obtained with 2.5-mm-thick sections.

Among all 47 adult living donors who underwent right hepatic lobe retrieval at our institution since the start of adult living related liver transplantation in January 2000, significantly fewer subjects required conventional intraoperative cholangiography after the introduction of CT cholangiography (three of 24 subjects [12%]) than before the introduction of CT cholangiography (23 of 23 subjects [100%]; P < .0001) (Fig 4). Of the three subjects who underwent intraoperative cholangiography after CT cholangiography, all three had variant second-order biliary tract branching anatomy, and the subtype of anatomic variants was identical as determined at CT and conventional cholangiography. It is at the transplant surgeon’s discretion whether to perform intraoperative cholangiography after CT cholangiography. Two of the three subjects underwent intraoperative cholangiography in November 2001, within 1 month of the introduction of CT cholangiography. The third subject underwent intraoperative cholangiography in November 2002. This subject was determined to have conventional biliary anatomy at the initial interpretation of the CT cholangiogram. However, further review of the CT cholangiogram immediately before surgery revealed variant anatomy with aberrant insertion of the right posterior duct into the left main duct.
DISCUSSION

We found that CT cholangiography enables consistently good-to-excellent visualization (mean second-order biliary branch visualization score of 2.9 on a scale of 0–3) of second-order biliary branches in living liver donor candidates. Among the 24 subjects who underwent right hepatic lobe retrieval and CT cholangiography, second-order biliary branch anatomy as determined with CT cholangiography was concordant with intraoperative findings in 23 (96%) subjects. In the discordant case, an aberrant right biliary branch insertion was missed at CT cholangiography. At surgery, two right-sided second-order bile ducts were found, separated by less than 1 mm of tissue. This biliary tract anatomic variant was not visible at CT cholangiography, possibly because of limited resolution in the z-axis or slight respiratory blurring. The use of thinner sections and shorter scanning times may be useful for evaluating biliary branching variants when bile ducts run close together. Larger patient series, however, are needed to address this important issue.

We also found, by using the Fisher exact test, that the use of conventional intraoperative cholangiography decreased significantly after the introduction of CT cholangiography: Intraoperative cholangiograms were obtained in all 23 (100%) subjects who underwent right hepatic lobe retrieval at our institution before the introduction of CT cholangiography but in only three of 24 (12%) subjects thereafter. In two of those three subjects, intraoperative cholangiography was performed in November 2001 (within 1 month of the introduction of CT cholangiography) and may partially reflect the surgeons’ initial lack of familiarity with this test. In the third subject, intraoperative cholangiography was performed as a result of the discrepancies in the initial interpretation of the CT cholangiogram before surgery. The decrease in the need for intraoperative cholangiography likely reflects the diagnostic quality of CT cholangiography and the increasing familiarity and confidence of our transplant surgeons with this test. It is possible that a reduction in the rate of intraoperative cholangiography decreases the surgery time, potentially improving donor outcome. However, because of improvements in surgical approach and increasing surgical experience with right hepatic lobe retrieval during the study period, we did not evaluate surgery time as an endpoint in this study.

The high prevalence of variant biliary anatomy in our study (27 of 62 subjects [44%] who underwent CT cholangiography and 13 of 24 subjects [54%] who underwent right hepatic lobe retrieval) is in agreement with that in previous reports (3,4) and underscores the importance of preoperative biliary imaging in living donor candidates for right-lobe liver transplantation. Reports of previous studies have described the use of CT cholangiography in the evaluation of living potential liver donors. Cheng et al (9) reported results in 16 living potential liver donors, 10 of whom had confirmation of CT cholangiographic findings with either endoscopic retrograde cholangiography or conventional intraopera-

Figure 2. Images in a 45-year-old woman examined for potential liver donation. (a) Coronal volume-rendered CT cholangiogram (image acquired with 2.5-mm section thickness) shows a right posterior biliary branch (arrow) inserting into the common hepatic duct. (b) Intraoperative cholangiogram shows similar second-order biliary branch (arrow) anatomy.

Figure 3. Coronal maximum intensity projection CT cholangiogram (image acquired with 2.5-mm section thickness) in a 43-year-old man examined for potential liver donation. Although the image shows conventional anatomy, two right-sided biliary branches separated by less than 1 mm of tissue were found at surgery.

Figure 4. Diagram shows the number of living right-lobe liver donors who underwent intraoperative cholangiography before and after the introduction of CT cholangiography at our institution. Before the introduction of CT cholangiography, all 23 donors underwent intraoperative cholangiography for the assessment of second-order biliary tract anatomy. After the introduction of CT cholangiography, only three of 24 donors underwent intraoperative cholangiography.
tive cholangiography. Schroeder et al (10) also reported excellent results in 16 living potential liver donors with multidetector row CT cholangiography, but only four had intraoperative confirmation of CT cholangiographic biliary tract anatomic findings.

In addition to providing a consistent and accurate depiction of biliary tract anatomy, multi–detector row CT cholangiography offers several other advantages in the preoperative evaluation of living liver donor candidates. CT cholangiography is minimally invasive and simple to perform. The image data enable visualization of the biliary anatomy, as well as the relationship between bile ducts and hepatic vasculature, and can be readily reformatted into three-dimensional displays (10). CT angiography and liver parenchyma evaluation can be performed on the same day.

Intravenous cholangiography is rarely performed in North America, owing in part to the perceived high risk of contrast material reactions (28,29). In several recent studies of CT cholangiography, however, minor contrast material reactions were encountered in only 1%–3% of patients, a rate similar to that with conventional intravenous contrast material–enhanced CT (30–33). No major reactions were noted. The low frequency and mild reactions in these studies may reflect the use of slow contrast material infusion rates as well as intravenous administration of diphenhydramine before imaging (12,30). In our series of 62 living donors, only two reactions occurred; both were minor and self-limiting. However, attention to potential contrast material reaction in future studies is important to ensure patient safety.

An alternative technique, oral CT cholangiography, does not appear to be a sufficiently robust alternative to intravenous CT cholangiography for determining second-order biliary tract anatomy (13). In several centers, magnetic resonance (MR) cholangiography has been used in the preoperative evaluation of the biliary tract in potential living liver donors (34–36). The spatial resolution of MR cholangiography is lower than that of CT cholangiography and does not allow for consistent visualization of the biliary system in living liver donors (6,36,37), even when a combination of conventional and excretory MR cholangiography is used (38).

There are several limitations to this study. This was a single-institution retrospective study. Of the 62 subjects who underwent preoperative CT cholangiography, 24 had surgical confirmation of the second-order biliary anatomy; we do not have independent confirmation of biliary branching anatomy in the other subjects. Because of the small number of subjects who underwent imaging with 2.5-mm-thick sections (n = 8), we did not study whether section thickness had an effect on accuracy in cases of closely spaced parallel bile ducts. Future studies with larger patient series are needed to address whether thinner sections and faster scanning may further improve accuracy in such cases.

In conclusion, CT cholangiography accurately depicts biliary tract anatomy in living donor candidates for right hepatic lobe transplantation, and donors who undergo preoperative CT cholangiography are unlikely to need conventional intraoperative cholangiography.

References
Multi–Detector Row CT in Evaluation of 94 Living Renal Donors by Readers with Varied Experience

**PURPOSE:** To retrospectively assess the accuracy of four-section multi–detector row computed tomography (CT) in the evaluation of renal transplant donors when scans are read by one of multiple readers with varied levels of expertise, by using surgery as the reference standard.

**MATERIALS AND METHODS:** This retrospective study was approved by the institutional review board and complied with the Health Insurance Portability and Accountability Act. Informed consent was waived. Between October 1999 and March 2003, 94 renal donors (42 men, 52 women; mean age, 44 years) underwent four-section multi–detector row CT. Unenhanced scanning of the abdomen was performed with 5-mm section thickness and table speed of 15 mm per rotation. Next, 135–150 mL of nonionic iodinated (300 mg/mL) contrast material was injected intravenously at a rate of 4–5 mL/sec. Contrast material–enhanced CT was initiated 20–25 seconds, 65–70 seconds, and 10 minutes after start of injection. Arterial phase scanning was performed with 1.25-mm section thickness and 7.5-mm table speed. Venous and excretory phase scanning was performed with 2.5-mm section thickness and 15-mm table speed. Each scan was evaluated independently by one of 11 readers for renal vascular and ureteral anatomic variants. Findings at CT were compared with those at surgery. Sensitivity and specificity (with 95% confidence intervals) and accuracy of CT were calculated on the basis of presence or absence of variant anatomy at surgery.

**RESULTS:** CT depicted 107 of 114 renal arteries confirmed at surgery; seven accessory arteries were missed in six donor kidneys. CT depicted 95 of 98 renal veins confirmed at surgery. Sensitivity and specificity of CT were 66% and 100%, 75% and 100%, and 50% and 100%, and overall accuracy was 94%, 97%, and 99%, for identification of variant anatomy of renal arteries, veins, and ureters, respectively.

**CONCLUSION:** Multi–detector row CT as the sole imaging technique in the preoperative evaluation of living renal donors is accurate even when images are read by multiple readers with varied levels of expertise.

Renal transplantation is routinely performed in patients with end-stage renal disease. Living kidney donor transplantation is becoming more common because of the scarceness of cadaveric grafts and the ever-increasing demand for kidney transplantation (1,2). Comprehensive imaging evaluation of kidney donor anatomy is crucial for selecting candidates for living renal transplantation and the surgical technique for procuring the renal graft (3–8). This information is indeed crucial for facilitating successful laparoscopic donor nephrectomy, because of the limited field of view available during surgery (3–5,9). It is essential that radiologists be familiar with the surgical need and the surgery itself so that the pertinent anatomic details can be provided at imaging (10).

In most transplantation centers in the United States, computed tomography (CT) is used in the preoperative assessment of renal donors. The introduction of multi–detector row CT has further improved the performance of helical CT, not only with the speed of scanning...
but also with thin-section acquisition and superior-quality two- and three-dimensional images. The purpose of our study was to retrospectively assess the accuracy of four-section multi-detector row CT in the evaluation of renal transplant donors when images are interpreted by one of multiple readers with varied levels of expertise, by using surgery as the reference standard.

MATERIALS AND METHODS

Subjects

This retrospective study was approved by our institutional review board, and informed consent was waived. Our study was in compliance with the Health Insurance Portability and Accountability Act. Our study included 94 renal donors, representing all who were found suitable for kidney donation from the pool of potential donors between October 1999 and March 2003. Electronic reports of the 94 living renal donors who underwent multi-detector row CT were retrieved from the radiology and surgical databases (N.R.). There were 42 men and 52 women aged 19–73 years (mean age, 44 years). All donors underwent nephrectomy.

Imaging Technique

Scanning was performed in the craniocaudal direction with a four-section multi-detector row CT scanner (Lightspeed QX/i; GE Medical Systems, Milwaukee, Wis). No oral contrast material was administered. Unenhanced CT of the abdomen was performed first from vertebral T12 through L5 by using 5-mm section thickness and table speed of 15 mm per rotation. Subsequently, through an 18-gauge cannula placed in an antecubital vein, approximately 135–150 mL of a nonionic contrast material containing 300 mg of iodine per milliliter (iohexol, Isovue 300; Bracco Diagnostics, Princeton, NJ) was injected at a rate of 4–5 mL/sec. Scanning was initiated 20–25 seconds, 65–70 seconds, and 10 minutes after the start of injection to coincide with the arterial phase, venous phase, and excretory phase, respectively. The following technical parameters were selected for each phase of imaging: for arterial phase scanning, section thickness of 1.25 mm and table speed of 7.5 mm per rotation; for venous and excretory phase scanning, section thickness of 2.5 mm and table speed of 15 mm per rotation. Other parameters were kept constant for each phase of scanning, as follows: gantry rotation of 0.8 second, 140 kVp, 200–300 mA, and pitch of 6:1. Source images were reconstructed with 50% overlap.

Image Processing

The reconstructed images were processed with a commercially available workstation (ADW 3.1; GE Medical Systems) in 40-year-old man. Coronal MIP image shows two arteries (arrows) supplying the left kidney.

TABLE 1

Performance of Multi–Detector Row CT in the Evaluation of Donor Renal Arteries

<table>
<thead>
<tr>
<th>Finding at Surgery</th>
<th>Finding at Multi–Detector Row CT Artery</th>
<th>Normal</th>
<th>Variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal artery</td>
<td>76</td>
<td>6</td>
<td>82</td>
</tr>
<tr>
<td>Variant artery</td>
<td>0</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>76</td>
<td>18</td>
<td>94</td>
</tr>
</tbody>
</table>

Note.—Data are numbers of kidneys.

Image Analysis

Source images and the two- and three-dimensional data sets for each of the 94 donors were reviewed and interpreted independently by one of 11 possible readers (with 1–6 years of experience in interpreting abdominal CT scans [“dedicated readers”]) and had more experience in interpreting renal donor CT scans (“non-dedicated readers”) and had worked as active participants with the kidney transplantation team at the weekly transplantation-radiology rounds.

The scans from the 94 patient examinations were not equally distributed among the 11 readers; the number of scans interpreted varied from four to 21 per reader, with an average of 8.5 per reader. CT scans from 18 (19%) of the 94 patient examinations were read by one of the two dedicated readers, and scans from the remaining 76 patients (81%) were read by one of the nine nondedicated readers. A standardized report was generated for each study and included information about renal cysts, renal stones, renal artery fibromuscular dysplasia or other vascular or congenital abnormality, the number of renal arteries supplying each kidney and their location, the distance between the first-order branch of the renal artery and the aorta, and the number and size of accessory arteries. Accessory arteries were categorized as either polar (piercing the kidney directly) or hilar (entering the kidney at the hilum), depending on their course. The polar artery supplies either the upper or lower pole of the kidney. The hilar accessory artery usually arises from the aorta, close to the main artery, and enters the kidney through its hilum. Division of the main renal artery within 2.0 cm from the aorta was recorded as early branching. Likewise, the following information about renal veins and ureters was included in the report: location and number of renal veins; presence of circumaortic or retroaortic anomalies; distance between gonadal, adrenal, and lumbar veins and the inferior vena cava; number of ureters; and any anomaly of the collecting system.

Reference Standard

The period between CT and surgery varied from 6 to 545 days, with an average of 80 days. Findings at surgery were...
used as the reference standard. Donor nephrectomy was performed by multiple transplantation surgeons (n = 7), including one of the authors (D.K.), with 6–30 years of experience in renal transplantation surgery. The surgeons were aware of the CT findings before surgery. Findings at surgery were dictated to generate a standard electronic report that included details of donor anatomy and surgical technique. The surgical report contained information regarding all items evaluated at CT. All findings discovered at surgery and not recorded in the preoperative CT reports were recorded. In addition, any change in surgical management from what was originally planned was recorded in the report. On the basis of the available findings of CT imaging and evidence at surgery, a database was designed and data were compared (N.R.). Discrepant readings of CT scans (in donors in whom surgery showed findings not reported initially at CT) were retrospectively reviewed by one reader (D.V.S.) who was blinded to findings at surgery and at the previous CT reading.

Statistical Analysis

On the basis of the presence or absence of renal vascular and ureteric anomalies at surgery, the sensitivity, specificity, and accuracy of CT were calculated. The exact 95% confidence interval (CI) for the sensitivity and specificity of CT for detection of anatomic variations of renal arteries, veins, and ureters was calculated by using statistical software (SAS, version 8; SAS Institute, Cary, NC). Interobserver agreement was not calculated, because the number of CT scans read by the nondedicated and dedicated readers varied from a minimum of four to a maximum of 21 per reader.

RESULTS

On the basis of the imaging findings, nephrectomy was performed in the left kidney in 77 of the 94 donors and in the right kidney in 17. In these 17 donors, the right kidney was chosen mainly because of the presence of variant anatomy in the left kidney. Nephrectomy was performed with an open flank procedure in 68 of the 94 donors and with a laparoscopic approach in 26. In five of the 68 donors who underwent an open flank procedure, laparoscopy was initially attempted, but the procedure was subsequently converted to open surgery. At CT, small simple renal cysts (with a diameter of less than 2 cm) were present in six of the 94 donor kidneys. None of the selected donor kidneys was noted to have a renal stone, renal artery fibromuscular dysplasia, or any other vascular or congenital abnormality at multi-detector row CT.

Renal Arteries

A total of 107 renal arteries in 94 donor kidneys were depicted at CT (Fig 1). Only 12 kidneys were shown to have more than one artery; 11 of these had two renal arteries (Fig 2), and one had three renal arteries (total number of accessory arteries at CT, 13). At surgery, 114 arteries were identified in 94 kidneys. Seventy-six of the 94 kidneys (81%) had a single artery, and 18 (19%) had more than one artery (Table 1), including two arteries in 16 kidneys and three arteries in two kidneys (total number of accessory arteries at surgery, 20). Seven accessory arteries in six donor kidneys (including one in each of five kidneys and two in one kidney) were initially missed at CT. Four were classified as superior polar arteries, and three, as inferior polar arteries. The sensitivity and specificity of CT for the detection of variant anatomy of renal arteries were 66% (95% CI: 41%, 87%) and 100% (95% CI: 95%, 100%), respectively.

The overall accuracy of CT for the detection of renal arteries was 94% (107 of 114 arteries in 88 of 94 kidneys), which shows good agreement with surgical findings. All seven accessory arteries that were missed at the initial interpretation were confidently identified at retrospective review of the transverse CT scans (Fig 3). On the postprocessed three-dimensional images, only three of the seven missed accessory arteries could be seen. Two missed accessory arteries appeared as early branching on the three-dimensional images because of their close origin and course with the main renal artery from the aorta. Five of seven accessory arteries initially missed at CT measured 2.0–2.5 mm, and two measured 1.5 mm in size. At surgery, four of these seven accessory arteries were anastomosed successfully in the recipient. The remaining three arteries were sacrificed: two, because they were considered too small (1.5-mm diameter) for anastomosis, and one, because it had refractory spasm.

Renal Veins

Ninety-five renal veins were identified at CT in 94 donor kidneys. The anatomic variations recorded in nine of the 94 donor kidneys were as follows: circumaortic left renal vein (Fig 4) in seven kidneys, retroaortic left renal vein in one kidney, and accessory left renal vein (Table 2) in one kidney. At surgery, 98 veins, including 12 with variant anatomy, were identified in 94 kidneys. Of these 94 kidneys, 90 (96%) had one vein each (including seven circumaortic left renal veins and one retroaortic left renal vein) and four (4%) had two veins each. In three donor kidneys, three accessory renal veins (one in each kidney) were not correctly recorded at CT. The sensitivity and specificity of CT for the identification of variant anatomy of renal veins were 75% (95% CI: 43%, 95%) and 100% (95% CI: 96%, 100%), respectively. At surgery, the overall CT findings were concordant in 91 of the 94 kidneys (accuracy, 97%). At the retrospective review of the CT data sets, however, missed accessory renal veins could be confidently seen in all three donors (Fig 5). Of these three missed accessory veins, two were obvious on source images; in one donor, how-
ever, the accessory vein was better appreciated on a coronal three-dimensional MIP image. All three missed accessory renal veins were less than 6 mm in diameter and were ligated during surgery because of the presence of multiple intrarenal venous communications.

**Pyelocalyceal System and Ureters**

Both the pyelocalyceal system and the ureter above the iliac bifurcation were adequately opacified at CT in all 94 donor kidneys. At CT, ureteral anomaly, in the form of complete ureteral duplication, was seen in only one of the 94 donor kidneys (Fig 6). At surgery, however, two of the 94 donor kidneys were noted to have a ureteral anomaly in the form of complete ureteral duplication: one duplicate ureter (the one detected at CT) in a right kidney, and the other in a left kidney. At retrospective review of the CT scans, complete ureteral duplication in the left donor kidney could be seen on source images. The ureter draining the upper pole, however, was not opacified. On the basis of the findings at surgery, pyeloureteral findings at CT were concordant with surgical observations in 93 of the 94 donor kidneys (accuracy, 99%). The sensitivity and specificity of CT in the identification of variant ureteral anatomy were 50% (95% CI: 1%, 99%) and 100% (95% CI: 96%, 100%), respectively.

**Surgical Management**

In five donors, nephrectomy was initially attempted with laparoscopy, but the procedure was converted to open surgery. This change was not due to bleeding in any of these donors. In one donor, the open surgical approach was used to preserve the accessory artery because it was supplying a substantial portion of the kidney. In two donors, this approach was used because there was more than one perforating lumbar vein draining into the left renal vein, and the duplicate lumbar veins rendered laparoscopic ligation technically difficult. In the remaining two donors, the open surgical approach was used because of the difficulty of hilar dissection.

**Data Summary**

CT demonstrated an overall accuracy of 94%, 97%, and 99% in the detection of renal arteries, veins, and ureters, respectively (Table 3). The two dedicated readers who reviewed 18 of the 94 CT scans failed to note the presence of two accessory arteries. Conversely, the nondedicated readers who read 76 scans failed to record the presence of five accessory arteries in four kidneys, one accessory vein in three kidneys, and a duplicate ureter in one kidney. Discrepant reading of CT scans, however, did not result in any segmental or complete loss of graft kidney viability.

**DISCUSSION**

The only curative therapy for patients with end-stage renal disease is renal transplantation. In the past decade, there has been a substantial increase in living kidney transplantations (1–3). It is crucial to evaluate the donor’s vascular anatomy before removal of the donor kidney with any surgical technique, and it is particularly crucial to do so with the laparoscopic approach because of limited exposure during the procedure (6,11). The most crucial data to be gathered before renal transplantation are data about the vascular anatomy of the donor kidney. The anatomic information required before conventional open and minimally invasive surgery in living kidney donors includes the number, length, location, and branching pattern of the renal arteries and the status of the donor kidney and its venous and collecting systems (12–17). Familiarity with the surgeon’s perspective is a prerequisite for the reliable provision of the required details about a kidney donor at imaging. Often, the choice of surgical approach is influenced by the findings at imaging (4,6,7).

In the past, several investigators have used single-detector row helical CT for predicting the renovascular anatomy (9,18), and some have reported good correlation with both catheter angiography and surgery (5,19–22) for images interpreted by experienced readers. Since then, several advances have been made in CT technology, as well as in postprocessing methods. The introduction of multi-detector row CT has enabled an increase in the speed of scanning and in spatial resolution, compared with those achievable with single-detector row helical CT (23). Simultaneously, multi-detector row CT provides greater volume coverage with superior-quality three-di-
radiologic practice at most academic institutions in the United States and the relative shortage of radiologists, it is difficult to have dedicated readers freely available to interpret these focused studies. Therefore, we wanted to study the effect these developments and improved multi-detector row CT scanners have on the performance of multiple readers with varied levels of expertise in the interpretation of kidney donor CT scans.

Our results show that, for the renal arteries, CT findings were concordant with those from surgery in 88 of the 94 donors. Although seven accessory arteries measuring 1.5–2.5 mm in diameter were missed at the initial interpretation, retrospective review of the CT data helped confirm their presence on transverse CT scans alone. This observation emphasizes the importance of careful review of the transverse images along with two- and three-dimensional reconstructions. Likewise, readers accurately identified renal veins in 91 of the 94 donor kidneys at CT. Even though three accessory renal veins in three donor kidneys were initially missed at CT, one accessory vein in each of those kidneys could be confidently seen at retrospective review.

Our results are comparable to those from recently published articles about studies of multi-detector row CT in which motivated expert readers interpreted the scans. For example, Kawamoto et al (24), in their series of 74 donors, reported agreement between CT and surgical findings in reference to renal arteries in 69 of 74 donors (accuracy of 93%, average for three readers; accuracy range, 89%–97%). Accessory renal arteries were missed in four, two, and six kidneys by the first, second, and third readers, respectively. The sensitivity and accuracy of CT in reference to renal vein anomalies were, respectively, 92% and 99% (average for three readers; accuracy range, 96%–100%). Likewise, Kim et al (25) reported that, in their series of 77 renal donors, multi-detector row CT had an overall depiction rate of 98% (89 of 91 arteries and 83 of 85 veins), with sensitivity and specificity of 86% (12 of 14 accessory arteries) and 100% (65 of 65 accessory arteries), respectively, for accessory arteries, and 75% (six of eight accessory veins) and 100% (69 of 69 accessory veins), respectively, for accessory veins.

Our study emphasized the performance of multi-detector row CT as the sole imaging test for the comprehensive evaluation of living kidney donors by multiple readers with varied levels of expertise. Results of our analysis indicate that four-section multi-detector row CT, in comparison with objective observation at surgery, has high diagnostic accuracy. We believe that our experience reflects a true performance of multi-detector row CT in the comprehensive evaluation of renal donor anatomy outside a controlled setting of dedicated or highly experienced readers. The use of a predefined template to systematically record renal vascular and ureteral findings at donor CT angiography can ensure that the pertinent anatomic details are evaluated with imaging. This approach may reduce the number of discrepant interpretations of CT scans, independently of the reader’s experience.

There are few relative risks to CT, one of which is the remote possibility of renal damage associated with the contrast material (26). Likewise, CT also poses the risk of exposure to ionizing radiation, which should be a consideration in a healthy young adult.

Magnetic resonance (MR) angiography is an acceptable alternative in donors with a history of allergy to iodinated contrast material (3,22,27). Gadolinium-enhanced MR angiography in the evaluation of accessory arteries has been shown to have a sensitivity, specificity, and ac-

---

TABLE 3
Sensitivity, Specificity, and Accuracy of Multi–Detector Row CT in the Evaluation of Renovascular and Ureteral Anatomic Variants

<table>
<thead>
<tr>
<th>Variant</th>
<th>Sensitivity (%)</th>
<th>95% CI for Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>95% CI for Specificity (%)</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Renal artery</td>
<td>66 (12/18)</td>
<td>41, 87</td>
<td>100 (76/76)</td>
<td>95, 100</td>
<td>94 (88/94)</td>
</tr>
<tr>
<td>Renal vein</td>
<td>75 (9/12)</td>
<td>43, 95</td>
<td>100 (82/82)</td>
<td>96, 100</td>
<td>97 (91/94)</td>
</tr>
<tr>
<td>Ureter</td>
<td>50 (1/2)</td>
<td>1, 99</td>
<td>100 (92/92)</td>
<td>96, 100</td>
<td>99 (93/94)</td>
</tr>
</tbody>
</table>

Note.—Data in parentheses are numbers of kidneys.
Accuracy of 89%, 94%, and 91%, respectively (27).

There were limitations in our study. First, although each CT scan was reviewed independently by one of 11 readers, 19% of scans were reviewed by two dedicated readers. Second, the number of CT scans interpreted by each reader was not equally distributed. It is conceivable that dedicated readers with interobserver agreement would have provided even better results; at a retrospective review by a dedicated reader, all surgical findings were deemed visible at CT. Furthermore, objective evidence at surgery constituted the reference standard for renal vascular and ureteral anomalies. Moreover, we preferred to select kidneys with a normal anatomy or a less intricate anomaly for donor nephrectomy. Therefore, the performance of multi–detector row CT in the evaluation of more complex vascular and excretory anatomy and anomalies could not be compared. In addition, we used an empirical scanning delay rather than a more accurate bolus timing technique or automated software technique. We, in our experience with CT angiography of the abdomen in healthy adults, as well as investigators in a previous study (24), demonstrated excellent results with this empirical scanning delay. It is conceivable, however, that use of a bolus timing technique or automated scanning software could have provided more consistent opacification of the renal vasculature, which could have improved our performance in the detection of small renal vessels. Finally, modification of CT protocols to generate thinner sections (<1 mm), or use of more than four detector rows with a smaller detector configuration, may improve the detection of small accessory arteries (23).

In conclusion, multi–detector row CT used as the sole imaging technique in the preoperative evaluation of living renal donors provides high accuracy even when images are read by multiple readers with varied levels of expertise.

Acknowledgment: The authors acknowledge the input of Elkan Halpern, PhD, for the statistical analysis.

References
Diffusion-weighted MR Imaging of Kidneys in Healthy Volunteers and Patients with Parenchymal Diseases: Initial Experience¹

**PURPOSE:** To prospectively evaluate feasibility of diffusion-weighted (DW) magnetic resonance (MR) imaging in assessment of renal function in healthy volunteers and patients with various renal abnormalities and to prospectively evaluate reproducibility of DW MR imaging in volunteers.

**MATERIALS AND METHODS:** Study protocol was approved by local ethics committee; informed consent was obtained. Eighteen healthy volunteers and 15 patients underwent transverse fat-saturated echo-planar DW MR imaging of the kidneys during normal breathing. Freehand regions of interest were delineated in the cortex and medulla of the kidneys. The following apparent diffusion coefficient (ADC) values were calculated: ADC of all b values (ADCavg), ADC of low b values (b = 0, 50, 100 sec/mm²; ADClow), and ADC of high b values (b = 500, 750, 1000 sec/mm²; ADChigh). These values were calculated to differentiate influence of perfusion and diffusion. Reproducibility was assessed by repeating the same protocol in five randomly selected volunteers after 6 months. For statistical analysis, Student t tests were used.

**RESULTS:** In all volunteers, ADCavg and ADChigh were significantly higher in the cortex than in the medulla (P < .001). No difference between the cortex and medulla could be observed for ADClow. Patients with renal failure had significantly lower ADCavg (P < .001, P = .004), ADClow (P = .02, P = .03), and ADChigh (P = .02, P = .04) of cortex and medulla, respectively, than did volunteers. In the patient with pyelonephritis, all ADC values of cortex and medulla were substantially lower compared with the contralateral side, whereas patients with ureteral obstruction showed varying degrees of difference in all ADC values compared with the contralateral side. No statistically significant changes were found in the repeat study of the volunteers.

**CONCLUSION:** DW MR imaging is feasible and reproducible in the assessment of renal function, as shown in our initial experience with a small number of patients and volunteers.

Diffusion-weighted (DW) magnetic resonance (MR) imaging is an MR imaging technique used to show molecular diffusion, which is the brownian motion of the spins in biologic tissues (1). The apparent diffusion coefficient (ADC), as a quantitative parameter calculated from the DW MR images, combines the effects of capillary perfusion and water diffusion in the extracellular extravascular space (1). Thus, DW MR imaging provides information on perfusion and diffusion simultaneously in any organ, it can be used to differentiate normal and abnormal structures of tissues better, and it might help in the characterization of various abnormalities.
DW MR imaging is already an established method used routinely at several institutions in the diagnosis of acute stroke (2).

Only in recent years has DW MR imaging been used in extracranial organs; for example, it has been used both to monitor treatment response and tissue characterization and to perform functional evaluation of different organs, such as the parotid glands or kidneys (3–12). DW MR imaging of abdominal organs is much more difficult to perform as a result of physiologic motion artifacts and heterogeneous composition of the organs (10,13). The kidneys are one of the most interesting organs in which to measure ADC values. With their complex anatomic structure and physiology, they are extremely challenging for DW MR imaging. Previous studies generally involved the use of breath-holding sequences (9,11,13–15). This made it difficult to use DW MR imaging in severely ill or dysnephic patients. Furthermore, only a few MR sections were obtained through the kidneys (9,11,13). This harbors the inherent risk of missing small focal lesions.

Thus, the aim of our study was to prospectively evaluate (a) the feasibility of DW MR imaging in the assessment of renal function in healthy volunteers and in patients with various renal abnormalities and (b) the reproducibility of DW MR imaging in volunteers.

MATERIALS AND METHODS

Study Population

Eighteen healthy volunteers (13 men and five women; median age, 27 years; age range, 23–40 years) who had no history of renal disease, hypertension, or other vascular disease and 15 patients (eight men and seven women; median age, 58 years; age range, 24–78 years) with pyelonephritis (n = 1), ureteral obstruction (n = 3), and acute or chronic renal failure (n = 11) were included in our study. All volunteers were medical or paramedical personnel, such as medical students, physicians, and nurses. The patients with diffuse renal abnormalities were accrued consecutively from a group of patients with clinically and histologically proved diagnoses. Those with unilateral abnormalities that were diagnosed with clinical evaluation, intravenous urography, sonography, and laboratory parameters were randomly selected. No specific preparatory measures, such as fasting or drinking, were undertaken prior to the examination. Volunteers were examined within 15 minutes of agreeing to participate in the study. No clinical signs of overhydration (eg, peripheral edema, jugular venous distension, high blood pressure) were noted in the patients.

Serum creatinine values were obtained from all patients on the day of the MR examination.

The study protocol was approved by the local ethics committee, and informed consent was obtained from all volunteers and patients.

MR Imaging

MR imaging was performed with a 1.5-T MR imager (Sonata; Siemens, Erlangen, Germany) with a 40 mT/m maximum gradient capability with a combination of an anterior and a posterior six-channel body coil. Five randomly selected volunteers (three men and two women; median age, 29 years; age range, 25–36 years) underwent a repeat study with the identical protocol after a mean time of 6 months ± 1 (standard deviation) to determine the reproducibility of the DW MR imaging results.

For morphologic evaluation of the kidneys, a coronal T2-weighted half-Fourier rapid acquisition with relaxation enhancement—or RARE—sequence (HASTE; Siemens) was used, with an early and late echo. A transverse T1-weighted fast low-angle shot (or FLASH) gradient-echo sequence was also used, and both in-phase and out-of-phase images were acquired. The early echo T2-weighted images were acquired with the following parameters: section thickness, 4.0 mm; intersection gap, 1 mm; field of view, 400 × 400 mm; matrix size, 220 × 256; one signal acquired; voxel size, 1.8 × 1.6 × 4.0 mm; repetition time msec/echo time msec, 1040/63; partial Fourier factor, 5/8; 20 sections acquired; total acquisition time, 22 seconds. The late-echo T2-weighted images were acquired with identical parameters, except for an echo time of 361 msec and a partial Fourier factor of 7/8, which resulted in a total acquisition time of 23 seconds. For the T1-weighted images, the following parameters were used: section thickness, 5.0 mm; intersection gap, 1 mm; field of view, 380 × 380 mm; matrix size, 179 × 256; one signal acquired; voxel size, 2.1 × 1.5 × 5.0 mm; repetition time, 101 msec; in-phase echo time, 4.76 msec; out-of-phase echo time, 2.38 msec; flip angle, 70°; bandwidth, 350 Hz per pixel; parallel imaging with modified sensitivity encoding (or mSENSE) with a parallel imaging reduction factor of two; number of sections acquired, 25; acquisition time, 20 seconds.

Transverse DW multisection echo-planar MR imaging was performed with the following diffusion gradient b values: 0, 50, 100, 150, 200, 250, 300, 500, 750, and 1000 sec/mm². These were applied in three orthogonal directions and subsequently averaged to minimize the effects of diffusion anisotropy. The following parameters were used for this sequence by applying modified sensitivity encoding: parallel imaging reduction factor of two; 3200/71; section thickness, 5 mm; intersection gap, 1 mm; voxel size, 3.0 × 3.0 × 5.0 mm; matrix size, 128 × 128; field of view, 380 × 380 mm; partial Fourier factor, 6/8; bandwidth, 1502 Hz per pixel; six signals acquired. Fat saturation was used to avoid chemical shift artifacts, and two presaturation slabs were positioned perpendicular to the anterior and posterior sections, respectively, to suppress motion influences. The whole sequence consisted of 25 sections, with an acquisition time of 9 minutes 22 seconds. The study was performed during normal respiration. ADC maps were calculated automatically with the MR system. Calculated ADC values are expressed in square millimeters per second.

Image Analysis

Image analysis was performed off-line at a Linux workstation (Dell, Round Rock, Tex) with dedicated software (Biomap; Novartis, Basel, Switzerland). No major distortion artifacts due to susceptibility or eddy currents or N/2 ghosting artifacts were observed in the DW images of all subjects.

Morphologic evaluation of the images was performed in consensus by two experienced radiologists who were blinded to the DW MR images and data (H.C.T., R.H.O.). The presence and size of parenchymal lesions, focal or diffuse parenchymal signal alteration, corticomedullary differentiation (normal, reduced, absent), peritumoral changes, and dilatation of the collecting system (mild, moderate, marked) were noted.

In the transverse ADC map, circular regions of interest (ROIs) were placed in the cortex and medulla on several sections (upper, middle, and lower poles) in each kidney by means of consensus of two observers (H.C.T., F.D.K.). For each kidney, the ROIs in the cortex were merged into a single ROI, and the same procedure was performed for the medulla. This yielded four ROIs per subject (one ROI for cortex [on average, 3.2

Image analysis was performed off-line at a Linux workstation (Dell, Round Rock, Tex) with dedicated software (Biomap; Novartis, Basel, Switzerland). No major distortion artifacts due to susceptibility or eddy currents or N/2 ghosting artifacts were observed in the DW images of all subjects. Morphologic evaluation of the images was performed in consensus by two experienced radiologists who were blinded to the DW MR images and data (H.C.T., R.H.O.). The presence and size of parenchymal lesions, focal or diffuse parenchymal signal alteration, corticomedullary differentiation (normal, reduced, absent), peritumoral changes, and dilatation of the collecting system (mild, moderate, marked) were noted.
cm² ± 1.1] and one ROI for medulla [on average, 2.7 cm² ± 0.7] for each kidney), from which the average ADC values were calculated. Afterward, these ROIs were copied on the original DW images, from which the average intensities for each b value could be obtained.

In addition, ADCs of the kidneys were calculated separately for the low (ADClow; b = 0, 50, and 100 sec/mm²) and high (ADChigh; b = 500, 750, and 1000 sec/mm²) b values to enable differentiation of the relative influence of the perfusion fraction and true diffusion (1). ADChigh reflects almost only diffusion, whereas ADClow is composed of both diffusion and perfusion (1).

The ADC values were calculated by using a least squares solution of the following system of equations: S(i) = S₀ x exp(-b x ADC), where S(i) is the signal intensity measured on the ith b value image and b is the corresponding b value. S₀ is a variable estimating the exact (without noise induced by the MR measurement) signal intensity for a b value of 0 sec/mm². To reduce the influence of the noise in the measured signal intensity of the original DW MR images in the ADC calculations, the diffusion images obtained with three b values were used (0, 50, and 100 sec/mm² for ADClow; 500, 750, and 1000 sec/mm² for ADChigh) when calculating the ADC values.

### Statistical Analysis

Statistical analysis was performed with the Excel 9.0 (Microsoft, Seattle, Wash) and Analyse-It, version 1.68 (Analyse-It Software, Leeds, England) software packages. The ADC values of the volunteers and patients with renal failure are reported as the mean ± standard deviation.

The group of patients with renal failure was divided into two groups according to their serum creatinine level by using an arbitrary threshold of 2.5 mg/dL (221 μmol/L) to compare diffusion differences at different serum creatinine levels. The serum creatinine level was below this level in 11 patients and above this level in nine.

Statistical analysis was performed by using two-tailed paired Student t tests for the comparison of cortex and medulla in the volunteers. Two-tailed unpaired Student t tests were used when comparing the patients with renal failure with the volunteers. A P value of less than .05 was considered to indicate a statistically significant difference.

For the patients with unilateral abnormalities, the reported values are the percentage differences between the normal and the abnormal kidney. Differences of more than 10% were arbitrarily chosen as substantial.

**RESULTS**

### Morphologic Evaluation

Morphologic evaluation of the kidneys of the volunteers showed no abnormalities, with the exception of a cortical cyst that was 1 cm in diameter and located in the upper pole of the left kidney in one individual.

Of the three patients with ureteral obstruction, two had ureteral calculi, and one had transitional cell carcinoma in the left proximal ureter, with moderate dilatation of the collecting system and some perirenal fluid. One patient had a calculus in the left proximal ureter, with moderate dilatation of the collecting system. The other had a calculus in the right proximal ureter, with only a slightly dilated collecting system, but perirenal fluid surrounded the lower pole because of fornix rupture.

The patient with acute unilateral multifocal pyelonephritis showed slight enlargement of the affected left kidney.

Eleven patients had renal failure of various causes (Table 1). Four patients had normal findings at morphologic analysis of the kidneys (patients 1–4); two of these patients had a fresh hematoma at the lower pole of the left kidney less than 7 days after biopsy (patients 3 and 4). Patient 5 had enlarged kidneys: The left kidney measured 149 × 66 mm, and the right kidney measured 145 × 53 mm. One patient (patient 6) who abused an analgesic (phenacetin nephropathy) had a small irregular right kidney and a shrunken hydronephrotic left kidney; the shrunken left kidney was caused by external compression of the ureter by a huge tumor mass and was therefore not included in DW MR analysis. One patient had undergone nephrectomy on the left side 50 years previously for renal tuberculosis. The right kidney was reduced in size, with reduced corticomedullary differentiation and perirenal stranding. In addition, a 7-cm-diameter cortical cyst at the middle pole and a 1-cm-diameter cyst at the upper pole were visible (patient 7). Another patient had reduced corticomedullary differentiation, perirenal stranding, and a small (1 cm in diameter) cortical cyst at the lower pole on the right side (patient 8).

The patient with Wegener granulomatosis showed loss of corticomedullary differentiation and perirenal stranding and a small hematoma in the posterior paranephric space at the level of the lower pole.

---

**TABLE 1** Characteristics of Patients with Renal Failure

<table>
<thead>
<tr>
<th>Patient No./Sex/Age (y)</th>
<th>Diagnosis</th>
<th>Serum Creatinine Level (mg/dL)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/M/64</td>
<td>Chronic renal insufficiency due to hypercalcemia (bone metastases)</td>
<td>2.12 (0.70–1.30)</td>
</tr>
<tr>
<td>2/F/55</td>
<td>Diabetic nephropathy</td>
<td>2.60 (0.60–1.10)</td>
</tr>
<tr>
<td>3/F/45</td>
<td>Chronic renal insufficiency due to diuretic abuse</td>
<td>2.12 (0.60–1.10)</td>
</tr>
<tr>
<td>4/F/43</td>
<td>Acute glomerulonephritis</td>
<td>1.56 (0.60–1.10)</td>
</tr>
<tr>
<td>5/M/67</td>
<td>Acute renal insufficiency due to interstitial nephritis</td>
<td>5.41 (0.70–1.30)</td>
</tr>
<tr>
<td>6/F/76</td>
<td>Chronic interstitial nephritis due to anagietic abuse</td>
<td>3.41 (0.60–1.10)</td>
</tr>
<tr>
<td>7/M/78</td>
<td>Acute renal failure due to cardiac insufficiency</td>
<td>2.24 (0.70–1.30)</td>
</tr>
<tr>
<td>8/M/65</td>
<td>Intestinal nephritis due to cyclosporine toxicity</td>
<td>1.41 (0.70–1.30)</td>
</tr>
<tr>
<td>9/M/58</td>
<td>Wegener granulomatosis</td>
<td>3.37 (0.70–1.30)</td>
</tr>
<tr>
<td>10/M/45</td>
<td>Acute tubular necrosis</td>
<td>1.95 (0.70–1.30)</td>
</tr>
<tr>
<td>11/M/68</td>
<td>Immunoglobulin A nephritis</td>
<td>4.74 (0.70–1.30)</td>
</tr>
</tbody>
</table>

* Values are from the day of MR examination; data in parentheses indicate the normal range. To convert to Système International units (micromoles per liter), multiply by 88.4.
after biopsy (patient 9). One patient with acute tubular necrosis had loss of corticomedullary differentiation (patient 10). Another patient presented with bilateral hemorrhagic cysts and a decrease in corticomedullary differentiation (patient 11).

**Functional Evaluation**

The group of volunteers used for analysis of reproducibility had ADCavg values of $2.00 \pm 0.07 \times 10^{-3}$ mm$^2$/sec and $(3.73 \pm 0.43) \times 10^{-3}$ mm$^2$/sec, and ADClow values of $(3.40 \pm 0.50) \times 10^{-3}$ mm$^2$/sec and $(1.53 \pm 0.25) \times 10^{-3}$ mm$^2$/sec, respectively. The ADChigh values for the cortex and medulla, respectively, were found when comparing any of the corresponding areas and the different ADC values with a paired Student t test.

In the volunteers, the ADCavg which was calculated from the entire range of $b$ values, was significantly greater in the cortex than in the medulla ($P < .001$), as shown in Table 2.

After separate analysis of ADClow and ADChigh, the ADClow of all volunteers did not show a significant difference ($P = .77$) between the cortex and medulla (Table 2). However, the ADChigh in volunteers was significantly greater in the cortex than in the medulla ($P < .001$) (Table 2). In the volunteers, the ADClow was always found to be greater than the ADChigh in any one region ($P < .001$).

**Patients**

Unilateral abnormality.—The results for the patients with a unilateral abnormality are listed in Table 3. The patient with pyelonephritis had the most substantial differences between the kidneys for both the cortex and the medulla and for all calculated ADC values, which ranged from 22.3% for ADChigh in the medulla to 49.3% for ADClow in the cortex (Fig 2).

For the patients with ureteral obstructions, no substantial differences were found in the ADCavg when compared with the contralateral kidney. The patient with a transitional cell carcinoma in the ureter showed only a substantial difference (25.7%) for ADChigh in the cortex. The patient with a calculus and a fornix rupture showed substantial differences in ADChigh, both in the cortex and in the medulla. The other patient with a calculus also showed a substantial difference in ADChigh in the medulla (18.5%) but not in the cortex (Fig 3). On the other hand, a substantial difference was found in ADClow in the cortex (12.9%).

Bilateral abnormality.—Kidneys in which renal failure was diagnosed ($n = 20$) showed ADCavg values of $(1.82 \pm 0.22) \times 10^{-3}$ mm$^2$/sec in the cortex and $(1.71 \pm 0.21) \times 10^{-3}$ mm$^2$/sec in the medulla, which were significantly less than those of the volunteers ($P < .001$ and $P = .004$, respectively). The ADChigh values for the cortex $(3.31 \pm 0.64) \times 10^{-3}$ mm$^2$/sec and medulla $(3.41 \pm 0.58) \times 10^{-3}$ mm$^2$/sec were also found to be significantly lower when compared with those of the volunteers ($P$ values of .02 and .03, respectively). A similar significance level was present for the ADChigh of both the cortex $(1.53 \pm 0.25) \times 10^{-3}$ mm$^2$/sec and the medulla $(1.38 \pm 0.23) \times 10^{-3}$ mm$^2$/sec in comparison with the volunteer values ($P$ values of .02 and .04, respectively).

When dividing the group of patients with renal failure into two groups according to serum creatinine level (an arbitrary threshold of 2.5 mg/dL [221 µmol/L] was used), the group of patients with a creatinine level lower than this threshold ($n = 11$) had ADC values lower (except the ADChigh in the medulla) than those in volunteers (Table 2). The reported $P$ values for this group were as follows: .03 and .09 for ADCavg of the cortex and medulla, respectively; .37 and .37 for ADChigh of the cortex and medulla, respectively; and .26 and .61 for ADClow of the cortex and medulla, respectively.

The group of patients with a creatinine level higher than the threshold ($n = 9$) showed significant differences from the volunteers for all ADC values, except for ADChigh in the medulla (Table 2). In this group, ADCavg values were $(1.73 \pm 0.24) \times 10^{-3}$ mm$^2$/sec for the cortex ($P = .005$) and $(1.61 \pm 0.26) \times 10^{-3}$ mm$^2$/sec for the medulla ($P = .02$). ADClow and ADChigh were $(3.07 \pm 0.68) \times 10^{-3}$ mm$^2$/sec ($P = .02$) and $(1.44 \pm 0.24) \times 10^{-3}$ mm$^2$/sec ($P = .02$) for the cortex and $(3.10 \pm 0.52) \times 10^{-3}$ mm$^2$/sec ($P = .005$) and $(1.31 \pm 0.26) \times 10^{-3}$ mm$^2$/sec ($P = .06$) for the medulla, respectively (Fig 4).

**DISCUSSION**

In this study, repeated measurements of the kidneys in healthy volunteers showed

---

**Figure 1.** Transverse ADC maps calculated from echo-planar DW MR images (3200/71) with $b$ values between 0 and 1000 sec/mm$^2$ at the midpole of the kidneys in a healthy 30-year-old male volunteer. (a) First ADC map. (b) ADC map obtained at the same level in the same volunteer 7 months later. The similar appearance of the renal parenchyma after a 7-month interval indicates the reproducibility of the DW MR imaging technique.

**Figure 2.** Transverse ADC map calculated from echo-planar DW MR images (3200/71) with $b$ values between 0 and 1000 sec/mm$^2$ at the midpole of the kidneys in a 24-year-old female patient with pyelonephritis of the left kidney. Note the decrease in intensity of medulla and cortex of the left kidney (arrow) compared with the contralateral side (arrow-head).
reproducible results of DW MR imaging. This clarifies an important issue for the validation of this method for follow-up of different abnormalities. DW MR imaging of the kidneys was found to provide information on renal function and be suggestive of the presence and degree of obstruction or inflammation. DW MR imaging was performed without breath holding, thus allowing examination of severely ill, old, or obese patients who were unable to hold their breath for a long time.

In several previously published articles, authors analyzed the feasibility of DW MR imaging of the kidneys by using different technical approaches (10,15,16); however, only a few studies were performed in patients (9,11).

The present study was performed to investigate DW MR imaging of the kidneys in healthy volunteers and patients with various diffuse renal abnormalities in a normal hydration state and without breath holding.

DW MR imaging provides information on diffusion and perfusion at the same time (17). When applying only high $b$ values, the influence of perfusion is largely cancelled out, and the ADC value approximates the true diffusion. Low $b$ values are influenced by both perfusion and diffusion. Thus, calculating the ADC of low and high $b$ values separately provides more specific information on kidney function.

The ADCs of the healthy volunteers calculated with all $b$ values and those calculated with high $b$ values ($b$ value of 500, 750, and 1000 sec/mm²) were significantly higher in the cortex than in the medulla. Similar differences were observed in a study in which diffusion tensor imaging was used and gradients were applied in six directions (14) and in a study performed without breath holding.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Sex</th>
<th>Age (y)</th>
<th>ADCavg (%)*</th>
<th>ADClow (%)†</th>
<th>ADChigh (%)‡</th>
<th>Serum Creatinine Level (mg/dL)§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pyelonephritis</td>
<td>Female</td>
<td>24</td>
<td>44.3</td>
<td>31.7</td>
<td>49.3</td>
<td>38.2</td>
</tr>
<tr>
<td>Transitional cell carcinoma proximal ureter,</td>
<td>Female</td>
<td>61</td>
<td>5.8</td>
<td>-1.7</td>
<td>25.7</td>
<td>-2.7</td>
</tr>
<tr>
<td>moderate dilatation</td>
<td>Female</td>
<td>39</td>
<td>-5.0</td>
<td>-3.4</td>
<td>18.5</td>
<td>28.8</td>
</tr>
<tr>
<td>Calculus proximal ureter, fornix rupture</td>
<td>Female</td>
<td>44</td>
<td>2.2</td>
<td>1.2</td>
<td>-1.8</td>
<td>34.3</td>
</tr>
<tr>
<td>Calculus proximal ureter, mild dilatation</td>
<td>Male</td>
<td>44</td>
<td>2.2</td>
<td>1.2</td>
<td>-1.8</td>
<td>34.3</td>
</tr>
</tbody>
</table>

Note.—ADC data are the difference between the normal and the abnormal kidney. Positive values indicate a lower value for the abnormal kidney compared with the normal kidney, whereas negative values indicate a higher value for the abnormal kidney.

* ADCavg values were calculated from the following $b$ values: 0, 50, 100, 150, 200, 250, 300, 500, 750, and 1000 sec/mm².
† ADClow values were calculated from the following $b$ values: 0, 50, and 100 sec/mm².
‡ ADChigh values were calculated from the following $b$ values: 500, 750, and 1000 sec/mm².
§ Serum creatinine values were acquired at the date of the MR examination; data in parentheses indicate the normal range. To convert to Système International units (micromoles per liter), multiply by 88.4.

**TABLE 2**
ADC of the Kidneys of Healthy Volunteers and Patients with Renal Failure

<table>
<thead>
<tr>
<th>Group</th>
<th>ADCavg ($\times 10^{-4}$ mm²/sec)*</th>
<th>ADClow ($\times 10^{-4}$ mm²/sec)†</th>
<th>ADChigh ($\times 10^{-4}$ mm²/sec)‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volunteers (n = 36)</td>
<td>2.03 ± 0.09</td>
<td>1.87 ± 0.08</td>
<td>1.67 ± 0.11</td>
</tr>
<tr>
<td>Patients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum creatinine level &lt;2.5 mg/dL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(221 μmol/L) (n = 11)</td>
<td>1.90 ± 0.18</td>
<td>1.79 ± 0.14</td>
<td>1.60 ± 0.24</td>
</tr>
<tr>
<td>Serum creatinine level &gt;2.5 mg/dL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(221 μmol/L) (n = 9)</td>
<td>1.73 ± 0.24</td>
<td>1.61 ± 0.26</td>
<td>1.44 ± 0.24</td>
</tr>
</tbody>
</table>

Note.—Data are mean ± standard deviation. Data in parentheses are number of kidneys.

* ADCavg values were calculated from the following $b$ values: 0, 50, 100, 150, 200, 250, 300, 500, 750, and 1000 sec/mm².
† ADClow values were calculated from the following $b$ values: 0, 50, and 100 sec/mm².
‡ ADChigh values were calculated from the following $b$ values: 500, 750, and 1000 sec/mm².
in which gradients were applied in three directions (16). In contrast, studies (9,15) in which b values of up to 198 and 300 sec/mm², respectively, were applied showed the ADC of the medulla to be higher than that of the cortex. In several other studies, no effort was made to differentiate between ADC of the cortex and ADC of the medulla (4,11,16).

When comparing ADClow and ADChigh of the cortex and medulla separately in the current study, a significant difference could only be observed in ADChigh. The difference in ADChigh is probably due to the presence of more free diffusion-inhibiting structures in the medulla than in the cortex. However, the lack of difference between the cortex and medulla in ADClow might be because the effect of higher true diffusion in the cortex is largely cancelled out by the greater anisotropy due to the radial orientation of the structures in the medulla.

A group studying DW MR imaging of the kidneys without breath holding observed high interindividual differences of ADC values in healthy volunteers (16). This finding was attributed to the hydration state. Since no mean age of the subjects was mentioned in this study, however, these differences could also be related to differences in age. Indeed, our results in a relatively homogeneous age group of volunteers did not show large differences, even when we disregarded hydration state.

A study by Müller et al (15) showed a significant decrease in ADC in volunteers who were dehydrated compared with volunteers who were hydrated. Under normal conditions, this may be less important, since hydration status in volunteers did not seem to substantially influence the results. The results were reproducible even after an interval of several months. The standard deviation of the ADC values in our group of volunteers was relatively small.

Most of the previously published studies dealing with DW MR imaging of the kidneys were performed with breath holding and covered only portions of the kidneys (9,11,13,15). In contrast, our data do not show major motion artifacts, the image quality was good (even in obese patients), and the examinations were performed during normal respiration. A pilot study was performed in two volunteers; DW MR imaging without and with pulse triggering (in the same individual) did not show substantial differences in calculated ADC values or image quality (data not shown). Thus, our investigation was performed with free breathing and without pulse triggering. Repetition times in a pulse-triggered sequence change in patients who are nervous, have an irregular pulse, or both. This can change the signal intensities of the DW MR images with a possible negative effect on the accuracy and image quality of the ADC maps. Obtaining images during normal respiration is a major advantage in the clinical routine.

In our study protocol, the entire kidneys were covered with a small section thickness. This allows visualization of smaller focal changes and is clinically important in patients with small renal lesions or in the detection of early vascular changes in patients with transplanted kidneys.

We know from previously published data that the medulla of the kidney is anisotropic because of the radial orientation of its structures (11,12,14). To our knowledge, there is only one study in which diffusion tensor imaging of the kidneys was performed (14). Performing DW MR imaging in only one direction (z axis direction) because of anisotropy in the medulla might lead to a loss of information under certain conditions. This is an important aspect, for instance, in kidney transplants. Thus, our investigation was performed by averaging the three orthogonal directions and has the following advantages. First, ADC fluctuations due to different kidney positions are avoided. Second, DW MR imaging is easier to perform and postprocess than diffusion tensor imaging. Third, most of the chronic parenchymal renal diseases (eg, glomerulonephritis, glomerulosclerosis, interstitial nephritis) originate in the cortex. Anisotropy is not a main issue there; therefore, diffusion tensor imaging does not provide additional information, and averaging of the three orthogonal directions to minimize the influence of anisotropy seems to be justified.

In patients with high serum creatinine values, Fukuda et al (11) observed a decrease in ADC compared with those with normal values. This is confirmed by the results in the present study. Acute and chronic renal failure, as well as renal artery stenosis, also showed decreased ADC values compared with those of healthy volunteers (9). This corresponds to our findings in the group of patients with renal insufficiency. Independent of the underlying abnormality—which causes chronic renal failure—interstitial fibrosis, tubular atrophy, and scarring of glomeruli are the final results (18). This implies a restriction of free water in the extravascular extracellular space and causes a decrease in ADC. However, an overlap of the ADC in the cortex and medulla between patients and volunteers was observed when only patients with serum creatinine levels below 2.5 mg/dL [221 µmol/L] were analyzed.

A study performed in pigs with renal artery stenosis or ureteral obstruction also showed a decrease in ADC (15). Our investigation in three patients with ureteral obstruction confirmed a decrease in ADClow of variable extent. The ADCavg showed only slight changes in all three patients, and the ADChigh changed to a varying extent. Thus, differentiation of ADCavg, ADClow, and ADChigh may allow more subtle analysis of the different abnormalities. These findings are likely to be attributed to differing degrees and durations of the obstruction. This could be important for the clinical care of the patient. Mechanical obstruction to urinary outflow causes a rise in luminal pressure and dilation of the proximal collecting system and consequent increased interstitial pressure in the cortex and medulla, as well as decreased renal blood flow, which is reflected in a decrease in ADClow. However, further investigations in a larger series are necessary to quantify the degree of obstruction.

In the patient with acute pyelonephritis, a lower ADC (ADCavg, ADClow, and ADChigh) value was observed in the cortex and medulla when compared with the opposite side. This corresponds to zones of inflammation involving the papilla and cortex (18). A combination of conventional MR imaging and DW MR imaging provides information on morphologic and functional changes at the same time. With functional evaluation, DW MR imaging is able to depict even early alterations (9).

Future applications for DW MR imaging of the kidneys might include children with nephritis, follow-up of patients undergoing chemotherapy with nephrotoxic agents or immunosuppressive drugs, and follow-up of patients with glomerulonephritis or tubulointerstitial disease.

There are limitations of our study. First, the number of patients with the same abnormality is small. To analyze different abnormalities in more detail, larger series must be conducted, possibly with histopathologic correlation.

Second, our volunteers are of a homogeneous younger age group, whereas the patients have a wider range of ages. Thus, no age matching between volunteers and patients could be performed. This is important because the ADC of the kidneys probably decreases slightly over the years.
In conclusion, DW MR imaging of the kidneys gives reproducible noninvasive information on renal function in healthy volunteers, and it is also feasible in severely ill patients. It may provide information as to the degree of kidney dysfunction, assist in the differentiation of various renal abnormalities, and be applicable for follow-up of patients with various abnormal conditions. However, larger scale studies are needed for confirmation.

References
Treated Ovarian Cancer: MR Imaging, Laparotomy Reassessment, and Serum CA-125 Values Compared with Clinical Outcome at 1 Year

**PURPOSE:** To compare retrospectively the use of magnetic resonance (MR) imaging, laparotomy reassessment, and serum CA-125 values in predicting the presence of residual tumor in women who have been treated for ovarian cancer.

**MATERIALS AND METHODS:** This study was approved by the institutional review board, and informed consent was waived. The study was compliant with the Health Insurance Portability and Accountability Act. Seventy-six women (mean age, 59 years) with treated ovarian cancer underwent preoperative MR imaging of the abdomen and pelvis with intravenous gadolinium-based and intraluminal barium contrast material. MR findings were compared with surgical and histopathologic findings, serial and static serum CA-125 values, and clinical follow-up results. Tumor absence was proved with normal surgical results and by following up patients for at least 1 year, with no evidence of residual tumor at serial CA-125 analysis or subsequent laparotomy. McNemar test for correlated proportions was used for statistical analysis.

**RESULTS:** Sixty-eight women had residual tumor proved at laparotomy and biopsy or at clinical follow-up. Eight patients had no evidence of residual tumor. Gadolinium-enhanced MR imaging depicted residual tumor in 61 patients (sensitivity, 90%; specificity, 88%; accuracy, 89%) compared with laparotomy, which demonstrated residual tumor in 60 patients (sensitivity, 88%; specificity, 100%; accuracy, 89%) and CA-125 values, which demonstrated residual tumor in 44 patients (sensitivity, 65%; specificity, 88%; accuracy, 67%) (P < .01). The positive predictive values for MR imaging, laparotomy, and serum CA-125 values were 98%, 100%, and 98%, respectively, whereas the corresponding negative predictive values were 50%, 50%, and 23%, respectively. In 14 patients, there was a discrepancy between the results of MR imaging and those of laparotomy. In seven patients, MR imaging depicted residual tumor that was not found at laparotomy but was proved at subsequent biopsy or clinical and imaging follow-up, with an increasing serum CA-125 level. In six patients, MR findings were normal, and subsequent laparotomy revealed small-volume residual tumor. Residual tumor was incorrectly predicted with MR imaging in one patient who had no surgical or clinical evidence of residual tumor for 1 year.

**CONCLUSION:** Gadolinium-enhanced spoiled gradient-echo MR imaging depicts residual tumor in women with treated ovarian cancer, with an accuracy, positive predictive value, and negative predictive value that are comparable to those of laparotomy and superior to those of serum CA-125 values alone.

Epithelial ovarian cancer is the most common gynecologic malignancy and is the fifth most frequent cause of cancer-related death in women. In the United States, there were an estimated 25,400 new cases of ovarian cancer in 2003 (1,2). Despite progress in treatment,
Serious ovarian cancer remains the leading cause of gynecologic cancer deaths, accounting for an estimated 14,300 deaths in 2003 (3). Initial diagnosis and treatment with surgical staging and tumor debulking, followed by platinum-based chemotherapy, form the foundation of patient care for women with ovarian cancer.

After patients receive treatment, critical clinical decisions require an accurate assessment of tumor response to initial treatment. The presence of residual tumor dictates the use of additional consolidative chemotherapy; patients with complete clinical response will be entered into a program of close clinical and imaging surveillance (2–4). Subsequent tumor recurrence is treated with salvage chemotherapy. Repeat clinical, imaging, or surgical assessment is then performed to confirm or exclude the presence of residual tumor. The assessment of residual tumor following treatment for primary ovarian cancer or recurrent ovarian cancer is a clinical and imaging challenge that regularly confronts oncologists, radiologists, and surgeons.

Serial measurements of serum CA-125 values obtained during chemotherapy correlate with tumor response to therapy in 80% of women who are seropositive for ovarian cancer. A declining tumor marker correlates with tumor response, whereas a persistently elevated tumor marker is a strong indicator of residual tumor (5–10). The positive predictive value of an elevated serum CA-125 value is nearly 100%. Unfortunately, the sensitivity of serum CA-125 value is poor, and the negative predictive value of a normal CA-125 value is low. Findings from multiple studies have confirmed that a serum CA-125 value that is within the normal range does not exclude residual tumor. Makar et al (6) evaluated 208 patients undergoing laparotomy and found that the sensitivity of the serum CA-125 value in predicting residual tumor was 58%, with a negative predictive value of 43%. Clearly, decisions regarding the complete clinical response of ovarian cancer to chemotherapy cannot be based on serum CA-125 values alone.

Laparotomy reassessment following chemotherapy is now less commonly performed at many institutions (11–14). In prior reports, researchers have confirmed that many patients with a normal laparotomy result will eventually develop recurrent tumor, presumably from the progression of a subclinical tumor that was not found at the time of reassessment (12–14). The value of repeat laparotomy and surgical cytoreduction in changing long-term survival has also been questioned. In 1995, the National Institutes of Health Consensus Conference on Ovarian Cancer recommended that laparotomy be limited to patients in clinical trials or to patients in whom the results of laparotomy would affect clinical management (4).

The role of helical computed tomography (CT) and magnetic resonance (MR) imaging in evaluating patients with primary ovarian cancer and patients who have been treated for ovarian cancer has been described (15–25). At our institution, MR imaging has been used to follow up all patients with ovarian cancer after initial chemotherapy and to assess recurrence. To determine tumor response to therapy and to make management decisions regarding the need for additional treatment, our oncologists now routinely use information from MR imaging combined with serial CA-125 values. We undertook this study to compare retrospectively the use of MR imaging, laparotomy, and serum CA-125 values in predicting the presence of residual tumor in women with treated ovarian cancer.

**MATERIALS AND METHODS**

This retrospective study was approved by our Institutional Review Board, and informed consent was waived. Our study was compliant with the Health Insurance Portability and Accountability Act. Identifying patient data were removed from study records to protect patient confidentiality.

**Patients**

Eighty-one consecutive women with treated ovarian cancer who underwent MR imaging and subsequent laparotomy reassessment between 1994 and 2003 were included in this study. All patients subsequently underwent surgical exploitation and restaging within 6 weeks of MR imaging. Five patients whose serum CA-125 levels were negative for ovarian cancer at the time of diagnosis were excluded from the study. The remaining 76 women form the patient group for this retrospective study. Twenty-six patients (mean age, 59 years; age range, 34–88 years) had been included in a prior study comparing MR imaging results with serial CA-125 values in women with ovarian cancer (21). All 76 patients were previously identified as having epithelial ovarian cancer of the following histologic subtypes: undifferentiated (n = 25), papillary (n = 11), serous (n = 14), papillary serous (n = 18), mucinous (n = 2), endometrioid (n = 4), and clear cell (n = 2) adenocarcinoma.

In this retrospective review, a power analysis was not performed. Sample sizes were determined by including all consecutive patients with ovarian cancer who underwent concurrent MR imaging and subsequent laparotomy reassessment at our institution between 1994 and 2003. The starting point of the study coincided with the initiation of MR imaging techniques described later. The markedly decreased use of laparotomy reassessment in patients with ovarian cancer by the end of 2003 effectively determined the end point of the study.

Staging at the time of initial diagnosis prior to therapy was used to establish stage I ovarian cancer in no patients, stage II ovarian cancer in five patients, stage III ovarian cancer in 64 patients, and stage IV ovarian cancer in seven patients. All 76 patients were evaluated for the presence of residual tumor following treatment. Fifty-three patients (mean age, 59 years; age range, 32–79 years) were evaluated for residual tumor following surgical and/or chemotherapeutic cytoreduction of primary or persistent ovarian cancer. The mean time interval between primary surgery and MR imaging in these patients was 6 months. Twenty-three patients (mean age, 55 years; age range, 34–88 years) whose ovarian cancer had been in clinical remission were evaluated for residual tumor following treatment for recurrent ovarian cancer. The mean time interval between primary surgery and MR imaging in these patients was 40 months.

At the time of MR imaging and laparotomy, patients were evaluated to determine the clinical response to therapy and to establish the need for additional consolidative chemotherapy or intraperitoneal chemotherapy. The patient’s oncologist, as part of the routine clinical evaluation, ordered all MR imaging examinations. Preoperative MR imaging was routinely performed to assess the volume and location of the residual tumor. This information was used for presurgical planning, to direct biopsies, and to establish maximal response to chemotherapy prior to surgical cytoreduction or intraperitoneal chemotherapy. This article includes the results of MR imaging performed immediately prior to laparotomy reassessment. The rationale for subsequent laparotomy reassessment was to perform surgical cytoreduction, to obtain histopathologic proof of residual tumor,
to lyse adhesions, or to perform intraperitoneal chemotherapy. In these 76 patients, the results of MR imaging were not used to direct management to not perform laparotomy reassessment.

MR Imaging

MR imaging of the abdomen and pelvis was performed by using the body coil on a 1.5-T imager (Signa; GE Medical Systems, Milwaukee, Wis) equipped with high performance gradients (23 mT/m, 120 [mT · m⁻¹]/sec). For this study, 900–1350 mL of dilute barium sulfate (Readi-Cat 2; E-Z-Em, Westbury, NY) was used as an oral contrast agent in all patients. Patients were instructed to drink one bottle (450 mL) of contrast material every 20 minutes, starting 40 minutes to 1 hour before MR imaging. A rectal enema (500–1000 mL) was used to distend the colon in all patients.

For the transverse and sagittal T2-weighted fast spin-echo sequence, either respiratory-triggered (n = 22) or breath-hold (n = 54) images were acquired. Respiratory-triggered fast spin-echo imaging parameters were as follows: 4600/96 (repetition time msec/echo time msec), 256 × 256 matrix, two signals acquired, echo train length of eight, ±32-kHz receiver bandwidth, 7-mm section thickness, and 3-mm intersection gap. Breath-hold fast spin-echo imaging parameters were as follows: 2500/94, 256 × 192 matrix, one signal acquired, echo train length of 17, and ±32-kHz receiver bandwidth. For all fast spin-echo acquisitions, flow compensation and fat saturation were used for artifact reduction, and zero interpolation was used to increase the number of pixels to 512 in the frequency direction. For respiratory-triggered images, acquisition time ranged from 3 minutes 45 seconds to 5 minutes 0 second for each set of 24 sections. For breath-hold images, acquisition time was 24 seconds for each set of 12 images.

Dynamic gadolinium-enhanced fat-suppressed two-dimensional spoiled gradient-echo imaging was performed during suspended respiration following rapid bolus administration of 0.2 mmol of gadolinium chelate (gadodiamide, Omniscan, Nycomed, Princeton, NJ; gadopentetate dimeglumine, Magnevist, Berlex, Wayne, NJ; or gadoversetamide, Optima 200, Mallinckrodt, St Louis, Mo) per kilogram body weight. Transverse spoiled gradient-echo images were obtained immediately after the intravenous injection of gadolinium chelate, with additional delayed transverse images obtained at 3–5 minutes. Coronal fat-suppressed spoiled gradient-echo images were obtained prior to the delayed transverse images in all patients. Imaging parameters included 140–165/1.9–2.6, a 512–256 × 192 matrix with a three-quarter rectangular field of view, one signal acquired, a section thickness of 8 mm with no intersection gap, a receiver bandwidth of ±16–20 kHz, and a flip angle of 70°. Noninterleaved sets of 12 sections were obtained during each 22–24-second breath hold. The time for the entire MR examination by using current breath-hold T2-weighted MR imaging and gadolinium-enhanced spoiled gradient-echo MR imaging was 20 minutes.

Interpretation of MR Images and Comparison with Findings from Laparotomy Reassessment

All MR images were prospectively interpreted by one of three radiologists (including R.N.L.), each with 12–14 years experience in MR imaging of the body. It is our policy to always interpret MR images without knowledge of current CA-125 values. Unenhanced fast spin-echo T2-weighted MR images and fat-suppressed gadolinium-enhanced spoiled gradient-echo MR images were evaluated for the presence of residual tumor masses in the abdomen and pelvis that could represent either local tumor recurrence or distant metastases. The gadolinium-enhanced spoiled gradient-echo MR images were evaluated for peritoneal thickening and abnormal enhancement that was more intense than that of the liver parenchyma. Enhancement is most prominent on delayed gadolinium-enhanced images and may be smooth and continuous or nodular and discontinuous in appearance. Serosal tumor was noted if there was mural thickening of the bowel wall and if mural enhancement was greater than that of the liver parenchyma. The presence of lymphadenopathy and ascites was also noted on unenhanced fast spin-echo T2-weighted MR images and gadolinium-enhanced spoiled gradient-echo MR images.

One author (R.N.L.) reviewed the prospective MR image interpretation, as recorded in the written report for each study. Patients whose MR reports described definite or probable residual tumor were categorized as having residual tumor. Patients whose MR reports described no residual tumor or equivocal findings were categorized as not having residual tumor.

The anatomic location of the residual tumor, as described in the written reports, was then compared (R.N.L.) with the location of the residual tumor that was confirmed at laparotomy and histopathologic evaluation or at subsequent biopsy.

Laparotomy

Exploration of the abdomen and pelvis was performed in all patients following MR imaging. Seventy laparotomies were performed by oncologic surgeons (B.D., R.M.B.), and six laparotomies were performed by general surgeons. During laparotomy, visual inspection and palpation of all abdominal organs, peritoneal surfaces, and bowel serosa were performed. Biopsy was performed on all suspicious nodules and masses, and tissue samples were sent for histopathologic evaluation. Random peritoneal biopsies in the abdomen and pelvis were also performed according to standard surgical technique for laparotomy. Peritoneal washings were obtained for cytologic evaluation.

Serum CA-125 Values

One author (R.N.L.) reviewed each patient’s medical charts. Serum CA-125 levels were recorded, with the immediate preoperative serum CA-125 value noted for each patient. Serum CA-125 levels less than 35 U/mL were considered normal. Serum CA-125 levels greater than or equal to 35 U/mL were considered elevated. For those patients with a normal laparotomy result, continued surveillance of serial CA-125 values was performed for at least 1 year. All 76 patients were seropositive for ovarian cancer at the time of initial diagnosis for ovarian cancer.

Definitions

The diagnosis of residual tumor was made on the basis of surgical findings, if positive, or by using a combination of postoperative biopsy findings, clinical follow-up results, and serial CA-125 values. Tumor presence was recorded if (a) results from laparotomy and histopathologic evaluation confirmed residual tumor in the abdomen or pelvis, (b) serum CA-125 levels were persistently elevated on serial evaluations, or (c) there was clinical or subsequent surgical or histopathologic proof of residual tumor during the 12 months following a negative laparotomy result. Clinical evidence of residual tumor included an elevated and increasing serum CA-125 level that was at least double that of the initial baseline.
value or the presence of a tumor that was palpable at physical examination.

The use of a persistently elevated serum CA-125 level as proof of tumor is supported by findings from prior studies, which show that a persistently elevated serum CA-125 level of greater than 35 U/mL is nearly 100% predictive of residual or recurrent tumor in women who were initially seropositive for epithelial ovarian cancer (5–9). Follow-up MR imaging was not used by itself to prove residual tumor. In patients with a false-negative laparotomy result for which subsequent clinical or biopsy evidence proved residual tumor within 12 months, follow-up MR imaging results were used to confirm that the location of the progressive tumor agreed with that of the residual tumor predicted at initial MR imaging.

Tumor absence was recorded if laparotomy and histopathologic evaluation showed no evidence of residual tumor and if clinical follow-up during the subsequent 12 months showed no clinical evidence of residual tumor during serial CA-125 analysis and physical examination.

Data and Statistical Analysis

On the basis of these definitions, findings from each MR examination, laparotomy, and preoperative serum CA-125 analysis were recorded as true-positive, false-negative, true-negative, or false-positive.

True-positive MR findings were recorded if MR images showed a tumor mass or enhancing peritoneal tumors and if the surgical and histopathologic results confirmed residual tumor at the same anatomic location. In patients with a normal laparotomy result who subsequently developed evidence of residual tumor, true-positive MR findings were recorded if patients had (a) an increasing serum CA-125 level that was at least double that of the initial baseline value at serial evaluations, (b) a palpable mass, and/or (c) subsequent biopsy results that confirmed residual tumor at the anatomic site predicted at MR imaging. In the absence of tissue diagnosis, findings from follow-up MR imaging were used to confirm progressive tumor at the site predicted at initial MR imaging.

True-positive laparotomy results were recorded (R.N.L.) if findings from histopathologic evaluation confirmed residual tumor that was surgically identified in the abdomen or pelvis. True-positive serum CA-125 values were recorded if the values were greater than or equal to 35 U/mL.

False-negative MR findings and serum CA-125 values were recorded if MR images showed no evidence of residual tumor and if the serum CA-125 value was less than 35 U/mL, but the patient either had residual tumor proved at laparotomy or had developed evidence of residual tumor during the subsequent 12 months. False-negative laparotomy results were recorded if surgical findings and results from histopathologic evaluation showed no residual tumor, but the patient later developed clinical evidence of residual tumor, with an elevated and increasing (twofold or greater increase) serum CA-125 value, or if residual tumor was proved at subsequent laparotomy or biopsy.

False-positive MR findings were recorded if MR images were interpreted as showing residual tumor, but results from laparotomy reassessment, serum CA-125 analysis, and clinical follow-up for 12 months were negative. A false-positive serum CA-125 value was recorded if a transiently elevated CA-125 level returned to normal without any treatment following a normal laparotomy result.

Normal findings from MR imaging and laparotomy and normal serum CA-125 values were accepted as true-negative only after a 12-month disease-free interval without additional treatment, as was determined at clinical follow-up. Follow-up imaging results were also evaluated to confirm the absence of tumor for 12 months.

On the basis of these definitions, the sensitivity, specificity, accuracy, positive predictive value, and negative predictive value for MR imaging, laparotomy, and serum CA-125 values were determined in predicting the presence of residual tumor in women with treated ovarian cancer were calculated. Combinations of these three tests were similarly evaluated. The diameter of the largest tumor was determined from surgical and histopathologic reports. Tumors were categorized as microscopic, small-volume (<1 cm), moderate (1–2 cm), or bulky (>2 cm). A tumor was classified as microscopic if no tumor was noted by the surgeon following visual inspection and palpation, but a tumor was confirmed at histopathologic evaluation. The diameter of the largest tumor at MR imaging was determined from the written report by using the same size categories.

For patients with a normal laparotomy result who developed residual tumor within 12 months, tumor size was determined on the basis of a review of the tumor size indicated on concurrent MR images. If residual tumor was not seen at laparotomy or MR imaging but developed within 12 months, the tumor was categorized as microscopic. The sensitivity, specificity, and accuracy of MR imaging, laparotomy, and serum CA-125 values were also determined according to the size of the residual tumor.

Statistical Analysis

The sensitivity and accuracy of MR imaging, laparotomy, and serum CA-125 values were compared by using the McNemar test of correlated proportions (INSTAT 2.0; Graph Pad Software, San Diego, Calif). For this application of the McNemar test, data were paired for each patient, comparing MR imaging and laparotomy, MR imaging, and serum CA-125 values. Two-tailed P values were reported, with the null hypothesis rejected for P values less than .05.

RESULTS

Sensitivity, Accuracy, and Predictive Values for Results in All Patients

Residual tumor was proved at laparotomy, biopsy, or clinical follow-up in 68 women (mean age, 59 years; age range, 32–88 years). Eight patients had no evidence of tumor at laparotomy or for the 12 months following laparotomy. These eight patients had normal serial CA-125 values for 12 months and normal physical examination results. Seven of the eight patients also underwent follow-up MR imaging, which showed no evidence of residual tumor.

Gadolinium-enhanced MR images correctly depicted residual tumor in 61 patients (Fig 1), with seven false-negative interpretations, one false-positive interpretation, and seven true-negative interpretations (sensitivity, 90%; specificity, 88%; accuracy, 89%; positive predictive value, 98%; and negative predictive value, 50%). For laparotomy, there were 60 true-positive findings of residual tumor, eight false-negative findings of residual tumor, and eight true-negative findings of residual tumor (sensitivity, 88%; specificity, 100%; accuracy, 89%; positive predictive value, 100%; and negative predictive value, 50%). For serum CA-125 analysis, there were 44 true-positive results, 24 false-negative results, seven true-negative results, and one false-positive result (sensitivity, 65%; specificity, 88%; accuracy, 67%; positive predic-
Two of the three patients whose MR reports were equivocal for residual tumor were later shown to have subcentimeter tumor at laparotomy. These equivocal MR reports were classified as false-negative interpretations.

The eight false-negative laparotomy findings occurred in five patients with microscopic tumor, two patients with small-volume (<1 cm) tumor, and one patient with moderate (1–2 cm) residual tumor. In four patients, residual tumor was revealed during subsequent tissue diagnosis at surgical biopsy in one patient, at CT-guided biopsy combined with elevated CA-125 value at the time of laparotomy in two patients, and at CT-guided biopsy combined with a doubling of serum CA-125 value in one patient. Residual tumor was proved in the remaining four patients, with at least a doubling of serum CA-125 levels during serial testing and with follow-up MR imaging results showing progressive tumor at the site predicted on preoperative MR images. Two of these latter four patients also developed a palpable mass that was identified by their oncologist. In the eight patients with a false-negative laparotomy result, MR imaging enabled correct prediction of tumor in seven patients, and CA-125 analysis enabled correct prediction of tumor in five patients. In no patient were the results from laparotomy, MR imaging, and CA-125 analysis all incorrect.

Thirty-two patients were in clinical remission at the time of the MR examination and had a normal serum CA-125 value and no palpable tumor (Fig. 1). Residual tumor was proved in 25 patients at laparotomy exploration or clinical follow-up. In 20 (80%) of the 25 patients, MR imaging correctly depicted residual tumor compared with laparotomy, which demonstrated tumor in 21 (84%) of the 25 patients ($P > .05$). Both MR imaging and laparotomy were superior to serum CA-125 analysis ($P < .001$, McNemar test).

**Patients Evaluated for Residual Tumor after Treatment of Recurrent Ovarian Cancer**

Twenty-three patients whose disease had been in clinical remission were evaluated for response of recurrent tumor to treatment. Twenty-two of these patients had residual tumor proved at laparotomy and histopathologic evaluation or at subsequent follow-up. Gadolinium-enhanced MR imaging correctly depicted residual tumor in 20 patients, with two false-negative interpretations, no false-positive interpretations, and one true-negative interpretation (sensitivity, 91%; specificity, 100%; accuracy, 91%; positive predictive value, 100%; and negative predictive value, 33%). For laparotomy, there were 20 true-positive results, six true-negative results, and one false-positive result (sensitivity, 91%; specificity, 100%; accuracy, 91%; positive predictive value, 100%; and negative predictive value, 33%). For serum CA-125 analysis, there were 29 true-positive results, 17 false-negative results, and one false-positive result (sensitivity, 63%; specificity, 86%; accuracy, 66%; positive predictive value, 97%; and negative predictive value, 26%). There was no significant difference between the sensitivity, accuracy, positive predictive value, or negative predictive value of MR imaging and those of laparotomy ($P > .05$). The sensitivity and accuracy of MR imaging and laparotomy were superior to those of serum CA-125 analysis ($P < .01$, McNemar test).
**Anatomic Location and Size of Tumor**

In 55 of 68 patients with tumor, the presence and anatomic location of the tumor on MR images agreed with the location of the tumor at laparotomy and histopathologic evaluation. This was determined by comparing the sites of tumor described in the MR reports with the sites of tumor confirmed in the surgical and histopathologic evaluation reports. In seven patients, MR imaging depicted tumor while surgical results were normal (Figs 2, 3). In three of these patients, tumor was subsequently proved at biopsy. In four patients, tumor was subsequently proved at clinical and MR imaging follow-up, which revealed an increasing serum CA-125 value and progressive tumor at the site predicted at initial MR imaging. In six patients, MR findings were normal, and subsequent laparotomy revealed small-volume residual tumor. In one patient, small-volume tumor was incorrectly predicted with MR imaging in a patient who had no evidence of tumor at laparotomy or clinical follow-up for 1 year (Fig 4).

In general, the tumor size noted at MR imaging correlated with the tumor size at laparotomy. Seven patients had microscopic tumor, 17 had small-volume (<1 cm) tumor, 15 had moderate (1–2 cm) tumor, and 29 had bulky (>2 cm) tumor. The Table compares MR imaging, laparotomy, and serum CA-125 findings for the detection of residual ovarian cancer according to tumor size. There was no significant difference between MR imaging, laparotomy reassessment, and serum CA-125 values for any of the tumor sizes ($P > .05$, McNemar test). For the 24 patients with microscopic or small-volume (<1 cm) residual tumor, gadolinium-enhanced MR imaging correctly demonstrated residual tumor in 19 patients (79%) compared with laparotomy, which correctly demonstrated residual tumor in 17 patients (71%), and serum CA-125 values, which correctly demonstrated residual tumor in nine patients (38%). In the 44 patients with moderate (1–2 cm) or bulky (>2 cm) residual tumor, gadolinium-enhanced MR imaging correctly demonstrated tumor in 42 patients (95%) compared with laparotomy, which correctly demonstrated tumor in 43 patients (98%), and CA-125 values, which correctly demonstrated tumor in 35 patients (80%).

**Combinations of Tests**

Combinations of tests provided complementary information and improved the accuracy in predicting residual tumor. The combination of MR imaging and serum CA-125 values enabled correct prediction of residual tumor in 65 of 68 patients, with two false-positive interpretations and sensitivity of 96%, specificity of 96%, accuracy of 79%, and positive predictive value of 93%. The positive predictive value for the combination of MR imaging and serum CA-125 value was 97%, and the negative predictive value for the combination of MR imaging and serum CA-125 value was 67%. In comparison, the combination of laparotomy and CA-125 values enabled correct prediction of residual tumor in 65 of 68 patients, with one false-positive result and sensitivity of 96%, specificity of 88%, accuracy of 95%, positive predictive value of 98%, and negative predictive value of 70%. The combination of MR imaging and laparotomy enabled correct predic-
treated for ovarian cancer. The implication of this statement is reflected in the practice patterns our oncologists, who now routinely use findings from MR imaging to determine the need for additional consolidative or salvage chemotherapy in women with treated ovarian cancer. In practice, the information from serial CA-125 analysis and serial MR examinations is used to determine the degree of tumor response to chemotherapy.

Currently, patients with normal MR imaging results and normal serum CA-125 values undergo clinical and imaging surveillance. Patients with abnormal MR imaging results and elevated serum CA-125 values undergo consolidative chemotherapy. For patients in whom there is a discrepancy between MR imaging results and serum CA-125 values, the oncologist may initiate consolidative chemotherapy if the MR findings or clinical presentation is compelling for the presence of residual tumor; the patient may also undergo close interval follow-up. We currently perform serial MR examinations every 3–4 months in patients with treated ovarian cancer. In practice, patients with definite tumor depicted at MR imaging are treated with consolidative chemotherapy even if the serum CA-125 value is normal.

This approach to the care of women with treated ovarian cancer has helped to fill the information void created by the reduced use of follow-up laparotomy. While laparotomy reassessment is still used in selected patients for the lysis of adhesions or for intraperitoneal chemotheraphy, reassessment is no longer routinely employed in all patients with ovarian cancer. Because a normal CA-125 value is not useful in excluding residual tumor following treatment (5–7), MR imaging has assumed a critical role in patient care decisions. In our study, the combination of MR imaging and serum CA-125 values was more sensitive and accurate than serum CA-125 values alone in predicting the presence of residual tumor.

The use of MR imaging in lieu of laparotomy to follow up patients with treated ovarian cancer has several advantages. Patients can be spared the morbidity associated with laparotomy if treatment decisions are based on accurate cross-sectional imaging. The cost of combined serial CA-125 evaluation and MR imaging is considerably less than the cost associated with laparotomy. In patients who undergo surgical exploration, preoperative MR imaging can add valuable information by facilitating the assessment of tumor volume and by helping to direct biopsies in patients with minimal disease.

To be used in this manner, MR imaging must be able to distinguish patients who are in complete remission from those who have small-volume residual tumor. Both the positive predictive value and negative predictive value of the test must be high to ensure that correct treatment decisions are being made. The positive predictive value for MR imaging and laparotomy reassessment was high (98% and 100%, respectively). The negative predictive value of MR imaging and laparotomy, however, was only 50% for either procedure, but this value was still superior to that of serum CA-125 evaluation (negative predictive value, 22%).

In a prior study of patients with treated ovarian cancer, we compared the role of MR imaging with that of serial serum CA-125 evaluation and physical examination (21). In this earlier study (21), the relatively small number of patients who underwent both laparotomy reassessment and MR imaging precluded any meaningful comparison. In the current study, our conclusions are based on data from patients who were imaged at our institution within a 10-year time period.

To our knowledge, this study is the first to demonstrate that findings from cross-sectional imaging can approximate those of laparotomy reassessment in patients with ovarian cancer.

Gadolinium-enhanced MR imaging has been shown to be sensitive in depicting subtle peritoneal tumor and carcinomatosis (20–22,24). The superior contrast resolution of MR imaging makes the enhancing peritoneal tumors very conspicuous, especially on delayed gadolinium-enhanced images obtained 5 minutes following a double-dose injection of gadolinium chelate. In the past, the depiction of centimeter peritoneal metastases has been a limitation of cross-sectional imaging. Our MR technique used a section thickness of 7–8 mm. Thinner sections acquired with three-dimensional gradient-echo sequences may be useful but were not available for this study. In our experience, however, patients with peritoneal carcinomatosis often have thin sheets of diffuse peritoneal tumor that are easily depicted on fat-suppressed gadolinium-enhanced spoiled gradient-echo images. Subcentimeter peritoneal tumors are routinely depicted on gadolinium-enhanced MR images by using a section thickness of 7–8 mm.

At most institutions, helical CT is the

**DISCUSSION**

Findings from this study illustrate that by using currently available technology, sensitivity of gadolinium-enhanced MR imaging can equal the sensitivity of laparotomy reassessment in predicting residual tumor in women who had been

---

**Figure 4.** Transverse gadolinium-enhanced spoiled gradient-echo MR image (165/2.1, 70° flip angle) of 61-year-old woman with treated ovarian cancer, acute bowel obstruction, and normal serum CA-125 value. Image shows 2-cm enhancing soft-tissue mass (arrow) at point of bowel obstruction. Surgical and histopathologic findings demonstrate fibrotic mass and adhesions without evidence of tumor. MR images were interpreted as false-positive for tumor recurrence.
imaging modality used in patients with ovarian cancer. Helical CT is typically faster, more readily available, and more familiar to radiologists and oncologists. A comparison of helical CT and MR imaging was not the purpose of this study. The role of helical CT in ovarian cancer, however, has been reported \((15,16,25)\). Moderate or bulky tumor is certainly well depicted on helical CT scans. Small-volume tumor and carcinomatosis can be more challenging to distinguish on CT scans because CT has a more limited soft-tissue contrast. In a study of 64 patients with ovarian cancer, Coakley et al \((25)\) noted that the sensitivity of helical CT in depicting peritoneal metastases 1 cm or smaller was only 25%–50%; this number was significantly less than the overall sensitivity of helical CT in depicting peritoneal metastases, which was 85%–93%. In patients with treated ovarian cancer, the ability to accurately depict small residual tumors is essential for making clinical decisions regarding the need for additional chemotherapy.

For all of its strengths, the MR technique also has limitations that should be understood. Peritoneal and serosal thickening and enhancement are not specific for peritoneal tumor. Patients with a postoperative course complicated by abdominal abscesses and fistulas will show a similar pattern of peritoneal and serosal thickening and enhancement that is indistinguishable from tumor. In our experience, patients with a routine postoperative course do not show residual peritoneal or serosal enhancement on follow-up MR images. Patients who undergo peritoneal stripping and intraperitoneal chemotherapy may also show a similar pattern of peritoneal thickening and enhancement. While the distinction is incomplete, peritoneal masses that are nodular, irregular, confluent, or have more diffuse peritoneal thickening and enhancement are more likely to represent residual tumor. In practice, we also look for progressive changes in peritoneal disease on serial MR images.

Limitations of our study should be acknowledged. We used a persistently elevated serum CA-125 value on serial measurements to prove tumor recurrence in patients with a negative laparotomy result. This is certainly valid when assessing the accuracy of laparotomy or MR imaging in demonstrating tumor presence or absence. The use of an elevated serum CA-125 value to assess the accuracy of the same tumor marker, however, may seem circular. In these cases, we also based the determination of tumor recurrence on the oncologist’s clinical impression, combining physical examination results, serial CA-125 values, and, in some cases, biopsy or aspiration results and histopathologic findings.

Because this was a retrospective review of the initial MR interpretations, we cannot state definitively that the radiologist was never aware of the patient’s serum CA-125 value. To our knowledge, in 13 years of clinical practice at our institution, this information has not been avail-

<table>
<thead>
<tr>
<th>Tumor Size</th>
<th>MR Imaging</th>
<th>Laparotomy</th>
<th>Serum CA-125 Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microscopic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>True-positive</td>
<td>6</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>False-negative</td>
<td>1</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>86</td>
<td>29</td>
<td>43</td>
</tr>
<tr>
<td>Small volume</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>True-positive</td>
<td>13</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>False-negative</td>
<td>4</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>76</td>
<td>88</td>
<td>35</td>
</tr>
<tr>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>True-positive</td>
<td>13</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>False-negative</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>87</td>
<td>93</td>
<td>82</td>
</tr>
<tr>
<td>Bulky</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>True-positive</td>
<td>29</td>
<td>29</td>
<td>24</td>
</tr>
<tr>
<td>False-negative</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>100</td>
<td>100</td>
<td>83</td>
</tr>
</tbody>
</table>

Note.—Values are the number of patients in whom each of the three tests correctly predicted the presence of residual ovarian cancer according to tumor size. A tumor was classified as microscopic if no tumor was noted by the surgeon following visual inspection and palpation, but a tumor was confirmed at histopathologic evaluation. Small-volume tumors measured <1 cm in diameter, moderate tumors measured 1–2 cm in diameter, and bulky tumors measured >2 cm in diameter.
able at the time of the MR examination. We purposely do not ask for this information because a normal CA-125 value is of little use in excluding tumor and may incorrectly influence the interpretation of MR images. The three prospective interpreters of the MR images were all highly experienced in reading MR images of the body. The use of their prospective written interpretations of the MR images reflected their impressions prior to laparotomy reassessment. This approach avoided any bias or knowledge that might have occurred in a retrospective review of the same cases.

Selection bias may have occurred because the oncologist used the information from MR imaging to determine which patients would undergo laparotomy reassessment. This likely occurred later in the study because our oncologists gained confidence in MR imaging in patients with ovarian cancer. Patients with definite residual tumor may not have undergone laparotomy reassessment but were treated with consolidative chemotherapy. Certainly, the role of laparotomy in ovarian cancer changed during the course of our study. Early in our study, laparotomy reassessment was performed routinely, whereas now reassessment is performed only in selected cases. Because the information from MR imaging was available to the surgeon who was performing the operation, these MR results may have guided the surgeon in verifying the presence of disease. Finally, we found no difference between MR imaging and laparotomy reassessment in predicting residual tumor for the 76 patients in our study. Given the small differences between MR imaging and laparotomy, however, a much larger patient group may have revealed differences between the two tests.

In conclusion, MR imaging with gadolinium enhancement provides important clinical information in patients with treated ovarian cancer. The results of MR imaging are comparable to those of laparotomy reassessment in depicting residual tumor. Patients with MR images showing residual tumor undergo consolorative chemotherapy. Those patients with normal MR images and normal serum CA-125 values undergo close interval follow-up. At our institution, routine laparotomy reassessment is no longer performed. The results of MR imaging combined with serum CA-125 values are now used in lieu of laparotomy reassessment to assess tumor response to chemotherapy and to direct clinical management decisions.

References
Subclavian Steal Syndrome: Diagnosis with Perfusion Metrics from Contrast-enhanced MR Angiographic Bolus-timing Examination—Initial Experience

PURPOSE: To retrospectively determine whether differential temporal changes in signal intensity of the vertebral arteries, measured at a bolus-timing examination with a test dose of a gadolinium-based contrast agent, are present in patients with subclavian steal syndrome.

MATERIALS AND METHODS: Institutional review board exemption was obtained, and informed consent was not required for this retrospective study. The study complied with the Health Insurance Portability and Accountability Act. Twenty-five patients with known or clinically suspected atherosclerotic disease of the aortic arch and branch vessels underwent breath-hold contrast material-enhanced magnetic resonance (MR) angiography with circulation time derived from a timing examination by using a test bolus of a gadolinium-based contrast agent. Eight patients (three men and five women aged 54–80 years; mean, 70 years) had subclavian stenosis or occlusion with retrograde vertebral artery flow confirmed with time-of-flight MR angiography, nine patients (eight men and one woman aged 31–91 years; mean, 70 years) had mild to severe ostial stenosis of a single vertebral artery, and eight patients (including four men and four women aged 53–86 years; mean, 73 years) had normal vertebral arteries. The difference in time to peak signal intensity between the right and left vertebral arteries was compared among the three groups by using Fisher exact and Cochran–Mantel-Haenszel tests.

RESULTS: The delay in peak enhancement in the ipsilateral vertebral artery ranged from 2 to 4 seconds (mean, 2.5 seconds) in all eight patients with subclavian steal syndrome. In eight of nine patients with ostial vertebral artery stenosis and eight controls, both vertebral arteries filled simultaneously. The presence of unilateral delayed vertebral artery enhancement was significantly associated with retrograde flow in patients with subclavian steal syndrome, compared with patients with normal flow ($P < .01$) and those with ostial vertebral artery stenosis ($P < .01$).

CONCLUSION: A bolus-timing examination performed with a test bolus of the gadolinium-based contrast agent via the neck vessels that demonstrates at least a 2-second delay in peak contrast enhancement in the right or left vertebral arteries may, in the appropriate clinical setting, indicate subclavian steal syndrome.

The term subclavian steal describes a vascular disorder in which occlusion or stenosis of the subclavian artery or the brachiocephalic trunk proximal to the vertebral artery origin causes altered vascular hemodynamics that result in low-velocity and/or retrograde flow in the ipsilateral vertebral artery distal to the subclavian artery narrowing (1–5). The vertebral...
artery effectively steals blood from the posterior cerebral circulation. Subclavian steal syndrome may be manifested clinically by arm claudication or hand numbness and a decrease of at least 20 mm Hg in blood pressure in the upper limb on the affected side, symptoms that may be exacerbated by exercise of the ipsilateral upper limb. Cerebral symptoms may involve nonhemispheric regions of the brain and may lead to dizziness, vertigo, and visual disturbances.

Historically, the reference standard for the definitive diagnosis of subclavian steal syndrome was demonstration of vertebral artery blood flow reversal at invasive angiography, which can also depict narrowing in the subclavian artery or the brachiocephalic trunk. Noninvasive continuous-wave Doppler ultrasonography (US) of the neck can help to determine the direction of flow in the vertebral artery but cannot reliably image the proximal intrathoracic subclavian artery (4–6). In addition, evaluation of the vertebral arteries with duplex US may be difficult in obese patients and requires skilled technicians (7,8).

Several studies have shown the utility of magnetic resonance (MR) angiography for the effective and noninvasive diagnosis of subclavian steal syndrome (7,9–12). Several MR angiography techniques that have been described in the literature can demonstrate the reversal of flow in a vertebral artery. These include two-dimensional time-of-flight MR imaging with a selective presaturation pulse applied first above and then below the volume of interest. Phase-contrast MR imaging can also be used to detect reversal of flow, but the images may be degraded by susceptibility effects or aliasing artifact. More recently, three-dimensional contrast material–enhanced MR angiography has been used to evaluate the proximal aortic arch vessels as well as the vertebral arteries (2). However, while this technique can be used to evaluate the vertebral arteries for stenosis and occlusion, it cannot be used to directly determine flow direction.

A bolus-timing examination with a test-bolus injection is widely used to optimize contrast-enhanced MR angiography, is easy to interpret, and is vendor specific (13,14). When performing MR angiography in the transverse plane at the level of the neck vessels, we noticed a delay in peak vertebral artery enhancement, compared with that in the contralateral vertebral artery, in patients with occlusive disease of the subclavian artery or the brachiocephalic trunk. We hypothesized that this delay might be due to retrograde vertebral artery flow. Thus, the purpose of our study was to retrospectively determine whether differential temporal changes in signal intensity of the vertebral arteries, measured at a bolus-timing examination with a test dose of a gadolinium-based contrast agent, are present in subclavian steal syndrome.

**MATERIALS AND METHODS**

**Patients**

We retrospectively searched an MR database at our institution by using the terms subclavian steal, vertebral artery stenosis, and normal aortic arch, which yielded records of 30, 68, and 138 patients, respectively. Twenty-two of 30 patients with subclavian steal syndrome were excluded because bolus-timing examination was performed at a level inferior to the origins of the vertebral arteries (at the level of the aortic arch). This left a total of eight patients, three men and five women aged 54–80 years (mean, 70 years), in whom subclavian artery stenosis or occlusion was depicted at three-dimensional contrast-enhanced MR angiography with ipsilateral retrograde vertebral artery flow and confirmed at two-dimensional time-of-flight MR angiography. Of these eight patients, one had mild (<50%) stenosis, four had severe (>75%) stenosis, and three had occlusion of the subclavian artery determined with electronic caliper measurements at prospective image interpretation. Another nine consecutive patients, including eight men and one woman aged 31–91 years (mean, 70 years), who had unilateral vertebral artery stenosis and bilateral antegrade vertebral artery blood flow demonstrated at two-dimensional time-of-flight MR angiography, were selected as a control group. Of these nine patients, two had mild (<50%) stenosis, four had moderate (50%–75%) stenosis, and three had severe (>75%) stenosis of a single vertebral artery. Another group of eight consecutive patients, including four men and four women aged 53–86 years (mean, 73 years) who had no lesion in any of the aortic arch or proximal cervical vessels, were selected as the normal-aortic-arch control group; four of these patients, however, had unilateral internal carotid artery disease. Age distributions according to sex are reported in Table 1. Institutional review board exemption was obtained, and informed consent was not required for the retrospective study. Our study complied with the Health Insurance Portability and Accountability Act.

**Bolus-timing Examinations**

All examinations were performed by using a 1.5-T MR system (Sonata or Symphony; Siemens, Erlangen, Germany) with a 300–600-msec rise time and 30–40 mT/m maximum gradient strength. After informed patient consent was obtained, an intravenous catheter was placed in the patient’s right arm, and a phased-array quadrature body coil was positioned to include the chest and neck.

After a localizer sequence was applied, a timing examination to measure circulation time from the antecubital vein to the common carotid artery was performed with breath holding at end expiration. In each patient, a 1-mL test bolus of gadopentetate dimeglumine (Magnevist; Berlex Laboratories, Wayne, NJ), followed by 20 mL saline, was infused at 2 mL/sec by using an MR-compatible injector (Spectris MR Injector SBT 200; MEDRAD, Pittsburgh, Pa). During the infusion of contrast material and the saline flush, a

---

**TABLE 1**

**Patient Age and Delay in Time to Peak Signal Intensity as a Function of Sex and Clinical Status**

<table>
<thead>
<tr>
<th>Location of Delay</th>
<th>Patient Group</th>
<th>No. in Group</th>
<th>Mean Age (y)</th>
<th>Vertebral Artery</th>
<th>Common Carotid Artery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>10</td>
<td>65.7 ± 15.7</td>
<td>5 (50)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>15</td>
<td>74.5 ± 8.0</td>
<td>4 (27)</td>
<td>3 (20)</td>
<td></td>
</tr>
<tr>
<td>Clinical status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal vertebral arteries (controls)</td>
<td>8</td>
<td>73.0 ± 10.2</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Vertebral artery stenosis</td>
<td>9</td>
<td>69.8 ± 16.9</td>
<td>1 (11)</td>
<td>2 (22)</td>
<td></td>
</tr>
<tr>
<td>Subclavian steal</td>
<td>8</td>
<td>70.4 ± 8.6</td>
<td>8 (100)</td>
<td>1 (12)</td>
<td></td>
</tr>
</tbody>
</table>

Note.—Unless otherwise indicated, data are numbers of patients. Numbers in parentheses are percentages.
transverse magnetization-prepared turbo fast low-angle shot sequence (repetition time msec/echo time msec/inversion time msec, 500/1.3–4.2/260–300; flip angle, 20°–25°; section thickness, 10 mm) was applied in a single location at the middle level in the neck, below the carotid bifurcation, every second for 30–60 seconds. The field of view and matrix were 30 cm and 128 × 256, respectively. The turbo fast low-angle shot sequence was used because unenhanced blood yields low signal intensity, thereby minimizing flow-related enhancement. Data acquisition for the bolus-timing examination commenced at the same time. Breath-hold contrast-enhanced three-dimensional MR angiography of the aortic arch and neck vessels was subsequently performed during end expiration, by using the circulation time derived from the bolus-timing examination.

Analysis of Images Obtained during Bolus Timing

To determine the time to peak arterial contrast enhancement, circular regions of interest were manually drawn over both the common carotid and vertebral arteries by a single reader (C.W., 1 year of experience with MR angiography) who was not involved in any of the clinical studies. The regions of interest were of a fixed size (10 pixels) and shape (circle). An additional region of interest was placed over skeletal muscle in the neck to serve as background.

Data and Statistical Analysis

Signal intensity values that were measured and reported automatically by the workstation software supplied by the MR imager manufacturer (Syngo; Siemens) were recorded for the entire duration of the timing examination. Signal intensity values for the arteries were divided by the value for the adjacent muscle, to obtain values that were corrected for background. The MR angiographic data recorded for every subject consisted of the time to peak signal intensity (TP) in the left and right vertebral arteries (denoted as TPVL and TPVR, respectively) and in the left and right common carotid arteries (TPCL and TPCR, respectively). Delay in time to peak signal intensity between vertebral arteries (DV) was calculated for each subject as DV = TPVL – TPVR, and delay in time to peak signal intensity between carotid arteries (DC), as DC = TPCL – TPCR. For the eight patients with subclavian steal syndrome, the recorded data also included the peak signal intensities achieved in the vertebral arteries ipsilateral and contralateral to the subclavian artery narrowing (PL1, PL2). Relative peak signal intensity (RPI) in the vertebral artery ipsilateral to the stenosed subclavian artery was calculated as a percentage of the peak signal intensity in the contralateral vertebral artery, as RPI = (PL1/PL2) · 100.

The Mann-Whitney test was used to assess age differences between the sexes or between patients with a delay in time to peak signal intensity and patients without such a delay. Sex differences with respect to vertebral and carotid artery delays to peak enhancement were assessed by using both the Mann-Whitney and the Fisher exact tests: The Mann-Whitney test was used to test actual values of delays in the vertebral and carotid arteries; the Fisher exact test was used to assess results in the vertebral and carotid arteries that were encoded as 0 (negative delay) or as greater than 0 (positive delay). Subject groups defined in terms of clinical status (presence of subclavian steal syndrome, presence of vertebral artery stenosis, or control) were compared with respect to distribution of the sexes by using the Fisher exact test and with respect to age distribution by using the Kruskal-Wallis test. There was no significant association of subject age or sex with either clinical status or delay in time to peak enhancement. However, because the association of sex with clinical status approached statistical significance, even though the study had low statistical power to detect differences according to sex, the analyses to assess the association of clinical status with delay in time to peak signal intensity were conducted both with and without control for the potential confounding effect of sex. Specifically, the Fisher exact test and the Cochran-Mantel-Haenszel test were used to evaluate the data for an association between clinical status and positive delay in time to peak signal intensity without adjustment for sex and with adjustment for sex, respectively. Spearman rank correlation coefficients were used to assess the associations between delay in time to peak enhancement in the vertebral arteries and time to peak enhancement values measured in the ipsilateral and contralateral vertebral arteries, relative peak signal intensity, and subject age in the group of patients with subclavian steal syndrome. All statistical analyses were conducted by using software (SAS for Windows, version 9.0; SAS Institute, Cary, NC). Each reported P value is two sided and is preceded by the name of the statistical test used to obtain it. Results were declared statistically significant with a two-sided significance level (α) of .05 (ie, P < .05).

RESULTS

Time to Peak Enhancement

In all eight patients with subclavian steal syndrome, a discrepancy in time to peak signal intensity was found between the left and right vertebral arteries (Table 2). The average delay in vertebral artery peak enhancement in patients with sub-

<p>| TABLE 2 | Results of Bolus-timing Examination in Vertebral Arteries in Patients with Subclavian Steal |
|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>Patient No./Age (y)</th>
<th>Side of Subclavian Artery Narrowing</th>
<th>Delay in Time to Peak Signal Intensity (sec)*</th>
<th>Time to Peak Signal Intensity (sec)</th>
<th>Peak Signal Intensity(^{\text{a}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/71 Right</td>
<td>2</td>
<td>8</td>
<td>3.16</td>
<td>1.72 (\text{LVA}) 54.4</td>
</tr>
<tr>
<td>2/77 Left</td>
<td>2</td>
<td>15</td>
<td>2.53 (\text{LVA}) 3.35</td>
<td>74.5</td>
</tr>
<tr>
<td>3/78 Left</td>
<td>2</td>
<td>16</td>
<td>2.37 (\text{LVA}) 4.04</td>
<td>58.7</td>
</tr>
<tr>
<td>4/67 Left</td>
<td>2</td>
<td>19</td>
<td>1.99 (\text{LVA}) 2.50</td>
<td>79.6</td>
</tr>
<tr>
<td>5/64 Left</td>
<td>2</td>
<td>17</td>
<td>4.09 (\text{LVA}) 2.74</td>
<td>67.0</td>
</tr>
<tr>
<td>6/54 Right</td>
<td>3</td>
<td>8</td>
<td>2.37 (\text{LVA}) 3.74</td>
<td>63.4</td>
</tr>
<tr>
<td>7/72 Left</td>
<td>4</td>
<td>23</td>
<td>2.28 (\text{LVA}) 2.57</td>
<td>88.7</td>
</tr>
</tbody>
</table>

* Delay in time to peak signal intensity was measured in the vertebral artery ipsilateral to subclavian artery narrowing.
\(\text{a}\) Data are ratios of peak signal intensity in the artery to that in adjacent muscle.
\(\text{b}\) Relative peak signal intensity in the ipsilateral vertebral artery was calculated as a percentage of peak signal intensity in the contralateral vertebral artery.
\(\text{c}\) Vessel with delay in time to peak signal intensity.
clavian steal syndrome was 2.5 seconds (range, 2–4 seconds), and the delay was always seen in the vessel ipsilateral to the diseased subclavian artery (Fig 1, Table 2). In one of the nine patients in the group with vertebral artery stenosis, a difference in time to peak enhancement was found between the vertebral arteries (Table 3). In this patient, there was a 1-second delay in peak signal intensity in the right vertebral artery compared with the time to peak signal intensity in the left vertebral artery, which was moderately stenosed. None of the eight patients in the normal-aortic-arch (control) group had a delay in peak enhancement in a vertebral artery (Fig 2).

The results of timing examinations with regard to delay in peak enhancement in the common carotid artery are detailed in Table 4. One of the eight patients in the group with subclavian steal syndrome had a 1-second delay in peak signal intensity. Two of the nine patients in the group with vertebral artery stenosis had a difference in time to peak enhancement in the carotid artery, and in both of these patients, the duration of the delay was 1 second. None of the eight patients in the control group had a delay in peak enhancement in a carotid artery.

Table 1 indicates the percentage of patients who tested positive for a delay in peak enhancement of the vertebral or carotid arteries, according to sex and clinical status. With respect to age distribution, there were no significant differences between the sexes (Mann-Whitney test, \( P = .27 \)), between patients with positive delay and those with negative delay (Mann-Whitney test, \( P = .32 \) and \( .18 \) for vertebral arteries and carotid arteries, respectively), or among subject groups defined according to clinical status (Kruskal-Wallis test, \( P = .84 \)). There were no significant differences between the sexes with regard to delay in peak contrast enhancement either in the vertebral arteries (Mann-Whitney test, \( P = .30 \); Fisher exact test, \( P = .40 \)) or in the carotid arteries (Mann-Whitney test, \( P = .14 \); Fisher exact test, \( P = .25 \)). There were no significant differences in sex distribution among the patient groups according to clinical status (Fisher exact test, \( P = .08 \)). Both with and without adjustment for sex, there were statistically significant differences among the clinical status groups with respect to delay in peak enhancement in vertebral arteries (Fisher exact test and Cochran–Mantel-Haenszel test, \( P < .001 \)), but no significant clinical status group differences were found with respect to delay in peak enhancement in the carotid arteries (Fisher exact test, \( P < .75 \); Cochran–Mantel-Haenszel test, \( P = .51 \)). The results of post hoc comparisons showed that the incidence of vertebral artery delay was significantly higher among patients with subclavian steal syndrome than among patients with vertebral artery stenosis (Fisher exact test, \( P < .001 \); Cochran–Mantel-Haenszel test, \( P = .001 \)) or control subjects (Fisher exact test, \( P < .001 \); Cochran–Mantel-Haenszel test, \( P = .002 \)).

Correlations

In the group of patients with subclavian steal, there were no significant correlations (Spearman correlation, \( P > .28 \)) between patient age and delay in peak enhancement in vertebral arteries (\( r = 0.192 \)), relative peak signal intensity (\( r = \text{Figure 1. Left-sided subclavian steal in 67-year-old woman. (a) Thin maximum intensity projection coronal image from three-dimensional contrast-enhanced MR angiography of aortic arch and great vessels demonstrates occlusion (arrow) of the proximal left subclavian artery and a normal-appearing left vertebral artery (arrowhead) that originates from the left subclavian artery. (b) Transverse image from two-dimensional time-of-flight MR angiography (27/6.5; flip angle, 35°) of the neck vessels, with a presaturation band placed above the volume of interest, shows normal signal intensity in the common carotid arteries (arrowheads) and right vertebral artery (long arrow). There is no signal in the left vertebral artery (short arrow), a finding that indicates either occlusion or retrograde flow. (c) Transverse image from two-dimensional time-of-flight MR angiography (27/6.5; flip angle, 35°) of the neck vessels, with a presaturation band placed below the volume of interest, shows normal signal intensity of blood flowing in the internal jugular veins (arrowheads). As expected, there is no signal in the right vertebral artery (long arrow), whereas signal in the left vertebral artery (short arrow) indicates retrograde flow. Note that the signal in the left vertebral artery is weaker than that in the right vertebral artery in b (Fig 1 continues).
0.152), or time to peak enhancement in the contralateral ($r = 0.339$) and ipsilateral ($r = 0.436$) vertebral arteries. In six of eight patients with subclavian steal, there was a concomitant decrease of more than 20% (range, 20.4%-45.6%) in peak signal intensity in vertebral arteries in the affected side, compared with that in the contralateral vertebral arteries (parvus tardus phenomenon) (Table 2, Fig 1e).

**DISCUSSION**

Gadolinium-based contrast-enhanced three-dimensional MR angiography is widely performed for simultaneous evaluation of the aortic arch vessels and carotid arteries (13-18). Methods used to optimize selective arterial enhancement include a timing examination (13,14), MR fluoroscopy (17), automated bolus detection schemes (19), and time-resolved imaging (20). The benefit of a bolus-timing examination (13,14) is the ability to selectively enhance the aortic arch vessels and carotid arteries, while minimizing artifact due to extraluminal enhancement. This is particularly important in the setting of subclavian steal, where the peak signal intensity in the vertebral arteries is decreased in the affected side.

**Figure 1 (continued).** (d) Transverse image from bolus-timing examination (500/1.8/260; flip angle, 20°) in the neck vessels, acquired at the time of peak signal intensity in the right vertebral artery ($t = 16$ seconds), shows no contrast enhancement of the left vertebral artery (short arrow), while both common carotid arteries (arrowheads) and the right vertebral artery (long arrow) are enhanced. (e) Transverse image from bolus-timing examination in the neck vessels, obtained 2 seconds later than d ($t = 18$ seconds), shows delayed contrast enhancement in the left vertebral artery (short arrow) and residual enhancement in the common carotid arteries (arrowheads) and right vertebral artery (long arrow). Note that the left vertebral artery is enhanced to a lesser degree than is the right vertebral artery in d. (f) Signal intensity–time curve for vertebral artery enhancement, derived from bolus-timing examination, shows 2-second delay in peak arterial enhancement in the left vertebral artery (LVA) ($t = 18$ seconds) compared with that in the right vertebral artery (RVA) ($t = 16$ seconds). Peak signal intensity in the left vertebral artery is also decreased (parvus tardus) when compared with that in the right vertebral artery in c and e.
lus-timing examination is that it can be performed with any MR imager, without proprietary software or additional software expense. We exploited the utility of this examination for gathering physiologic information regarding flow characteristics in the vertebral arteries in patients with and without subclavian steal syndrome.

Our study demonstrated a delay of at least 2 seconds in time to peak arterial enhancement in the vertebral artery ipsilateral to the side of subclavian steal in all patients with subclavian steal syndrome. This finding was not seen in patients with vertebral artery stenosis or those with normal vertebral arteries, and, therefore, the finding had 100% specificity for the diagnosis of subclavian steal syndrome.

The delay to peak enhancement was presumably due to increased bolus transit time, which in turn was caused by retrograde flow. We also noticed a concomitant decrease in peak vertebral arterial signal intensity in the affected side compared with that in the contralateral side in six (75%) of eight patients, which indicates a relative reduction of blood flow through the vertebral artery ipsilateral to the diseased subclavian artery. These findings are consistent with those in early experiments by Reivich et al (3), who used angiography and intraoperative flowmeters to demonstrate that retrograde vertebral arterial flow in patients with subclavian steal syndrome is decreased both in rate and in volume.

In our study patients with vertebral artery stenosis, a difference in time to peak enhancement was found in only one (11%) of nine patients. In this patient, the time to peak signal intensity in the stenotic vertebral artery was 1 second earlier than that in the contralateral vessel. This occurrence can be explained by the observation in prior US studies (6,21) that hemodynamic changes lead to increased flow velocity and earlier peak enhancement even in vessels with moderate stenosis. Conversely, the degree of stenosis may be so severe as to impede blood flow (as with stenoses that cause a reduction of 75% or more in the vessel cross-sectional area) or, in cases of occlusion, completely eliminate it (6). This phenomenon was not seen in our study, but we cannot exclude the possibility that severe stenoses may cause a delay in the time to peak signal intensity. Images from the patients in the control group demonstrated no significant differences in the time to peak signal intensity in the two vertebral arteries, and this finding indicates that equivalent circulation times may be expected through both vessels.

This study had several recognized limitations, including a retrospective design and a small number of patients. We attempted to minimize bias by consecutively selecting patients for each of the three study groups from our hospital database. Finally, we did not have conventional angiographic images for correlation with MR angiograms. However, breath-hold contrast-enhanced three-dimensional MR angiography has been accepted as a reliable test for evaluation of the great vessels and is currently the modality of choice at many institutions.

In conclusion, a bolus-timing examination in the neck vessels during which a difference of 2 seconds or more is observed between the times to peak enhancement in the vertebral arteries may indicate subclavian steal syndrome in the appropriate clinical setting. The absence of such a difference in time to peak enhancement probably excludes this diagnosis. In addition, this technique has potential for postoperative follow-up. Restoration of antegrade blood flow in the affected vertebral artery, seen as a disappearance of the delay in the time to peak signal intensity, would suggest successful correction of the syndrome. With the increasing use of MR angiography in imaging of vascular disorders, the bolus-timing examination can function as a supplemental tool for diagnosis or confirmation of subclavian steal syndrome.

References
6. Landwehr P, Schulte O, Voshage G. Ultra-
Diagnostic Imaging Costs: Are They Driving Up the Costs of Hospital Care?1

PURPOSE: To retrospectively determine how changes in utilization of computed tomography (CT), magnetic resonance (MR) imaging, and other imaging technologies between 1996 and 2002 influenced costs of inpatient hospital care at one large academic medical center.

MATERIALS AND METHODS: Institutional review board did not require its approval or patient informed consent for studies with use of billing data. Patient anonymity was protected by removal of potentially identifying information. Data on hospital costs for 17 139 patients admitted to Massachusetts General Hospital, Boston, Mass, between 1996 and 2002 were downloaded from hospital cost-accounting system; sample was restricted to inpatients with diagnoses in diagnosis-related groups 014–015 (Stroke and TIA [transient ischemic attack]), 164–167 (Appendectomy), 082 (Lung Cancer), 182–183 (Upper Gastrointestinal Conditions), 148–149 (Colon Cancer), and 243 (Back Problems). For each patient, data on demographics, all products and services used, and costs associated with each product or service were obtained. By using institutional codes, we calculated costs of CT, MR imaging, and total imaging relative to total hospital costs. Statistical analyses were performed with Student t test and multiple linear regression analysis.

RESULTS: Between 1996 and 2002, number of inpatient CT and MR images obtained at the hospital more than doubled. In 2002, hospital costs were 155% those of 1996 levels; inpatient imaging costs were 151% those of 1996 levels. Total costs increased an average of 7.8% per year; imaging costs increased 8.3% per year. Although highly variable over the study period, as a percentage of total imaging costs, CT and MR imaging costs appeared to remain stable relative to costs of other imaging modalities.

CONCLUSION: Despite substantial increases in utilization of inpatient CT, MR imaging, and other imaging technologies, diagnostic imaging costs increased at approximately same rate as did total costs for inpatients with several diagnoses. CT and MR imaging do not appear to be driving the cost increases seen between 1996 and 2002.

© RSNA, 2005

Computed tomography (CT) scanning is now the workhorse of most radiology departments and imaging centers in the United States. With many new indications and more efficient machines, CT scanning has forever changed our ability to diagnose and treat disease. Over the past 10 years, there have been substantial increases in the utilization of CT and magnetic resonance (MR) imaging (1–3). However, with medical care costs now composing 14% of the gross domestic product (4), there is concern that recent increases in the utilization of high-technology imaging studies are contributing heavily to the crisis in health care costs; some have even argued that diagnostic imaging has replaced prescription drugs as the new driver of health care costs. For instance, in a series of analyses of state-level data, “medical technology” (of which imaging is only a part) was one of the most important drivers of health care costs in a variety of settings. Medical technology accounted for 19% of the growth in hospital costs between 1998 and 2000 (5), 7% of the growth in outpatient costs between 1996 and 1999 (6), and 11% of the growth in physician services between 1990 and 2000 (7). Furthermore, in Massachusetts, diagnostic
imaging costs increased faster than total patient costs between fiscal years 1999 and 2001. These trends have made diagnostic imaging a potential target for cost-cutting initiatives through such mechanisms as restrictions on physician-self referral, tiered copayment systems, or the development of evidence-based practice guidelines (8). Though it is well established that the utilization of imaging technology has increased dramatically, few researchers have evaluated this trend in the context of the broader health care system. Findings in one study suggested a relationship between changes in the availability of CT, MR imaging, and other new technologies and health care spending by establishing an association between free-standing imaging centers and health care spending. The study, however, was limited by its short time span (9). Thus, the purpose of our study was to retrospectively determine how changes in the utilization of CT, MR imaging, and other imaging technologies between 1996 and 2002 have influenced costs of inpatient hospital care at a single large academic medical center.

MATERIALS AND METHODS

Inpatients

Between 1996 and 2002, more than 250,000 patients were admitted to Massachusetts General Hospital, Boston, Mass, an 875-bed urban teaching hospital that also serves as a tertiary referral center and a major trauma center. We selected a sample of these patients on the basis of diagnoses by using diagnosis-related groups (DRGs). Six DRGs were included in our study because they are associated with a high utilization of imaging studies, are in clinical areas where there were substantial advances in imaging technology over the study period, and represent diseases that affect several different organ systems and therefore that involve different imaging tests. Thus, our study included 17,139 inpatients with diagnoses in the following DRGs: 4888, DRG 014–015 (Stroke and TIA [transient ischemic attack]); 1686, DRG 164–167 (Appendectomy); 1014, DRG 082 (Lung Cancer); 4332, DRG 182–183 (Upper Gastrointestinal Conditions such as esophagitis, gastroenteritis, and diverticulitis); 4068, DRG 148–149 (Colon Cancer); and 1151, DRG 243 (Back Problems) (10).

Data Collection

For each patient, one author (M.T.B.) downloaded information from our institution’s cost-accounting database (Transition Systems Inc; TSI, subsidiary of Eclipsys, Boca Raton, Fla) as follows: age (sex not downloaded), length of hospital stay, length of intensive care unit (ICU) stay, DRG, and Charlson Comorbidity Score (11,12), as well as all products and services used and the costs associated with each product and service. At the time our study was conducted, our institutional review board did not require its approval or patient informed consent for studies using billing data. Patient anonymity, however, was protected by the removal of any potentially identifying information.

In the database, relative value units are assigned to each product or service to reflect the time and/or complexity and are updated annually. Relative value units are then converted to costs by using a set conversion value. Unit costs of products and services consist of both direct and indirect components. The direct cost of each product or service could be variable or fixed, depending on whether or not costs correlate with fluctuations in volume. For example, variable direct costs include consumable supplies and personnel time costs, while fixed direct costs include some portion of the acquisition costs (for major equipment) amortized over the lifetime of the equipment. To calculate these costs, representative time-and-motion studies are used for labor costs, and actual acquisition costs are used for supplies. A portion of indirect or overhead costs is also allocated to each product or service. Overhead costs are allocated on the basis of several factors, such as square footage and utility use, among others.

The sum of unit costs for all products and services used represents the total cost for a patient’s hospitalization. Physician’s fees are excluded because they are maintained separately by our institution’s physician organization. This cost-accounting method has been described previously (13–15).

Utilization and cost of imaging were determined by using institution-specific product codes for CT, MR imaging, and all other imaging modalities. In the cost-accounting database, some CT scans and MR images that are normally ordered as bundled images (ie, images of abdomen and pelvis or of head and neck) are counted as two individual images. As a result, data on the number of images obtained per patient presented in this study may appear inflated. Furthermore, for the purposes of this study, we calculated all costs relative to 1996 levels. Costs were not adjusted for inflation because, theoretically, the force of inflation would drive up both imaging and total hospital costs equally and would not affect the relationship between the two.

Statistical Analysis

All analyses were performed by using computer software (Microsoft Access and Excel 2000 for Windows, Microsoft, Redmond, Wash; SAS, version 8 for Windows, SAS Institute, Cary NC). Comparisons of inpatient demographics and length of hospital stay were made by using the Student t test, where appropriate (two-tailed test, α = .05). Multiple linear regression analysis was performed to determine the effect of covariates on total hospital costs and length of stay.

RESULTS

Inpatient Characteristics

In Table 1, demographic characteristics of the study sample according to DRG and of the overall sample are presented. In 2002, there were 752 more patients admitted with these diagnoses than there were in 1996. The mean age of the patients in the overall sample decreased significantly (from 59.6 to 57.9 years, P < .004) over the study period. There was also a significant decline in the mean Charlson Comorbidity Score, which declined from 2.05 in 1998 to 1.86 in 2002 (P < .001). Furthermore, the mean length of hospital stay decreased by 1.6 days (from 7.0 days in 1996 to 5.4 days in 2002, P < .001); however, the mean length of ICU stay remained stable at 0.4 day (P > .99).

CT and MR Imaging Utilization

The total number of CT scans and MR images obtained annually more than doubled over the study period (Table 2). Three factors contributed to this increase: the number of patients included in the sample increased, a greater percentage of these patients underwent CT and/or MR imaging, and more images were obtained per patient (among those who underwent imaging). For example, the overall percentage of patients who underwent CT increased from 47% in 1996 to 61% in 2002, while the average number of images obtained per patient increased from 1.9 to 2.5 images. Overall, there were...
only modest increases in the percentage of patients who underwent MR imaging (from 25% in 1996 to 26% in 2002); however, the average number of images obtained per patient increased (from 1.8 to 2.6 images).

Trends in utilization of CT and MR imaging varied over the study period, depending on the diagnosis (Table 2). The greatest change in CT utilization was observed among patients who underwent appendectomy (DRG 164–167). Over the study period, the number of CT scans obtained in patients who underwent appendectomy increased 694%, from 47 scans in 1996 to 373 scans in 2002. This was largely because of a 285% increase in the percentage of patients scanned (from 18% in 1996 to 69% in 2002). The number of CT scans obtained in patients with stroke and transient ischemic attack (DRG 014–015) more than doubled over the study period, from 751 scans to 1621 scans. This, however, was not caused by a change in the percentage of patients but rather by a 67% increase in the number of CT scans obtained per patient, which increased from 1.5 scans in 1996 to 2.5 scans in 2002.

We found that, in our sample, only stroke and transient ischemic attack and back problems (DRG 243) were associated with substantial utilization of MR imaging (ie, more than 50 images per year). Among patients with stroke and transient ischemic attack, the number of MR images obtained almost doubled between 1996 and 2002 because of a 47% increase in the number of MR images obtained per patient.

Costs

Between 1996 and 2002, hospital costs at our institution steadily rose at an average rate of 7.8% per year. By 2002, hospital costs were 55% higher than they were in 1996 (Fig 1). Although more variable, trends in inpatient imaging costs paralleled those of total costs. Over the study period, imaging costs increased an average of 8.3% per year. By 2002, imaging costs were 51% higher than they were in 1996.

As a percentage of total costs, imaging costs never increased beyond the 1996 level of 10.4% during the study period (Fig 2). Between 1997 and 2001, imaging costs declined as a percentage of total costs and never exceeded 9.3%. Between 2001 and 2002, however, imaging costs

### Table 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DRG 164–167 (Appendectomy)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients</td>
<td>168</td>
<td>219</td>
<td>197</td>
<td>230</td>
<td>269</td>
<td>291</td>
<td>312</td>
</tr>
<tr>
<td>Age (y)</td>
<td>29.7</td>
<td>32.1</td>
<td>33.7</td>
<td>33.0</td>
<td>33.8</td>
<td>33.8</td>
<td>33.3</td>
</tr>
<tr>
<td>Charlson Comorbidity Score</td>
<td>NA</td>
<td>NA</td>
<td>0.21</td>
<td>0.17</td>
<td>0.20</td>
<td>0.17</td>
<td>0.19</td>
</tr>
<tr>
<td>Length of total stay (d)</td>
<td>3.9</td>
<td>3.2</td>
<td>3.3</td>
<td>4.3</td>
<td>3.7</td>
<td>4.2</td>
<td>2.6</td>
</tr>
<tr>
<td>Length of ICU stay (d)</td>
<td>0.01</td>
<td>0.11</td>
<td>0.04</td>
<td>0.09</td>
<td>0.01</td>
<td>0.02</td>
<td>0.07</td>
</tr>
<tr>
<td><strong>DRG 014–015 (Stroke and TIA)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients</td>
<td>638</td>
<td>727</td>
<td>657</td>
<td>664</td>
<td>685</td>
<td>750</td>
<td>767</td>
</tr>
<tr>
<td>Age (y)</td>
<td>68.0</td>
<td>68.2</td>
<td>67.9</td>
<td>68.2</td>
<td>67.4</td>
<td>69.1</td>
<td>67.8</td>
</tr>
<tr>
<td>Charlson Comorbidity Score</td>
<td>NA</td>
<td>NA</td>
<td>2.10</td>
<td>2.18</td>
<td>2.10</td>
<td>2.15</td>
<td>2.18</td>
</tr>
<tr>
<td>Length of total stay (d)</td>
<td>6.4</td>
<td>6.8</td>
<td>6.2</td>
<td>6.6</td>
<td>7.3</td>
<td>7.4</td>
<td>5.3</td>
</tr>
<tr>
<td>Length of ICU stay (d)</td>
<td>0.4</td>
<td>0.4</td>
<td>0.7</td>
<td>0.8</td>
<td>0.7</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td><strong>DRG 243 (Back Problems)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients</td>
<td>119</td>
<td>116</td>
<td>144</td>
<td>176</td>
<td>191</td>
<td>196</td>
<td>209</td>
</tr>
<tr>
<td>Age (y)</td>
<td>57.0</td>
<td>59.5</td>
<td>56.8</td>
<td>58.3</td>
<td>58.8</td>
<td>57.4</td>
<td>58.7</td>
</tr>
<tr>
<td>Charlson Comorbidity Score</td>
<td>NA</td>
<td>NA</td>
<td>0.94</td>
<td>0.64</td>
<td>0.70</td>
<td>0.74</td>
<td>0.57</td>
</tr>
<tr>
<td>Length of total stay (d)</td>
<td>4.6</td>
<td>3.9</td>
<td>4.5</td>
<td>4.2</td>
<td>4.1</td>
<td>4.3</td>
<td>4.2</td>
</tr>
<tr>
<td>Length of ICU stay (d)</td>
<td>0.02</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0.05</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td><strong>DRG 148–149 (Colon Cancer)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients</td>
<td>551</td>
<td>547</td>
<td>562</td>
<td>553</td>
<td>621</td>
<td>566</td>
<td>668</td>
</tr>
<tr>
<td>Age (y)</td>
<td>59.8</td>
<td>59.5</td>
<td>58.4</td>
<td>58.0</td>
<td>57.4</td>
<td>58.6</td>
<td>58.0</td>
</tr>
<tr>
<td>Charlson Comorbidity Score</td>
<td>NA</td>
<td>NA</td>
<td>2.85</td>
<td>2.80</td>
<td>2.54</td>
<td>2.36</td>
<td>2.29</td>
</tr>
<tr>
<td>Length of total stay (d)</td>
<td>11.3</td>
<td>12.0</td>
<td>10.8</td>
<td>10.8</td>
<td>9.7</td>
<td>8.6</td>
<td>8.9</td>
</tr>
<tr>
<td>Length of ICU stay (d)</td>
<td>0.7</td>
<td>0.6</td>
<td>0.7</td>
<td>0.9</td>
<td>0.7</td>
<td>0.7</td>
<td>0.5</td>
</tr>
<tr>
<td><strong>DRG 182–183 (Upper Gastrointestinal Conditions)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients</td>
<td>524</td>
<td>547</td>
<td>534</td>
<td>581</td>
<td>621</td>
<td>743</td>
<td>782</td>
</tr>
<tr>
<td>Age (y)</td>
<td>59.1</td>
<td>59.5</td>
<td>59.5</td>
<td>58.1</td>
<td>57.8</td>
<td>59.7</td>
<td>57.4</td>
</tr>
<tr>
<td>Charlson Comorbidity Score</td>
<td>NA</td>
<td>NA</td>
<td>1.25</td>
<td>1.19</td>
<td>1.15</td>
<td>1.14</td>
<td>1.17</td>
</tr>
<tr>
<td>Length of total stay (d)</td>
<td>4.4</td>
<td>4.5</td>
<td>5.3</td>
<td>5.3</td>
<td>5.2</td>
<td>5.2</td>
<td>3.8</td>
</tr>
<tr>
<td>Length of ICU stay (d)</td>
<td>0.11</td>
<td>0.05</td>
<td>0.03</td>
<td>0.06</td>
<td>0.02</td>
<td>0.08</td>
<td>0.03</td>
</tr>
<tr>
<td><strong>DRG 082 (Lung Cancer)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients</td>
<td>149</td>
<td>136</td>
<td>107</td>
<td>132</td>
<td>138</td>
<td>189</td>
<td>163</td>
</tr>
<tr>
<td>Age (y)</td>
<td>64.9</td>
<td>65.5</td>
<td>63.5</td>
<td>62.3</td>
<td>65.3</td>
<td>65.2</td>
<td>65.0</td>
</tr>
<tr>
<td>Charlson Comorbidity Score</td>
<td>NA</td>
<td>NA</td>
<td>6.40</td>
<td>6.71</td>
<td>6.66</td>
<td>6.96</td>
<td>6.68</td>
</tr>
<tr>
<td>Length of total stay (d)</td>
<td>7.7</td>
<td>7.3</td>
<td>8.2</td>
<td>7.6</td>
<td>7.3</td>
<td>8.1</td>
<td>6.4</td>
</tr>
<tr>
<td>Length of ICU stay (d)</td>
<td>0.41</td>
<td>0.07</td>
<td>0.12</td>
<td>0.10</td>
<td>0.08</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td><strong>Overall sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total no. of patients</td>
<td>2149</td>
<td>2292</td>
<td>2201</td>
<td>2336</td>
<td>2252</td>
<td>2735</td>
<td>2901</td>
</tr>
<tr>
<td>Mean age (y)</td>
<td>59.6</td>
<td>60.0</td>
<td>59.8</td>
<td>58.7</td>
<td>58.2</td>
<td>59.2</td>
<td>57.9</td>
</tr>
<tr>
<td>Mean Charlson Comorbidity Score</td>
<td>NA</td>
<td>NA</td>
<td>2.05</td>
<td>2.02</td>
<td>1.92</td>
<td>1.94</td>
<td>1.86</td>
</tr>
<tr>
<td>Mean length of total stay (d)</td>
<td>7.0</td>
<td>5.7</td>
<td>6.9</td>
<td>6.9</td>
<td>6.8</td>
<td>6.6</td>
<td>5.4</td>
</tr>
<tr>
<td>Mean length of ICU stay (d)</td>
<td>0.4</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
</tbody>
</table>

* NA = not applicable.
increased sharply (from 9.3% to 10.3%) as a function of total costs.

The unit cost per CT scan steadily declined over the study period (Fig 3). By 2002, the cost per CT scan was 59% that of 1996. This decrease in the per-scan cost was due in part to the increased volume of CT scans obtained, and, thus, the per-scan fixed indirect costs decreased. This trend was not observed for the unit cost per MR image. The cost per MR image varied throughout the study period; however, by 2002, the cost per MR image was 110% that of 1996 levels.

In Table 3, imaging costs are classified according to modality. As a percentage of total imaging costs, CT costs remained relatively stable at approximately 20% throughout the study period. Although variable over the study period, MR imaging costs increased from 19% of total imaging costs in 1996 to 28% in 2002. Costs of imaging for other modalities, although highly variable throughout the study period, declined only slightly as a percentage of total imaging costs.

According to the results of the multiple regression models, only length of stay and length of ICU stay were predictive of increased total hospital costs (Table 4). An additional day of hospital stay and an additional day of ICU stay added $1172 and $4199, respectively, to total hospital costs, while controlling for all other covariates included in the model. Length of ICU stay was also predictive of overall length of stay, as was age, Charlson Comorbidity Score, and imaging costs. An additional day of ICU stay added 5.36 days to the total length of stay, an increase of 10 years of age added 0.6 day, an increase of one unit of Charlson Comorbidity Score added 0.54 day, and an additional $100 of imaging costs decreased the length of stay by 0.26 day, while controlling for all other covariates in the model. Put another way, spending an additional $385 on imaging, according to this model, was associated with a 1-day reduction in the length of stay. Both the model for total hospital costs and length of stay had high adjusted $R^2$ values of 0.86 and 0.90, respectively, a finding that suggests that a large portion of the variance could be explained by the covariates included in the models.

**DISCUSSION**

We found that imaging costs at our institution increased by more than 50% between 1996 and 2002 among patients included in this study. Over the study period, the number of CT and MR imaging examinations performed annually more than doubled, a trend that was driven by increases in three factors: the number of pa-
tients admitted, the percentage of these patients who were imaged during their admission, and the number of images obtained per patient.

Over the study period, total hospital costs increased at approximately the same rate as did inpatient imaging costs. As a percentage of total hospital costs, inpatient imaging costs remained relatively stable between 1996 and 2002, and they represented approximately 10% of the total costs. As a result, diagnostic imaging is unlikely to be a dominant driver of hospital costs. However, the rate of increase in imaging costs was tempered by a substantial reduction in the cost per examination, particularly per CT examination.

Imaging costs, broken down according to modality, were so highly variable over the study period that we were not able to determine whether there was any technology substitution of CT or MR imaging for other modalities. Although highly variable over the study period, costs of MR imaging appeared to represent a slightly greater percentage of imaging costs relative to CT, while the costs of other imaging modalities declined somewhat.

In multiple regression analyses, we found that imaging costs were not predictive of total hospital costs. We did find, however, that imaging costs were predictive of length of stay. According to the model, an increase of $385 in total imaging costs per patient was associated with a reduction of 1 day in total length of stay.

The principal limitations of this study relate to how costs were calculated and what costs were included in the study. In this study, we relied solely on the cost-accounting system at our institution. Although the costs derived from this system may accurately reflect costs at our institution, the generalizability of the findings to other institutions and health care settings may be limited. Specifically, how overhead costs are allocated and how relative value units are assigned to individual products or services may vary between institutions. For example, at our institution, imaging tests with higher relative value units, such as CT angiography and MR angiography, were introduced during the study period. This may have skewed the unit costs for CT and MR imaging in the later years of the study. Nevertheless, we believe that findings at our institution are likely comparable to those that would have been observed at other large urban teaching hospitals.

Another important limitation of our study relates to our analysis of only inpatient costs. Most imaging is now performed on an outpatient basis. If imaging studies that had been performed on an inpatient basis were shifted to the outpatient setting during the study period, our analysis would have underestimated the contribution of imaging costs to total costs.

Furthermore, costs presented in this study were not adjusted for inflation in medical care costs or changes in case mix. We did not think that adjustment for inflation was necessary because the objective of the study was to analyze how changes in diagnostic imaging costs have changed relative to total inpatient costs. Theoretically, the force of inflation would have the same effect on both di-

**Table 3**

<table>
<thead>
<tr>
<th>Year</th>
<th>CT</th>
<th>MR Imaging</th>
<th>Other Imaging</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>22</td>
<td>19</td>
<td>59</td>
</tr>
<tr>
<td>1997</td>
<td>22</td>
<td>22</td>
<td>56</td>
</tr>
<tr>
<td>1998</td>
<td>22</td>
<td>24</td>
<td>54</td>
</tr>
<tr>
<td>1999</td>
<td>25</td>
<td>31</td>
<td>44</td>
</tr>
<tr>
<td>2000</td>
<td>23</td>
<td>29</td>
<td>48</td>
</tr>
<tr>
<td>2001</td>
<td>25</td>
<td>36</td>
<td>39</td>
</tr>
<tr>
<td>2002</td>
<td>20</td>
<td>28</td>
<td>52</td>
</tr>
</tbody>
</table>

Note.—Numbers are percentages of total imaging costs.
agnostic imaging costs and total inpatient costs and thus would not change the nature of the relationship between the two. In addition, we thought that adjustment for case mix was not warranted in this study. Case mix indexes are typically used to control for changes in distribution of diagnoses (ie, severity of disease) over time and are calculated by using DRGs. In this study, we held diagnoses constant by selecting patients with a fixed set of six DRGs. We further attempted to control for severity of disease by collecting data on patient characteristics, such as age and Charlson Comorbidity Score. Interestingly, we found that both age and Charlson Comorbidity Score declined significantly (P < .004 and < .001, respectively) over the study period, and this decline suggested that, among the diagnoses we chose to include in the study, patients were actually less ill over time. Thus, it is unlikely that severity of disease can explain increases in either diagnostic imaging costs or total inpatient costs in this study.

In an evaluation of the value of improved imaging technology, ideally we would like to show that the use of today’s high-technology imaging services has somehow resulted in improved patient outcomes and/or decreased costs. Although our study cannot address patient outcomes, we have shown, at the very least, that imaging may not increase hospital costs at the same rate that imaging utilization has increased. Our results would therefore suggest that across-the-board limits on imaging utilization would be ill-advised.

In conclusion, in our study, although inpatient diagnostic imaging costs increased dramatically over the 7-year study period, total hospital costs increased at approximately the same rate. Inpatient imaging costs remained stable at approximately 10% of total hospital costs. Thus, although increases in imaging costs contribute to increases in hospital costs, so too do many other factors, and imaging costs cannot fully explain the observed trends.

References
Comparative Scatter and Dose Performance of Slot-Scan and Full-Field Digital Chest Radiography Systems

PURPOSE: To evaluate the scatter, dose, and effective detective quantum efficiency (DQE) performance of a slot-scan digital chest radiography system compared with that of a full-field digital radiography system.

MATERIALS AND METHODS: Scatter fraction of a slot-scan system was measured for an anthropomorphic and a geometric phantom by using a posterior beam-stop technique at 117 and 140 kVp. Measurements were repeated with a full-field digital radiography system with and without a 13:1 antiscatter grid at 120 and 140 kVp. For both systems, the effective dose was measured on posteroanterior and lateral views for standard clinical techniques by using dosimeters embedded in a female phantom. The effective DQEs of the two systems were assessed by taking into account the scatter performance and the DQE of each system. The statistical significance of all the comparative differences was ascertained by means of t test analysis.

RESULTS: The slot-scan system and the full-field system with grid yielded scatter fractions of 0.13–0.14 and 0.42–0.48 in the lungs and 0.30–0.43 and 0.69–0.78 in the mediastinum, respectively. The sum of the effective doses for posteroanterior and lateral views for the slot-scan system (0.057 mSv ± 0.003 [± standard deviation]) was 34% lower than that for the full-field system (0.086 mSv ± 0.001, P < 0.05) at their respective clinical peak voltages (140 and 120 kVp). The effective DQE of the slot-scan system was equivalent to that of the full-field system in the lung region but was 37% higher in the dense regions (P < 0.05).

CONCLUSION: The slot-scan design leads to marked scatter reduction compared with the more conventional full-field geometries with a grid. The improved scatter performance of a slot-scan geometry can effectively compensate for low DQE and lead to improved image quality.

In the past decade, an increasing number of dedicated digital imaging devices have been developed for chest imaging. Many of these devices have been of the flat-panel type, which includes arrays of thin-film transistors coupled to either photoconductors (direct detection technology) (1) or to photodiodes with scintillators (indirect detection) (2). Other methods have included arrays of charge-coupled device (CCD) or complementary metal oxide semiconductor detectors (3,4) and a selenium-coated drum scanned with electrometers (5). Described in terms of the detective quantum efficiency (DQE), a common metric used to describe the exposure efficiency of a detector on the basis of its response to the primary x-ray radiation, most of these detectors have shown DQEs superior to those of screen-film or computed radiographic detectors (2,6–8). This improved DQE has enabled improved image quality per unit exposure (or consistent quality at reduced exposure) and has made possible various advanced applications that rely on high signal-to-noise ratio (SNR) and quantum efficiency.

While the concept of the DQE applies to the detector response to primary radiation, the unwanted detection of scattered photons at chest radiography has long been recognized as

Abbreviations:
DQE = detective quantum efficiency
SNR = signal-to-noise ratio
TLD = thermoluminescent dosimeter

1 From the Duke Advanced Imaging Laboratories, Department of Radiology (E.S., J.Y.L., T.T.Y., J.L.J., J.T.D., C.E.F., H.P.M., C.E.R.), Department of Biomedical Engineering (E.S., J.Y.L., J.L.J., J.T.D., C.E.F.), Department of Physics (E.S.), and Radiation Safety Division (T.T.Y.), Duke University Medical Center, DUMC 3302, Durham, NC 27710. Re-ceived March 18, 2004; revision requested August 19. Address correspondence to E.S. (e-mail: samei@duke.edu). Authors stated no financial relationship to disclose.

© RSNA, 2005
a source of image quality degradation in terms of reduced image contrast and increased noise. Scattered x-rays can account for 95% of the detected x-ray flux in the mediastinum and up to 70% in the lung on radiographs acquired without a grid (9,10). By adding an unknown nonuniform background to the image, scatter degrades contrast and relative noise (expressed in terms of variance of the detector signal value divided by the mean detector signal value squared) on the image by a factor equal to the value of the scatter fraction. In other words, on a region of an uncompensated radiograph with a scatter fraction of 95%, only 5% of the original primary contrast remains. While this reduction in contrast can be remedied by means of postprocessing, the contribution of scattered radiation to image noise may not be readily correctable.

Over the years, various techniques have been developed to reduce the contribution of scattered photons to the image. The uses of antiscatter grids (11,12) and air gaps (13) have by far been the most common techniques. A grid can reduce the scatter fraction from 95% down to 65% in the mediastinum and from 70% down to 20% in the lung (11). However, while grids reject scattered photons, they also reduce the detection efficiency for primary photons, which leads to an increase in image noise for the same patient entrance exposure. The net result is an increase in SNR at the cost of an increase in patient dose. Air gaps similarly reduce detected scattered photons but can lead to increased patient dose and magnification blur. To overcome these fundamental limitations, the use of scanning beam and slit devices has been proposed (12,14–16). These techniques offer scatter rejection without compromising the detection of primary photons. However, a previous commercial implementation of the scanning slot chest radiography system did not reject scatter better than a standard grid, possibly because of the large width of the slot (4 cm) for that system (17) or because of suboptimal tuning of the beam and detector slots.

A dedicated digital chest radiography system with a scanning slot apparatus has recently become commercially available. The system uses a small field array of detector elements that travel with a scanning slot, four times narrower than that of the previous implementation, to construct a full-field chest radiograph. A report of a recent study provided the resolution and DQE performance and the preliminary scatter performance of the new system (18). The purpose of our current study was to evaluate the scatter, dose, and effective DQE performance of the slot-scan digital chest radiography system compared with that of a more conventional full-field radiography system.

**MATERIALS AND METHODS**

**Slot-Scan Imaging System**

Physical specifications of the slot-scan imaging system are shown in Table 1 and Figure 1a. The slot-scan system (ThoraScan; Delft Imaging Systems/Nucletron, Veenendaal, the Netherlands) is composed of a moving CCD-based slot detector in conjunction with a dedicated x-ray tube and high-frequency generator capable of producing a moving fan beam of x-rays. The detector consists of eight 5.5 × 1.1-cm subdetectors tiled together to create a 44 × 1.1-cm sensitive area. Each subdetector is made of a thallium-doped cesium iodide scintillation layer fiberoptically coupled to a CCD.

The system uses the slot scanning geometry to acquire the chest radiograph with no antiscatter grid in place. The narrow fan beam synchronized with the movement of the detector assembly scans the patient, moving upward. Automatic exposure control is achieved by means of a low-dose partial downward initial scanning from which the appropriate level of tube current for subsequent scanning is determined. The image data are continuously read from the CCDs as the patient is scanned by means of the time-integration method (19,20). After scanning, the image data are sent to an associated processing workstation, where the data are resampled into equal increments in the orthogonal directions and are corrected for nonuniformities through gain, offset, and bad-pixel calibration (21). The image data are finally postprocessed by using a multiscale image processing algorithm and are tone scaled for proper display (22).

Prior to image acquisition, the imaging system was calibrated according to the manufacturer's guidelines at 117 kVp and 2.4 mAs with 2.0 mm of added copper filtration. The copper filtration was removed for subsequent image acquisitions. The system performance was evaluated at 117 and 140 kVp, with 0.3 mm of added copper filtration. The latter is the setting suggested by the manufacturer for most patients, while the former is recommended for small patients.

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Basic Physical Specifications of the Evaluated Slot-Scan and Full-Field Systems</strong></td>
</tr>
<tr>
<td><strong>Characteristic</strong></td>
</tr>
<tr>
<td>Capture element thickness (mm)*</td>
</tr>
<tr>
<td>Detector fill factor (%)</td>
</tr>
<tr>
<td>CCD pixel pitch (mm)</td>
</tr>
<tr>
<td>Image pixel pitch (mm)</td>
</tr>
<tr>
<td>Image matrix size</td>
</tr>
<tr>
<td>Subdetector size (cm)</td>
</tr>
<tr>
<td>Detector size (cm)</td>
</tr>
<tr>
<td>Image area (cm)</td>
</tr>
<tr>
<td>Scanning time (sec)</td>
</tr>
<tr>
<td>Dwell time</td>
</tr>
<tr>
<td>Source to image distance (cm):</td>
</tr>
<tr>
<td>Air gap (cm)</td>
</tr>
<tr>
<td>Antiscatter grid</td>
</tr>
<tr>
<td>Lines per centimeter</td>
</tr>
<tr>
<td>Standard filtration (mm)</td>
</tr>
<tr>
<td>Half-value layer at 140 kVp (mm)</td>
</tr>
</tbody>
</table>

---

*Thallium-doped cesium iodide scintillator for both systems.
†Calculated as scanning width (1.1 cm) × scanning time (1.3 sec)/scanning length (44.3 cm).
‡Measured to the cover plate of the system.
Radiology on the geometric phantom images. Similar region designations were used in the mediastinum, and retrocardiac and subdiaphragmatic areas. Each beam stop comprised an area immediately surrounding the exposure behind the beam stop to that in the standard antiscatter grid in place. In previous experiments, this augmentation was found to provide scatter fractions that are closer in value to those measured in clinical practice (24). The geometric phantom consisted of interleaved slices of acrylic designed to approximate the attenuation and scatter characteristics of the chest, specifically 11.2 cm of acrylic in the lung, 20.0 cm in the mediastinum, and 22.3 cm in the subdiaphragm. Note that the geometric phantom did not include a retrocardiac region. Each phantom was placed in a posteranterior orientation. An array of beam stops (14 x 16 array, 25-mm spacing, each beam stop a cylinder of 3-mm diameter and 6-mm height embedded in a 6-mm sheet of acrylic) was placed posterior to each phantom (ie, on the tube side of the phantom), and images were acquired with postprocessing and tone scaling turned off and with automatic exposure control set to manual techniques. For the slot-scan system, the images were acquired at 117 and 140 kVp with exposure set at four times the predetermined photo-timed milliampere-second settings to ensure quantitative accuracy in measuring the low exposures behind the beam stops. For the full-field system, scatter measurements were acquired at 120 and 140 kVp at two times the photo-timed milliampere-seconds with and without the standard anticscatter grid in place.

For each beam stop, the scatter fraction was then measured as the ratio of the exposure behind the beam stop to that in an area immediately surrounding the beam stop by using a technique previously developed at our laboratory (9,25). Only beam stops in specific anatomic regions were used for analyses; all others were disregarded. The scatter fractions in each anatomic region were then computed from averages obtained from the beam stops located in that region. Nonetheless, the number of beam stops was always adequate to characterize the scatter fraction in each region. For the slot-scan system, the number of beam stops projected in the lung, mediastinum, retrocardiac region, and subdiaphragmatic region of the images were 62, 34, 0, and 22 in the geometric phantom; and 30, 6, 10, and 23 in the anthropomorphic phantom, respectively (Fig 2). The corresponding numbers for the full-field system were 52, 34, 0, and 11 in the geometric phantom and 32, 7, 11, and 18 in the anthropomorphic phantom, respectively. The scatter measurements were obtained primarily by two of the authors (J.Y.L., J.L.J.), with additional assistance from the other authors.

Scatter Measurements

The scattered radiation performance of the two imaging systems was evaluated in terms of scatter fraction by using an anthropomorphic (Humanoid Systems, Carson, Calif) and a geometric chest phantom (23). The anthropomorphic phantom, designed to correspond to a human of average build (175 cm in height, 73.5 kg in mass), was augmented with two additional 2.5-cm-thick slabs of acrylic (one posterior and one anterior). In previous experiments, this augmentation was found to provide scatter fractions that are closer in value to those measured in clinical practice (24). The geometric phantom consisted of interleaved slices of acrylic designed to approximate the attenuation and scatter characteristics of the chest, specifically 11.2 cm of acrylic in the lung, 20.0 cm in the mediastinum, and 22.3 cm in the subdiaphragm. Note that the geometric phantom did not include a retrocardiac region. Each phantom was placed in a posteranterior orientation. An array of beam stops (14 x 16 array, 25-mm spacing, each beam stop a cylinder of 3-mm diameter and 6-mm height embedded in a 6-mm sheet of acrylic) was placed posterior to each phantom (ie, on the tube side of the phantom), and images were acquired with postprocessing and tone scaling turned off and with automatic exposure control set to manual techniques. For the slot-scan system, the images were acquired at 117 and 140 kVp with exposure set at four times the predetermined photo-timed milliampere-second settings to ensure quantitative accuracy in measuring the low exposures behind the beam stops. For the full-field system, scatter measurements were acquired at 120 and 140 kVp at two times the photo-timed milliampere-seconds with and without the standard anticscatter grid in place.

For each beam stop, the scatter fraction was then measured as the ratio of the exposure behind the beam stop to that in an area immediately surrounding the beam stop by using a technique previously developed at our laboratory (9,25). Only beam stops in specific anatomic regions were used for analyses; all others were disregarded. The scatter fractions in each anatomic region were then computed from averages obtained from the beam stops located in that region. Nonetheless, the number of beam stops was always adequate to characterize the scatter fraction in each region. For the slot-scan system, the number of beam stops projected in the lung, mediastinum, retrocardiac region, and subdiaphragmatic region of the images were 62, 34, 0, and 22 in the geometric phantom and 30, 6, 10, and 23 in the anthropomorphic phantom, respectively (Fig 2). The corresponding numbers for the full-field system were 52, 34, 0, and 11 in the geometric phantom and 32, 7, 11, and 18 in the anthropomorphic phantom, respectively. The scatter measurements were obtained primarily by two of the authors (J.Y.L., J.L.J.), with additional assistance from the other authors.

Dose Measurements

The dose performance of each system was evaluated by using an adult female anthropomorphic phantom (CIRS, Norfolk, Va) embedded with thermoluminescent dosimeters (TLDs) (Fig 3). The phantom was different from those used for the scatter measurements because those phantoms were not designed for embedding dosimeters. The TLDs (Harshaw TLD-100; Thermo Electron, Santa Fe, NM) were embedded in 20 locations within the phantom. The locations included the left and right sides of the phantom. Four locations within each breast were also included and corresponded to 12-, 3-, 6-, and 9-o'clock positions while facing the breast. Two TLD chips were placed...
within each organ location, and the mean value of two readings was used for the assessment of dose at that location. The dose error bars associated with each TLD-estimated dose level were estimated from the data obtained in a separate experiment, in which relative error was assessed as a function of dose level by irradiating a series of five TLD chips (26). Prior to data acquisition, the TLDs were calibrated by measuring their response to a simulated x-ray beam with a comparable peak voltage and half-value layer. The details of the TLD calibration method can be found elsewhere (26). The dosimetry took into account the active marrow distribution in various bones within the body (27).

The phantom was placed in either a posteroanterior or a left-lateral orientation. For lateral views, the beam entered the phantom from the right side. The measurements on the slot-scan system were made at 117 and 140 kVp. Those on the full-field system were made at 120 and 140 kVp. For each system and technique, multiple images were acquired at milliampere-second settings higher than clinical settings to produce high signals in the TLDs.

After exposure, the TLD chips were extracted and read by using a TLD reader (Auto TLD Reader Model QS 5500; Harshaw, Solon, Ohio). Organ dose values were normalized by the total applied milliampere-second values and then were multiplied by the standard clinical photo-timed milliampere-second values for the phantom. In this manner, the measured dose values were scaled to be representative of clinical doses. For organs for which the dose was measured at multiple locations (eg, breast), the maximum measured dose value was used as the most conservative estimate of radiation risk and organ dose. The effective dose was then calculated by using the weight factors published in International Commission on Radiological Protection Publication 60 (28). The overall error for the effective dose was estimated by using the quadrature summing of the errors for each organ within the phantom. The dose measurements were obtained primarily by one of the authors (T.T.Y.), with additional assistance from the other authors.

Effective DQE Calculations

While the DQE is used to describe the performance of a digital radiographic detector without scatter, in the presence of scattered radiation and additional attenuating layers, an effective DQE, $DQE_{eff}$, of a radiographic system may be defined as $DQE_{eff} = t(1 - SF)DQE$, where $t$ is the transmission of the primary x-rays by the elements of the imaging chain before the detector, $SF$ is the scatter fraction, and DQE is the conventionally measured DQE in the absence of the antiscatter grid and scattered radiation (18). This definition is consistent with an earlier extension of the DQE concept by Wagner et al (29) to characterize the performance of antiscatter grids.

By assuming equivalent x-ray attenuation through the detector cover plates of the two systems, the effective DQEs of the systems were assessed from the measured scatter fractions with the anthropomorphic phantom, a grid transmission measurement, and prior reports of the DQE performance of the two systems. By neglecting the frequency-dependent components of the scattered radiation, the measured DQE values at 0.15 cycles per millimeter were used as surrogates for the DQE at zero frequency. Those values were previously reported to be 0.45, 0.18, and 0.16 for the full-field system at 120 kVp, the slot-scan system at 117 kVp, and the slot-scan system at 140 kVp, respectively (6,18). For the full-field system, it was assumed that the DQE at 140 kVp was lower than that at 120 kVp by the same fraction as that measured for the slot-scan system, that is, full-field DQE (0.15 cycles per millimeter, 140 kVp) is 0.45 - 0.16/0.18 = 0.4.

For the full-field system, the primary transmission through the grid was measured at the standard clinical beam quality (120 kVp, 0.2 mm of added copper filtration). This measurement was obtained by placing a calibrated ionization chamber (model 10×5–6 ionization chamber and model 1015 x-ray monitor; Radcal, Monrovia, Calif) on the central axis of the beam at a distance of 100 cm from the focal spot. The beam was collimated to only envelope the ionization chamber, with 3-cm margins on each side. The ratio of measured exposures at a fixed high milliampere-second with and without the grid at the collimator was used to calculate the grid primary transmission. The measurement indicated a grid primary transmission of 0.65. Because the chamber was placed on the central axis and away from the detector, the measurements were not affected by off-focus placement of the grid or back-scattered radiation. All the effective DQE calculations and required measurements were made by one author (E.S.) and the results were examined by all.

To further verify the relative effective DQE of the two systems, we acquired images of the anthropomorphic chest phantom with the two systems and their corresponding clinical techniques (slot scan, 140 kVp; full field, 120 kVp with grid) with photo-timed exposures. The images were acquired by one author (E.S.) and were subjectively evaluated by all authors, independently and by consensus, to identify any substantial errors in the conclusions drawn from the dose and the effective DQE results.

Statistical Analysis

For the scatter fraction measurements, the differences between the two imaging systems and those between the full-field system with and without the grid were examined by using a two-tailed t test, assuming unequal variances. For cases in which the two experimental conditions included an equal number of beam stops (ie, the full-field system with and without the grid and each system at different beam qualities), a paired t test was performed; otherwise, an unpaired t test was performed. Similar analysis was undertaken for statistical significance of differences observed in the calculated effective DQE figures, with the assumption that scatter fraction was the only source of uncertainty. For dose measurements, the differences between the two systems were similarly examined by using a t test. Statistically significant differences were defined at P values less than .05. The statistical analyses were performed by using a spreadsheet program (Excel 2000; Microsoft, Redmond, Wash). The calculations were performed primarily by three
of the authors (J.Y.L., T.T.Y., E.S.), with additional input from our statistical consultant.

RESULTS

Scatter Results

The measured scatter fractions for both imaging systems at two energy levels are presented in graphic and tabular form in Figure 4a and 4b for measurements with the anthropomorphic and geometric phantoms, respectively. The most dramatic difference in the measured scatter fractions is observed when comparing the values from the two systems. For the full-field system, the values are consistent with those expected with the use of a high-efficiency grid. In the region of the phantoms that represented the lung, scatter fractions of 40%–45% were measured, while in the mediastinum the values were larger, as expected: 70%–80%, with slight differences between the mediastinal, retrocardiac, and subdiaphragmatic regions. The slot-scan system showed dramatically reduced scatter fractions for all regions in both phantoms. For the lung region, the scatter fractions were about 13%. For the mediastinum, they were in the 30%–40% range, while in the subdiaphragmatic region, they ranged from 32% to 36%. When compared with measurements for the full-field system, these slot-scan measurements represent a marked decrease in scatter fraction by a factor of about three in the lung, a factor of two in the mediastinum, and a factor between two and three in the subdiaphragmatic region. For the retrocardiac region of the anthropomorphic phantom, this is a decrease by a factor just less than two.

Scatter fractions measured for the full-field system with the anticscatter grid removed were 67% for the lung and 86%–91% for the denser (mediastinal, subdiaphragmatic, and retrocardiac) regions. With both phantom results at 120 kVp, the grid reduced scatter fractions by approximately 22%–25% in the lung and 16%–18% in the denser regions; the corresponding values for the slot-scan system were 54% and 47%–57%, respectively, which suggests a substantial advantage of the slot-scan technique over the grid for scatter reduction. For each system, there is no appreciable difference between the measured scatter fractions for the two energy levels examined with either of the two phantoms.

In comparing the measured scatter fractions, the differences between the two systems, between the grid and no grid conditions, and between different beam qualities were all statistically significant by a large margin ($P < .001$).

Dose Results

Tables 2 and 3 summarize the dose performance of the two systems. Notable for the left-lateral exposure is the variation in organ doses because of geometric distance and attenuation. For example, for the slot-scan system, the lateral lung dose ratio was 4.2 (109.0 µGy in right lung ÷ 25.9 µGy in left lung) at 117 kVp and 3.7 at 140 kVp. For the full-field digital system, the lateral lung dose ratio was 3.7 at 120 kVp and 4.3 at 140 kVp. Similarly, the lateral 9-o’clock breast dose ratio was 5.4 at 117 kVp and 5.1 at 140 kVp for the slot-scan system and 5.8 at 120 kVp and 4.8 at 140 kVp for the full-field digital system. Overall, the effective dose for the lateral view with the slot-scan system was a factor of 1.3 higher than that for the posteroanterior view at 117 kVp and was a factor of 1.6 higher at 140 kVp. This factor was notably higher (ie, a factor of...
3.6 at both 120 kVp and 140 kVp) for the full-field digital system tested (Table 3). In the comparison of the two systems at their corresponding clinical peak voltage settings (120 kVp for full field and 140 kVp for slot scan), the slot-scan system delivered a 16% (1 – 0.022/0.019) higher effective dose for the posteroanterior exposure, but this difference was not statistically significant (P > .05). However, for the lateral exposure, the slot-scan system delivered a significantly (P < .05) lower effective dose, which led to an approximately 34% (1 – 0.057/0.087) lower total dose (posteroanterior and lateral) compared with the total dose delivered by the full-field system (P < .05).

**Effective DQE Results**

Figure 5 presents values for the effective DQE. For the full-field system, the presence of the grid was generally advantageous, particularly in the dense regions. In the lung region, the two systems at their corresponding operating conditions (slot scan, 140 kVp; full field, 120 kVp with grid) performed equally well. However, in the denser regions, the slot-scan system yielded superior performance compared with the full-field system by an average of 37%, which indicates that a more efficient scatter rejection of a system can effectively compensate for its lower DQE. The comparable dose figures of the two systems as noted here serve as a secondary confirmation of this observation. All the differences in the effective DQE measurements were statistically significant (P < .05).

Figure 6 shows the images from our subjective evaluation. Given the differences in the image processing techniques applied by the two systems, the images cannot be objectively compared. However, from a subjective standpoint, the images obtained with the slot-scan system show relatively more detail and less noise. The difference in the lung regions is less pronounced, as predicted from the comparable effective DQEs of the two systems in that region, and, in fact, the slightly improved details obtained with the slot-scan system may be attributed to the statistically insignificant 16% greater radiation dose for that system (0.022 and 0.019 mSv, P > .05; Tables 2, 3). However, in the mediastinal region, the difference between the effective DQE of the two systems was more pronounced, and, thus, the difference in the noise level of the corresponding images is more readily discernable. These subjective observations confirm the predictions of the effective DQE observations.

**DISCUSSION**

In the past 20 years, various technologies have been deployed to acquire digital chest radiographs. Among those, computed radiography, CCDs and complementary metal oxide semiconductor devices, and flat-panel digital detectors have become commercially viable and are gradually replacing more conventional analog screen-film systems. One of the major goals in the development of these detector technologies has been the improvement of the DQE, which enables the acquisition of higher quality images with greater SNR at a lower dose. While other factors such as scatter and ana-

### Table 2

<table>
<thead>
<tr>
<th>Organ</th>
<th>117 kV</th>
<th>140 kV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Posteroanterior at 3.0 mAs*</td>
<td>Left Lateral at 4.0 mAs*</td>
</tr>
<tr>
<td>Lung Right</td>
<td>109.0 ± 13.0</td>
<td>NA</td>
</tr>
<tr>
<td>Left</td>
<td>111.0 ± 18.0</td>
<td>25.9 ± 5.0</td>
</tr>
<tr>
<td>Breast Right</td>
<td>136.0 ± 20.0</td>
<td>NA</td>
</tr>
<tr>
<td>12 o'clock</td>
<td>41.2 ± 8.2</td>
<td>87.3 ± 14.9</td>
</tr>
<tr>
<td>3 o'clock</td>
<td>133.0 ± 20.0</td>
<td>NA</td>
</tr>
<tr>
<td>6 o'clock</td>
<td>170.0 ± 22.0</td>
<td>NA</td>
</tr>
<tr>
<td>9 o'clock</td>
<td>20.9 ± 4.5</td>
<td>NA</td>
</tr>
<tr>
<td>Bone marrow</td>
<td>28.8 ± 6.0</td>
<td>17.0 ± 3.7</td>
</tr>
<tr>
<td>Effective dose (mSv)</td>
<td>0.028 ± 0.003</td>
<td>0.038 ± 0.002</td>
</tr>
</tbody>
</table>

Note.—Data are ± 1 standard deviation, and except where noted, data are measured in micrograys. NA = not applicable.

* Effective tube current, calculated as follows: (indicated nominal tube current)/(beam-on time, 1.3 sec)/(scanning length, 44.3 cm). Calculations were made with the assumption of a rectangular x-ray beam profile.

### Table 3

<table>
<thead>
<tr>
<th>Organ</th>
<th>120 kV</th>
<th>140 kV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Posteroanterior at 2.07 mAs*</td>
<td>Left Lateral at 6.74 mAs*</td>
</tr>
<tr>
<td>Lung Right</td>
<td>118.0 ± 13.0</td>
<td>NA</td>
</tr>
<tr>
<td>Left</td>
<td>68.8 ± 1.7</td>
<td>32.0 ± 5.9</td>
</tr>
<tr>
<td>Breast Right</td>
<td>178.0 ± 14.0</td>
<td>NA</td>
</tr>
<tr>
<td>12 o'clock</td>
<td>111.0 ± 13.0</td>
<td>15.8 ± 0.4</td>
</tr>
<tr>
<td>3 o'clock</td>
<td>23.2 ± 0.6</td>
<td>22.7 ± 4.5</td>
</tr>
<tr>
<td>6 o'clock</td>
<td>193.0 ± 14.0</td>
<td>NA</td>
</tr>
<tr>
<td>9 o'clock</td>
<td>231.0 ± 14.0</td>
<td>NA</td>
</tr>
<tr>
<td>Bone marrow</td>
<td>29.3 ± 5.5</td>
<td>40.5 ± 3.2</td>
</tr>
<tr>
<td>Effective dose (mSv)</td>
<td>0.009 ± 0.000</td>
<td>0.006 ± 0.000</td>
</tr>
</tbody>
</table>

Note.—Data are ± 1 standard deviation and, except where noted, are measured in micrograys. NA = not applicable.
tomographic noise also affect image quality, in the absence of other objective measures, recent years have witnessed the emergence of the DQE as the primary index of image quality. Different technologies have different capabilities in that regard; flat-panel detectors have DQEs two to three times higher than those of computed radiography (6,21). CCDs and complementary metal oxide semiconductor devices have offered another solid-state alternative to flat-panel detectors for acquiring digital radiographs. However, because of the low efficiency of their required optical coupling, these detectors generally offer a lower DQE compared with that of flat-panel detectors (18,30,31).

While the DQE describes an important aspect of the performance of an imaging system, because it characterizes the inherent image quality characteristics of the detector employed by a system, it does so in the absence of scattered radiation. As described previously, scatter can account for up to 70% of the detected x-rays in the lung and 95% in the mediastinum on a chest radiograph. Most digital detector systems employ the traditional antiscatter grid or air gap for scatter rejection. However, there is still considerable scattered radiation even with the use of such devices; it accounts for up to 35% of the detected x-rays in the lung and 65% in the mediastinum. This residual scattered radiation reduces image contrast and increases noise. While digital images can be digitally postprocessed to improve contrast, most postprocessing approaches (such as window and level adjustments) also increase the perceivable noise on the image, so that the resulting contrast-to-noise ratio remains limited. Nonlinear postprocessing techniques to remove scatter while maintaining or improving contrast-to-noise ratios have been previously reported but have not been implemented commercially.

A different approach for scatter reduction is the use of scanning radiography with one or more slits or slots (14). While other digital slot-scan mechanisms have been explored in the past, the CCD-based imaging system evaluated in this study was the first commercial digital version of the slot scanning technology for dedicated chest radiography. The system offers a notably narrow slot width (1 cm) with an additional 9.3-cm air gap. Our scatter fraction measurements indicate that the slot-scanning method used with this system offers a notable reduction in scattered radiation, which leads to improved image SNR at a comparable patient dose. Compared with measurements obtained by using a full-field digital chest radiography system with a typical 13:1 antiscatter grid, the slot-scan system yielded scatter fractions that were less than one-half of those in the mediastinal and subdiaphragmatic regions and less than one-third of those in the lung regions.

The scatter fractions measured for this slot-scan system compare favorably with those of previous systems with a similar geometry. According to previous measurements obtained by our group, a commercial analog slot-scan system known as advanced multiple-beam equalization radiography, or AMBER, did not reduce scatter fractions in comparison with a full-field system when used without exposure equalization. Sorenson et al (14) reported on a noncommercial system with a 1-cm scanning slot, 18-cm air gap, and 12:1 grid. Compared with the results of the current study, the scatter fractions from that previous study were higher in the lung (0.21) and comparable or lower in denser regions (0.30–0.39). Recalling that the current system uses an approximately 1-cm scanning slot, a smaller 9.3-cm air gap, and no grid, the system likely has superior design elements and better tuning of the beam and detector slots, thereby providing effective scatter rejection comparable to or better than those in previous reports (ie, 0.13 scatter fraction in the lung and 0.35–0.43 in denser regions).

While our findings indicate a notable difference between the scatter rejections of the tested slot-scan and full-field imaging systems, that difference should be attributed to the image acquisition geometry used with the two systems and not to the specific performance of the detectors employed reflected in their relative DQE. The slot-scan system offers a DQE 2.5 times lower than that of the full-field system (6,18). However, our measurements indicate that the system imparts two to three times less scatter and a 34% lower overall examination dose. These observations suggest that relying on the DQE alone as a single index of image quality per unit exposure can lead to misleading conclusions when comparing the inherent performance of
various digital radiography systems. The limitation of the DQE stems from two aspects of the metric as follows: (a) The DQE is often used to characterize the performance of the detector alone, without the inclusion or consideration of other elements of the imaging system such as the antiscatter grid or air gap, and (b) the DQE only involves the detector response in terms of the detection of primary x-ray radiation. No scattering medium is present in the beam when the DQE measurements are obtained. As a consequence, although the DQE is an effective metric to characterize the intrinsic performance of an imaging detector, it does not reflect an important and ever present component of radiologic imaging and therefore falls short of characterizing the performance of an imaging system. Reliance on the DQE alone can lead to biases in rank ordering of different radiography systems (30,32).

The effective DQE metric defined in this article includes the effect of scattered radiation and transmission through supradetector elements, such as grids, on image quality. Our results show that while the native DQE of the slot-scan system is lower than that of the full-field system, the reduced scatter detection of the slot-scan system can compensate for its lower DQE, which leads to equivalent (in the lungs) and even improved (in the mediastinal and subcardiac-subdiaphragmatic areas) image quality in terms of SNR per unit exposure. Our dose measurements are, in general, consistent with this conclusion. However, it should be noted that the measured dose figures may not be directly compared with the effective DQE figures because the dose values were obtained with the systems in the photo-timed mode, while effective DQE values were exposure-normalized, reflecting the achievable SNR per unit ex-
posure. If a pseudo-linear relationship is assumed between effective dose and exposure in the 100–150 kVp range, the reported effective DQE figures can be used to calibrate the photo-timers to achieve consistent SNR values for clinical images across different imaging systems.

To our knowledge, this article is the first to present digital chest dosimetry results in terms of individual organ doses; this departs from a traditional skin entrance exposure description. Furthermore, we estimated the effective dose figures by using the International Commission on Radiological Protection Publication 60 (28) weighting factors. Although individual organ doses are on the order of tens of microgray, the TLD data showed clear differentiation of organ dose between the slot-scan and full-field digital systems. It is of interest to observe the dose reduction between right and left lungs and between right and left breasts in lateral views. The sum of posteroanterior and lateral effective dose figures was 0.057 mSv and 0.087 mSv for the slot-scan and full-field digital systems at their clinical settings of 140 and 120 kVp, respectively. The difference was statistically significant ($P < .05$). In terms of risk, these values may pose negligible risks, but they serve as a useful dose index in comparing two digital chest systems.

The dose and scatter results reported from our study and the associated effective DQE are only relevant to a population of average-sized patients. However, in the patient population imaged in the United States, there are increasing numbers of patients with larger body mass. Because the scatter fraction increases with patient size, the removal of scattered radiation for larger patients would be even more important. While we have not obtained scatter measurements for larger body types, the dense regions of the chest phantoms examined demonstrated relatively larger scatter fractions. The data showed that the relative advantage of the slot-scan system is more pronounced for larger scatter fraction regions. Thus, it is reasonable to predict that the slot-scan system would be even more effective for larger patients.

Notwithstanding these implications in terms of scatter removal, full-field digital radiography systems offer some potential advantages in terms of image acquisition and aptitude for advanced x-ray–based imaging applications. The full-field digital radiography systems would require a lower equivalent of tube loading than would an equivalent slot-scan system, since the x-ray flux in the latter system would need to be maintained during a longer period of image acquisition (1.3 seconds in the case of the system tested in this study vs a few hundred milliseconds for a full-field system). Contrary to common notion, however, the longer acquisition time does not translate into potential for more patient motion artifacts. The reason is that at any incremental period of time during the slot-scan image acquisition (ie, a dwell time of 32 msec; Table 1), only a small fraction of the image (ie, an approximately 1-cm-wide horizontal band) is acquired. If the patient moves during that time interval, only a 1-cm-tall portion of the image will be blurred, whereas with the full-field system, the entire image will be affected. In addition, the current implementations of dual-energy and digital tomosynthesis imaging modalities require a large-area detector for rapid acquisition of multiple sequential images. The future developments in advanced applications with slot-scan detector systems may alter these conclusions regarding the relative merits of the two imaging methods.

The present study had limitations that should be taken into consideration. First, the effective DQE calculations did not take into account the spatial frequency component of scattered radiation and the effect of the grid on the frequency content of the image. The concept of effective DQE ideally should be extended beyond zero frequency through direct measurements. Second, because of design considerations, we were unable to use the same phantom for the scatter and dose measurements. While the figures are generally comparable, they might not be quantitatively related. These limitations will provide opportunities for future research in this area.

In summary, compared with conventional chest imaging systems with anti-scatter grids, the digital slot-scan imaging system offers a marked advantage in terms of scatter reduction and an associated improvement in the effective DQE, which leads to improved image quality. This advantage is more pronounced in areas of the image with high scatter fraction, such as the mediastinum and the subdiaphragmatic region. Because scatter fraction usually increases as patient body mass increases, it is expected that this advantage will be more pronounced for larger patients. Our subjective comparison of the quality of clinical chest images acquired with the two systems confirms this conclusion.

Acknowledgments: The authors acknowledge David M. Delong, PhD, of Duke University for reviewing the statistical analyses of this study. The authors further thank Greta Toncheva, MS, Giao Nguyen, MS, Lottie Barnes, BS, and Maksudur Sarder, MS, of Duke University for TLD dosimetry support, and Robert Walders, MA, Nicole Ranger, MS, and Tom Boehringer, AAS, of Duke University for additional technical assistance.

References


Exogenous Gene Expression in Tumors: Noninvasive Quantification with Functional and Anatomic Imaging in a Mouse Model

PURPOSE: To assess whether a combination of functional (planar imaging and single photon emission computed tomography [SPECT]) and anatomic (magnetic resonance [MR] imaging) imaging techniques can be used to noninvasively quantify tumor expression of a somatostatin receptor type 2A (SSTR2A) gene chimera in vivo.

MATERIALS AND METHODS: All animal experiments were approved by the institutional animal care and use committee. Expression of the SSTR2A gene chimera was quantified in vitro, in vivo, and ex vivo. The epitope tag of the fusion protein was detected through an antibody, and the receptor portion was detected by using the Food and Drug Administration–approved radiopharmaceutical indium 111 octreotide. Six mice were injected with cells transfected with vector and with two clonal cell lines that each expressed different amounts of the gene chimera. With a dedicated small-animal gamma camera, planar imaging and SPECT were used for quantification of radiopharmaceutical uptake in vivo; 4.7-T MR imaging was used to derive tumor weight. After imaging, excised tumors were evaluated for uptake and weight. For statistical analysis, linear regression analysis, Wilcoxon rank sum test, and Kruskal-Wallis test were employed.

RESULTS: Different expression levels of the chimeric gene were confirmed in vitro. Radiopharmaceutical uptake assessed in excised tumors and that derived from in vivo planar \( r = 0.94, P < .05, n = 18 \) or SPECT \( r = 0.90, P < .05, n = 18 \) images correlated. Weight of excised tumors and that derived from MR images \( r = 0.98, P < .05, n = 18 \) correlated. MR images also allowed morphologic assessment. The biodistribution parameter of percentage of injected dose per gram of excised tumors correlated with the same measure derived from a combination of planar \( r = 0.90, P < .05, n = 18 \) or SPECT \( r = 0.87, P < .05, n = 18 \) images and MR images.

CONCLUSION: A combination of noninvasive functional and anatomic imaging can be used in vivo to quantify gene transfer in tumors.

Although noninvasive monitoring of gene therapy has wide applicability, its use in patients with cancer poses a particular problem because of the variability in the size and morphologic characteristics of tumors. These parameters change in time and may be altered dramatically with therapy. For evaluation of gene expression, destructive in vivo and ex vivo methods are normalized to cell number, protein amount, or tumor weight. However, nondestructive methods for localization and quantification of gene expression are needed in vivo to assess effectiveness, to assess toxic reactions, and to guide dosing regimens (1).

Because of excellent sensitivity, a majority of systems designed to image gene delivery in vivo are based on nuclear medicine techniques. For example, transfer of the herpes simplex virus type 1 thymidine kinase (HSV-TK) gene or the D_{2} dopamine receptor gene is
usually detected by using positron emission tomography (PET) (2–6). In comparison, transfer of the sodium iodide symporter (7), the dopamine transporter (8), or the somatostatin receptor type 2 gene (9,10) has been detected with gamma camera imaging.

Although a small number of reporter systems have been designed for detecting gene expression in vivo, most do not use radiopharmaceuticals that are Food and Drug Administration–approved for clinical use (2–6). Indium 111 (111In) octreotide currently is used in humans to detect tumors (eg, rare neuroendocrine tumors) that express the somatostatin receptor type 2 gene (11). Analogues labeled with technetium 99m (99mTc) also are available clinically. Both 111In- and 99mTc-octreotide analogues are approved by the Food and Drug Administration for human use. Previously, we demonstrated imaging of a somatostatin receptor type 2A (SSTR2A) gene chimera (10). The fusion protein product contains an epitope tag that can be detected with an antibody. This affords immunologic techniques and markedly decreases the expense of analysis. The receptor portion is amenable to radiopharmaceutical-based imaging in vivo. Expression in vitro assessed by using an antibody to the epitope tag correlates with biodistribution of 111In-octreotide evaluated in excised tumors (10).

Although powerful, nuclear medicine–based techniques lack true anatomic detail. These functional methods rely on detection of a radiopharmaceutical for localization of the reporter gene product, not on the underlying anatomy. For small-animal imaging, machines are often used near the limit of their anatomic resolution capabilities. Thus, the images obtained with these techniques are susceptible to volume-averaging artifacts that occur when the resolution of the object of interest is less than 2.7 times that of the imaging system (12). Dedicated small-animal PET units and gamma cameras have resolution limits in the range of a few millimeters.

In comparison, imaging can now be routinely performed at submillimeter resolution by using anatomic methods such as magnetic resonance (MR) imaging. Because MR imaging affords excellent soft-tissue visualization, both the size and internal structure of the lesion may be evaluated. Thus, we assessed whether a combination of functional and anatomic techniques can be used to noninvasively quantify tumor expression of an SSTR2A gene chimera in vivo.

**MATERIALS AND METHODS**

**Plasmid**

As described previously, the full-length human SSTR2A gene was inserted into a vector (pDisplay; Invitrogen, Carlsbad, Calif), downstream of the membrane localization sequence and the hemagglutinin A epitope tag sequence (10).

**Cell Line Production and Characterization**

HT1080 cells of human fibrosarcoma (American Tumor Cell Collection, Rockville, Md) were grown in Dulbecco modified Eagle medium containing 1× glutamine, penicillin, and streptomycin and 10% fetal bovine serum. For transfection, 1 μg DNA was added with liposomes (Lipofectin 2000; Invitrogen) to 1 × 10^5 cells according to the manufacturer’s instructions. After 5 hours, the liposome-DNA solution was removed, and the cells were incubated in growth medium. After G418 selection, single colonies were isolated. Prospective colonies were assayed for gene expression by using the enzyme-linked immunosorbent assay, and then individual clones were assessed for gene expression quantitatively. Enzyme-linked immunosorbent assay and receptor-binding studies were performed by two authors (D.Y., L.H.) as described previously (10).

**Western Blot Analysis**

For cells, confluent six-well dishes were exposed to a lysis solution (0.1% sodium dodecyl sulfate, surfactant [1% Triton X-100; Pfizer Scientific, Fairlawn, NJ], 0.1 mol/L tris [pH 8], 0.14 mol/L sodium chloride, 0.025% sodium azide, and 0.18% protease inhibitor cocktail [Complete Protease Inhibitor; Roche, Mannheim, Germany]) for 1 hour at 4°C. After 30-minute centrifugation at 14 000g, the supernatant was collected. For samples, samples were washed with phosphate-buffered saline and homogenized with 10 strokes in the lysis solution. After 15 minutes of centrifugation at 14 000g, the supernatant was collected. Protein concentration was determined by using the Bradford method (Bio-Rad Laboratories, Hercules, Calif). Twenty micrograms of cell protein or 50 μg of tissue protein was loaded per lane on 7% sodium dodecyl sulfate gels. The sample was then transferred to nitrocellulose by using a semidry apparatus (Fisher, Atlanta, Ga). Equal transfer was confirmed with Ponceau-S staining.

Next, the membrane was blocked and exposed to 50 mU/mL horseradish peroxidase–rat-anti-hemagglutinin A antibody overnight at 4°C or 1 hour at room temperature. After four 10-minute washes with phosphate-buffered saline, the membrane was covered with a chemiluminescent horseradish peroxidase substrate (Perkin Elmer Life Science, Boston, Mass) and exposed to film (D.Y.).

**Immunohistochemical Analysis**

Paraffin-embedded tumor tissue was probed with a 1:1000 dilution of a mouse anti-hemagglutinin A antibody (Babco, Richmond, Calif), washed, and then
stained with a horseradish peroxidase–conjugated antimouse secondary antibody. The tissue was counterstained with the Giemsa stain. The presence of a brown reaction product at the periphery of the cells was considered a positive reaction for the fusion protein (evaluated by a veterinary pathologist and an author [V.K.]).

Biodistribution and Imaging

All animal experiments were approved by the institutional animal care and use committee. In nine nude mice (approximate age, 8 weeks; weight, 25 g), subcutaneous injection of $5 \times 10^6$ cells produced palpable tumors after 1 week. Each mouse received three inoculations of tumor cells: right thigh, cells transfected with vector; right and left shoulders, clone 309 and clone 301 cells, respectively. These tumor cells express different levels of the same gene chimera. Next, six of the nine mice were randomly selected for injection of $13 \text{ MBq (350 } \mu \text{ Ci)}$ of $^{111}\text{In-octreotide (Mallinckrodt, St Louis, Mo)}$ into the tail vein.

Twenty-four hours later, anesthetized animals were imaged with a 4.7-T small-animal MR unit (Bruker, Billerica, Mass) by using a T2-weighted fast spin-echo sequence (repetition time msec/echo time msec, 4120/72; signals acquired, four; field of view, 3.5 cm; section thickness, 1 mm; intersection gap, 0.3 mm; matrix, 256 $\times$ 256; spatial resolution, 136 $\mu$m). Tumor measurements were performed by using software (Image J; National Institutes of Health, Bethesda, Md). On each image containing a tumor, the periphery of the mass was traced, and the area of the drawn region was calculated by one investigator (V.K.). The areas were then multiplied by the section thickness plus intersection gap to obtain the volume of each section that contained the object of interest. Each section volume was then added. To control for volume averaging, however, only one-half of the volume of the most superior images and one-half of the volume of the most inferior images that contained the object of interest were added. With the assumption of a tissue density of 1 g/mL, to derive weight, the volume of the object of interest in cubic millimeters was multiplied by 0.001 g of tissue per cubic millimeter. The same process was used to trace and calculate the weight of necrotic and/or hemorrhagic material, which was identified on the MR images as areas of increased or decreased signal intensity compared with tumor and containing fluid-fluid or fluid-debris levels. The weight of the necrotic and/or hemorrhagic material was then subtracted from the weight of the mass to calculate the corrected weight.

Next, the mice underwent planar imaging for 10 minutes by using a gamma camera (mCAM; Siemens Medical Solutions, Hoffman Estates, Ill) fitted with a medium-energy parallel-hole collimator. No attenuation correction was used. For single photon emission computed tomography (SPECT), 15-minute acquisitions were performed. Imaging consisted of 120 views (imaging time of 7.5 seconds per view, 128 $\times$ 128 matrix, pixel size of 2.4 mm) over a 360° rotation of a fixed $\frac{1}{15}$ rpm rotational device attached to the front of the collimator. Each animal was spun in the rotational device. For SPECT image reconstruction, back-projection was performed with 10th-order Butterworth filter with a cutoff of 1.2 cycles per centimeter. On planar and SPECT images, region-of-interest (V.K.) total count measurements were normalized to the number of pixels in the region of interest to obtain average counts per pixel. For the planar images, these values were subtracted from those obtained from the left thigh, which did not have a tumor. For both techniques, the values were then converted to counts per minute by using an equation derived from phantoms containing relevant amounts of activity. Phantoms consisted

Figure 2. Graph shows ex vivo biodistribution analysis after intravenous administration of $^{111}\text{In-octreotide. Greater biodistribution of the radiopharmaceutical was seen in tumors derived from clone 309 cells than was seen in those derived from clone 301 cells or from cells transfected with vector. Error bars represent standard deviation ($n = 6$ mice). * = vector versus clone 301 cells ($P < .05$), ** = clone 301 cells versus clone 309 cells ($P < .05$), %I.D./g = percentage of injected dose per gram.
of 1.5-mL Eppendorf tubes containing 500 μL of different amounts of 111In-octreotide (13 MBq) and imaged 24 hours later. A, Planar image. B, Coronal SPECT image. C, Transverse SPECT image. D, Sagittal SPECT image. Subcutaneous tumors derived from clone 309 cells (arrow), clone 301 cells (white arrowhead), or cells transfected with vector (yellow arrowhead) were in the right shoulder, left shoulder, or right thigh, respectively. All three tomographic planes were centered on tumor derived from clone 309 cells. Planar and SPECT images were obtained in the same representative mouse. E, F, Graphs show that counts in excised tumors correlate with measurements derived from region-of-interest analysis of planar (E, n = 18) or SPECT (F, n = 18) images. cpm = counts per minute. G, Image shows that the size of the object on the planar image may not correlate with the true anatomic size of the object. Planar image (top) of phantoms (bottom) of the same size contain 500 μL of serial 1:1 dilutions of 111In-chloride (93.0–0.03 μCi [3.44–0.001 kBq]).

Figure 3. Images obtained with gamma camera and graphs of imaging results. By using both planar imaging and SPECT, tumors that expressed the gene chimera were visible, whereas tumors derived from cells transfected with vector were not. Mice bearing subcutaneous tumors (in both shoulders and right thigh) were injected intravenously with 111In-octreotide (13 MBq) and imaged 24 hours later. A, Planar image. B, Coronal SPECT image. C, Transverse SPECT image. D, Sagittal SPECT image. Subcutaneous tumors derived from clone 309 cells (arrow), clone 301 cells (white arrowhead), or cells transfected with vector (yellow arrowhead) were in the right shoulder, left shoulder, or right thigh, respectively. All three tomographic planes were centered on tumor derived from clone 309 cells. Planar and SPECT images were obtained in the same representative mouse. E, F, Graphs show that counts in excised tumors correlate with measurements derived from region-of-interest analysis of planar (E, n = 18) or SPECT (F, n = 18) images. cpm = counts per minute. G, Image shows that the size of the object on the planar image may not correlate with the true anatomic size of the object. Planar image (top) of phantoms (bottom) of the same size contain 500 μL of serial 1:1 dilutions of 111In-chloride (93.0–0.03 μCi [3.44–0.001 kBq]).
Influence of Radioactivity

For assessment of the influence of the amount of radioactivity on image representation, 1.5-mL Eppendorf tubes were filled with 0.5 mL of phosphate-buffered saline containing serial 1:1 dilutions of 111In-chloride from 93.0 to 0.03 μCi (3.44–0.001 kBq). Planar imaging was performed as described previously. By using a variety of background and saturation display levels, the size of the phantom on the image was visually compared with the size of the actual phantom by one author (V.K.).

Statistical Analysis

Linear regression was used to analyze correlations of data. The analyses were performed by using spreadsheet software (Excel 2000; Microsoft, Bellevue, Wash). By using different software (SAS 2001, version 8.02; SAS Institute, Cary, NC), the Kruskal-Wallis test was used to compare trends in gene expression that differed among tumors. The Wilcoxon rank sum test (one-tailed) was used to compare gene expression in vitro or uptake among tumors that differed in gene expression in vivo. For all tests, a difference with $P < .05$ was considered significant.

RESULTS

Gene Expression

By using an antibody to the hemagglutinin A domain of the fusion protein, gene expression was confirmed in whole cells with the enzyme-linked immunosorbent assay (normalized for cell number) and in cell lysates with the Western blot analysis (normalized for protein). In a quantitative enzyme-linked immunosorbent assay of clonal cell lines transfected with the same SSTR2A gene chimera (Fig 1, A), clone 309 cells reacted more than did clone 301 cells. In comparison, no reaction was seen in cells transfected with vector only. As seen at Western blot analysis (Fig 1, B), a distinct band was observed in all lanes that represented cells transfected with the SSTR2A gene chimera but not in lanes that represented cells transfected with vector only. The band was more intense in the lane marked clone 309 cells than it was in the lane marked clone 301 cells, and this finding implied greater expression in clone 309 cells. No expression was seen in cells transfected with vector only.

To confirm proper function of the SSTR2A portion of the fusion protein, receptor-binding assays were performed. For the assay, a saturation dose of $10^{-7}$ mol/L 111In-octreotide was used (13,14). Because both HT1080 cell clones express the same fusion protein, any difference in the degree of binding is caused by the amount of gene expression. As seen in Figure 1, C, binding to 111In-octreotide was greater for clone 309 cells than it was for clone 301 cells. Unlabeled somatostatin competes with the labeled analogue, thus confirming specificity. No specific binding was seen in cells transfected with vector only. The receptor-binding data corresponded with the results of enzyme-linked immunosorbent assay and of Western blot analysis. Thus, expression levels that were based on the anti-hemagglutinin A antibody corroborated receptor-binding data that were based on 111In-octreotide. Data demonstrate that a greater amount of fusion protein per cell was present in clone 309 cells than was present in clone 301 cells and that the hemagglutinin A and SSTR2A domains of the fusion proteins were functional.

Biodistribution

Figure 2 demonstrates ex vivo biodistribution analysis of 111In-octreotide 24 hours after injection into the tail vein in six nude mice. Each mouse bore three subcutaneous tumors derived from clone 309 cells, clone 301 cells, or vector-transfected cells. Unbound 111In-octreotide is eliminated through the kidneys and liver, and among the mice, there was variability in the extent of excretion through each organ. When we compared findings in excised tumors, there was a statistically significant difference in the biodistribution parameter of percentage of injected dose per gram between tumors that originated from vector-transfected cells or those that originated from gene chimera–transfected cells (vector vs clone 301 cells, $P < .05$, $n = 6$; vector vs clone 309 cells, $P < .05$, $n = 6$) and between tumors from clone 301 cells and those from clone 309 cells ($P < .05$, $n = 6$). Thus, tumors that expressed different levels of the fusion protein were distinguished by using the invasive biodistribution methods.

Functional Imaging

Tumor-bearing nude mice were imaged 24 hours after injection of 111In-octreotide into the tail vein but before

---

**Figure 4.** Images obtained in a nude mouse and resultant graph. MR imaging helped to define the anatomic size of the object and depicted its internal morphologic features. (a) Representative T2-weighted fast spin-echo MR image (4120/72; signals acquired, four; field of view, 3.5 cm; section thickness, 1 mm, with 0.3 mm intersection gap; matrix, 256 × 256; resolution, 136 μm) of a subcutaneous tumor (arrow) in the right thigh. Bladder (arrowhead) is identified. (b) Graph shows that weight derived from MR images correlates with weight of the excised tumor ($n = 18$). (c) Representative T2-weighted fast spin-echo MR image obtained in a subcutaneous tumor in the left shoulder with same parameters as in a. Arrow points to one of the fluid-debris levels within the tumor. For MR imaging, mice were placed on the abdomen; images (a and c) reflect this positioning to demonstrate the fluid-debris levels in c.
Figure 5. Graphs show correlation of uptake with invasive and noninvasive methods. The percentage of injected dose per gram (\%/I.D./g) of excised tumor (tissue) correlates with injected dose per gram derived from images. (a) Planar image \((n = 18)\) combined with MR image. (b) SPECT image \((n = 18)\) combined with MR image.

The sacrificing of the animals for the biodistribution analysis. Results of the biodistribution analysis are depicted in Figure 2. A planar image obtained in a representative mouse (Fig 3, A) demonstrates that the tumor derived from clone 309 cells that expressed more fusion protein was better visualized than the tumor derived from clone 301 cells. Tumor derived from cells transfected with vector appeared similar to the background. To improve localization, tomographic methods were applied. SPECT imaging of gene expression in a representative mouse (Fig 3, B–D) demonstrated the tumors that expressed the fusion protein, whereas tumors derived from vector-transfected cells appeared similar to the background. For both planar imaging \((r = 0.94, P < .05, n = 18)\) (Fig 3, E) and SPECT \((r = 0.90, P < .05, n = 18)\) (Fig 3, F), uptake derived from region-of-interest analysis of the images correlated with radioactivity associated with excised tumors (coefficient of the \(x\) variable = 1.8 and 1.3, standard error of the \(x\) variable = 0.2 and 0.2, respectively). Among the imaging methods, the correlation coefficient that showed the correlation between the planar imaging and SPECT techniques was 0.96 \((P < .05, n = 18)\), coefficient of the \(x\) variable = 1.3, standard error of the \(x\) variable = 0.1. The size of the object on the functional image obtained with the gamma camera, however, may not have reflected the true size of the object.

The planar image in Figure 3, G, demonstrates that although all of the phantoms are of the same size, the apparent size of each phantom on the image increased with increasing amounts of radioactivity. In addition, there is apparent heterogeneity in the color representation of each object, although the phantoms have a uniform distribution of the radiopharmaceutical.

Anatomic Imaging

Before the animals were sacrificed for the ex vivo biodistribution analysis, the results of which are depicted in Figure 2, the mice also underwent MR imaging. The image at the level of the pelvis in a representative mouse (Fig 4a) shows that the tumor had intermediate signal and demonstrated excellent contrast with adjacent structures. Weight derived from volume assessment of the entire mass on the tomographic MR images (Fig 4b) correlated with the weight of the excised tumors \((r = 0.98, P < .05, n = 18)\), coefficient of the \(x\) variable = 1.03, standard error of the \(x\) variable = 0.06). Significant correlation was not seen between the weight of the excised tumor and the weight derived from volume assessment of the mass on the SPECT images \((r = 0, P > .05, n = 18)\) or area derived from planar images \((r = 0, P > .05, n = 18)\).

In some \((11 of 18)\) tumors, high T2 signal greater than that within the soft tissue of the mass was seen (Fig 4c). In addition, layering of low T2 signal was noted within these regions of high signal, a finding that was consistent with fluid-fluid or fluid-debris levels caused by hemorrhage and/or necrosis. Because these areas were not expected to contain significant numbers of live cells that expressed the fusion protein, they were subtracted for calculation of the corrected tumor weight.
Noninvasive versus Invasive Assessment of Uptake

The association between the in vitro and in vivo findings was further examined by using regression analysis. The biodistribution parameter (Fig 5) of percentage of injected dose per gram, evaluated by using excised tumors, correlated with the noninvasive image-derived values obtained with planar imaging \((r = 0.90, P < .05, n = 18, \text{coefficient of the } x \text{ variable} = 1.8, \text{standard error of the } x \text{ variable} = 0.2)\) or SPECT \((r = 0.87, P < .05, n = 18, \text{coefficient of the } x \text{ variable} = 1.5, \text{standard error of the } x \text{ variable} = 0.2)\) techniques for determination of uptake and with MR imaging for determination of the weight of the entire tumor.

The biodistribution of \(^{111}\text{In-octreotide}\) among excised tumors was distinguishable by using the percentage of injected dose per gram (Fig 6a) parameter. Normalization, instead, to the corrected weight derived from MR imaging allowed differentiation among tumors on the basis of biodistribution (Fig 6b). This was also found when completely image-based parameters were derived by using planar imaging (Fig 6c) or SPECT (Fig 6d) for assessing uptake and by using MR imaging for corrected tumor weight. With each method, as depicted in Figure 6, by using either excised tumors or the in vivo image-derived parameters, tumors that originated from clone 309 cells had statistically significant greater expression than did those that originated from clone 301 cells \((P < .05, n = 6)\); both of these demonstrated greater \(^{111}\text{In-octreotide}\) biodistribution than did tumors that originated from cells transfected with vector (vector vs clone 301 cells, \(P < .05, n = 6\); vector vs clone 309 cells, \(P < .05, n = 6\)). The Kruskal-Wallis test also indicated a difference in gene expression with any of the methods employed \((P < .05)\).

Ex Vivo Gene Expression Analysis

To further validate gene expression ex vivo, on the same day as the imaging, the three additional tumor-bearing mice were sacrificed. Portions of the excised tumors were analyzed by using Western blot analysis with an antibody to the hemagglutinin A tag portion of the fusion protein. Expression was greater in tumors derived from clone 309 cells than from those from clone 301 cells (Fig 7, A). No band was seen in tumors derived from cells transfected with vector only. Results of immunohistochemical analysis (Fig 7, B) confirmed that expression in tumors derived from clone 309 cells was greater than that in tumors derived from clone 301 cells. As expected, staining was seen at the periphery of the cells, which is consistent with cell membrane localization of the fusion protein. Background staining was seen in tumors derived from cells transduced with vector. These data, which demonstrate varying degrees of gene expression among the tumors derived from the clone cells, were consistent with the in vitro and in vivo findings.

DISCUSSION

To optimally use gene therapy for cancer treatment, methods for quantification of gene expression are needed to assess effectiveness, to assess toxic reactions, and to guide dosing regimens (1). After gene transfer, the degree of gene expression changes as a function of time. At transduction, there is a lag period. Gene expression then increases and may plateau before it wanes. Tracking such changes will prove fruitful because, as with drug therapy, a certain amount of the gene product will be needed to obtain a therapeutic effect. Theoretically, reporter genes may be used to monitor the level of gene expression. In vivo, one would place a region of interest over the target and then quantify the signal that reflects the amount of reporter gene present.

Such a strategy alone, however, is not sufficient for cancer treatment because, unlike organs in adults, tumors grow, and if therapy is successful, tumors regress. This change in size poses a problem for quantification of gene expression because the signal change may be a result of the number of cells present instead of the efficiency of induction. For example, in a time course experiment, apparently decreased uptake at a subsequent time point may be caused by a smaller tumor, less gene expression, or both. A combination of functional and anatomic imaging will allow one to distinguish these possibilities by normalizing uptake to tumor size. This will be beneficial when one uses reporter genes for assessment of the effectiveness of dosage regimens of conventional therapy or of gene therapy, in which a therapeutic gene is delivered in conjunction with a reporter gene. Common in vivo gene delivery vectors result in relatively inhomogeneous and temporally variable expression. For example, by using an adeno-associated virus for delivery, onset of gene expression may require
2–6 weeks (15,16). In comparison, by using adenovirus for delivery, gene expression occurs in days but then often decreases after approximately 2–3 weeks (17); however, multiple temporarily separated injections can improve therapeutic effectiveness (18). During this time and at subsequent examinations, the tumor may grow or regress.

This growth or regression becomes particularly important if the reporter gene is used for in vivo assessment of promoter activity. By using the tetracycline-inducible system, the HSV-TK reporter gene has been imaged at induction of a tetracycline-responsive promoter (19). By using the HSV-TK reporter gene, Qiao et al (20) demonstrated tumor-specific targeting with the carcinoembryonic antigen gene promoter. Findings in these studies imply that in vivo assessment of transcriptional regulation is possible. To assess changes in gene expression over time and to compare the activity among promoters, functional assays need to be normalized to cell number or target size.

At any time in the experimental setting, tumors almost always vary in size; however, a number of such lesions need to be assessed to obtain statistically valid results. For assessment of gene expression with functional means, the degree of gene expression in each tumor should be normalized to weight. Similarly, variability in the size of organs also requires that normalization be performed when gene expression among individuals is compared.

The ultimate aim of oncology is to kill the cancer. Tumor necrosis occurs when large tumors outgrow their blood supply or when therapy is effective. After gene therapy, the portion of the tumor that undergoes necrosis will not express the reporter gene and, thus, should be excluded in the assessment of gene expression. By using anatomic imaging techniques, such as MR imaging, these areas can be excluded when size or weight is assessed. This is superior to simply measuring the size of a mass with calipers because there may only be a shell of live cells that encases necrotic or liquid material. With anatomic imaging techniques such as MR imaging, the areas that do not contribute to the functional signal may be excluded.

To our knowledge, our study is the first to demonstrate SPECT imaging of a reporter gene product that is based on the somatostatin receptor. Tomographic imaging techniques improve localization of signal compared with planar imaging techniques, in which the object is reduced to two dimensions. By using a clinical gamma camera, Tjuvajev et al (21) performed SPECT imaging of tumors that expressed the HSV-TK gene in rats. Auricchio et al (8) demonstrated SPECT imaging of a dopamine transporter gene expressed in the thighs of mice. We performed quantitative analysis of images by using tumor-bearing mice. By employing PET, gene expression of the HSV-TK gene has been normalized to weight by using calculations of object size that were based on region-of-interest analysis of the PET image (22). We did not find that we could adequately assess tumor size in mice by using functional planar imaging or SPECT alone. On functional images, the size of the object may vary according to the amount of radiopharmaceutical present, as we noted by using phantoms. In addition, for small-animal imaging, the size of the object often is near the resolution limit of the camera, and this close relationship results in partial-volume effects. Even dedicated small-animal nuclear medicine imaging systems have spatial resolution limits of several millimeters. In mice, tumors are usually smaller than 1 cm in diameter at analysis.

The submillimeter resolution of MR imaging provided accurate measurement of tumor volume that was used to derive tumor weight. Knowing the tumor volume is also beneficial. For example, it may be used to determine what volume of fluid is needed to disperse the therapeutic substance throughout the tumor at direct injection.

A combination of functional and anatomic information obtained with imaging allowed normalization of gene expression to tumor weight. The in vivo image-derived parameters correlated with ex vivo findings obtained from excised tumors to quantify uptake per unit weight. MR imaging also allowed exclusion from the weight calculation of those portions of the tumor (ie, hemorrhage and/or necrosis) that were not contributing to the functional signal. These areas may be further delineated by using contrast material–enhanced imaging. When data from SPECT or planar images were combined with data from MR images, tumors that expressed high, low, and no levels of the fusion protein could be statistically separated. To validate these concepts, we derived tumors from cells that constitutively expressed the hemagglutinin A-STR2A fusion protein. To address what some may consider a limitation of this study, in future experiments, we will use common in vivo gene delivery systems, such as adenovirus, for reporter gene transfer into tumors.

Data suggest that noninvasive imaging criteria can be substituted for invasive methods for following gene expression in tumors. In addition, morphologic data can be used to identify and exclude regions of the tumor that cannot contribute to gene expression. We demonstrated that these measurements can be obtained in small animals such as mice. A combination of functional and anatomic imaging techniques for noninvasive determination of biodistribution should also be applicable to other functional imaging methods, such as PET and optical imaging, as well as to other anatomic modalities such as computed tomography. Because most of these instruments are available for patient evaluation, these techniques should have clinical utility.

References
Iron Oxide Nanoparticle–labeled Rat Smooth Muscle Cells: Cardiac MR Imaging for Cell Graft Monitoring and Quantitation

**PURPOSE:** To perform a quantitative analysis of anionic maghemite nanoparticle–labeled cells in vitro and determine the effect of labeling on signal intensity at magnetic resonance (MR) imaging.

**MATERIALS AND METHODS:** The study was approved by the institutional animal care and use committee at Hôpital Bichat. In vitro cell proliferation, iron content per cell, and MR signal intensity of cells were measured in agarose phantoms for 0–14 days of culture after labeling of rat smooth muscle cells with anionic maghemite nanoparticles. Next, iron oxide–labeled smooth muscle cells were injected into healthy hearts and hearts with ischemic injury in seven live Fisher rats. Ex vivo MR imaging experiments in excised hearts 2 and 48 hours after injection were performed with a 1.5-T medical imaging system by using T2-weighted gradient-echo and spin-echo sequences. Histologic sections were obtained after MR imaging. Correlation analyses between division factor of iron load and cell amplification factor and between 1/T2 and number of labeled cells or number of days in culture were performed by using linear regression.

**RESULTS:** Viability of smooth muscle cells was not affected by magnetic labeling. Transmission electron micrographs of cells revealed the presence of iron oxide nanoparticles in vesicles up to day 14 of culture. Intracellular iron concentration decreased in parallel with cell division ($r^2 = 0.99$) and was correlated with MR signal intensity ($r^2 = 0.95$). T2*-weighted MR images of excised rat hearts showed hypointense signal in myocardium at 2 and 48 hours after local injection of labeled cells. Subsequent histologic staining evidenced iron oxide nanoparticles within cells and confirmed the presence of the original cells at 2 and 48 hours after implantation.

**CONCLUSION:** Magnetic labeling of smooth muscle cells with anionic maghemite nanoparticles allows detection of cells with MR imaging after local transplantation in the heart.

© RSNA, 2005

Cell therapy appears a promising field for the treatment of human diseases. As part of this new field, transplantation of smooth muscle cells has undergone extensive investigation in recent years as a potential therapy, mainly for repair of aneurysm (1) or myocardial ischemia (2) or for cardiac graft (3). Lack of ability to track the cell transplants, however, remains a major problem that must still be overcome to understand and optimize cell therapy. In fact, most cell transplantation techniques involve the use of histologic analysis to evaluate cell transfection, proliferation, and migration (4–6). There is therefore a need to develop noninvasive methods for visualizing transplanted cells.

Investigators in early studies used positron emission tomography to track radionuclide-labeled cells in vivo (7). Improvements in these techniques have since been used to analyze...
metastatic tumor cell trafficking in vivo (8–11). While these techniques are sensitive, they lack high spatial resolution and require the use of nuclear labels. High-spatial-resolution magnetic resonance (MR) imaging performed with contrast agents can overcome this limitation and can be used to study the fate of magnetically labeled cells (12–16). Recent studies have demonstrated that single-cell resolution (17) and stem cell tracking are possible with contrast material–enhanced MR imaging (18).

Considerable advances have been made to enhance the efficiency of labeling and the contrast properties of labeled cells, but there have been few studies of labeling stability or quantitative assays for in vivo cell tracking. Questions such as the following are still in need of clear answers: What is the relationship between signal intensity and the number of labeled cells? How many cells can be detected in vivo? How is signal intensity related to cell proliferation and migration?

Dextran-free anionic magnetic nanoparticles are stable in colloidal suspension and are adsorbed through nonspecific electrostatic interactions with the membranes of most cell types, followed by spontaneous cell internalization (19). This label gives magnetic properties to the entire cell body, properties that potentially allow the detection of cells according to $T2$ and $T2^*$ values measured at MR imaging. Thus, the purpose of our study was to perform a quantitative analysis of iron oxide–labeled cells in vitro and to determine the effect of labeling on signal intensity at MR imaging.

MATERIALS AND METHODS

Our study was approved by the animal care and use committee at Hôpital Bi-chat.

Anionic Magnetic Nanoparticles: Synthesis and Characterization

Stable colloidal suspensions of negatively charged maghemite nanoparticles were used. The precursor iron ferrofluid was synthesized according to the Massart method (20) by alkalinizing an aqueous mixture of iron(II) chloride and iron(III) chloride. The resulting magnetite ($Fe_{3}O_{4}$) particles were then acidified, oxidized in maghemite ($γ-Fe_{2}O_{3}$), and dispersed in water. This process resulted in an acidic ferrofluid composed of magnetic particles positively charged with nitrate counterions. Particles were then chelated (21) with meso-2,3-dimercaptosuccinic acid—HOOC-CH(SH)-CH(SH)-COOH—a process that forms strong complexes. An aqueous solution was obtained that contained thiolated maghemite nanoparticles that bore net negative surface charges mainly because of the unbound carboxylate groups (COO$^-$). This colloidal solution is stable over a wide range of pH values (from 3 to 11) with suitable ionic strengths ($<0.35$ mmol/L) and in various buffers. For this study, the nanoparticles were resuspended in a buffer solution that contained $0.1$ mmol/L HEPES by using serial ultracentrifugation. The nanoparticles consist of monocrystalline ferrimagnetic monodomain maghemite with a crystalline diameter typically of $9$ nm, a magnetic core of $8.7$ nm, and a hydrodynamic diameter of $35$ nm.

Longitudinal and transverse relaxivities (at $1.5$ T, with a flip angle of $25^\circ$) of anionic magnetic nanoparticles dispersed in water are $10$ and $357$ sec$^{-1}$·mM$^{-1}$·L·mol for $r1$ and $r2$, respectively (22).

Ultrasmall superparamagnetic iron oxide particles coated with dextran (Sinerem; Guerbet, Aulnay-sous-Bois, France), of similar size ($30$ nm), were used as a reference.

Smooth Muscle Cell Cultures

Smooth muscle cells from Fisher rat thoracic aorta were isolated as reported previously (23). Cells were cultured in a mixture of RPMI 1640 and M199 media supplemented with $10$% fetal calf serum, $50$ U/mL penicillin, $40$ μg/mL streptomycin, and $0.3$ mg/mL $t$-glutamine (Gibco; Invitrogen, La Jolla, Calif). The presence of smooth muscle cells was confirmed by the appearance of focal overgrowth at confluence microscopy and by a positive result at immunohistochemical analysis for quantification of α-actin.

Magnetic Cell Labeling

Eighty-percent-confluent smooth muscle cells (passages 5–8) from the Fisher rats were incubated for $2$ hours in a filter-sterilized ($0.1$-$μm$ filter) suspension of anionic magnetic nanoparticles ($0.1$, $1.0$, $5.0$, and $10.0$ mmol/L iron) in $0.1$ mmol/L HEPES at $37^\circ$C. During incubation, anionic magnetic nanoparticles adsorbed to the cell membrane surface by means of electrostatic interactions and were internalized by the cell through an endocytic pathway (24). Cells were then washed and incubated for $1$ hour in serum-free medium for chasing. After labeling and chasing, cells were washed in phosphate-buffered saline with pH of $7.4$, treated with trypsin, washed again, and resuspended in culture medium by two authors independently (C.R., J.F.D.).

Quantification of Magnetic Labeling

Iron content in cells was quantified (C.R.) by using two methods: magnetophoresis and electron spin resonance (19,25). Briefly, magnetophoresis consists of measurement of the velocity of labeled cells attracted to a permanent magnet with a magnetic field of known configuration. Since the balance of magnetic and viscous forces was known, we were able to calculate the iron content of each visualized cell. To achieve reliable results, we repeated this measurement for $100$ different cells. In electron spin resonance, the resonance spectrum of a known number (ie, $10^6$) of labeled cells is recorded. Since the spectral area is proportional to total iron mass, the mean iron content per cell can be calculated. The two methods of measurement provide complementary results: Magnetophoresis provides a measure of the distribution of the whole iron load among the cells, while electron spin resonance provides a measure of average iron load determined from the cell population. The latter method yields a more reliable result if the distribution is broad or the iron content per cell is small.

Growth Kinetics and Iron Contents of Smooth Muscle Cell Cultures

To monitor the fate of the magnetic nanoparticles during cell culture, the iron content per cell was analyzed and quantified with consensus by two authors (C.R., J.F.D.) according to the following procedure: After labeling with the magnetic nanoparticles ($1.0$ mmol/L suspension), cells were cultured for $0$–$14$ days, with data for the different durations of culture being derived from independent cell cultures obtained from the same initial cell sample. On the indicated days, cells were treated with trypsin, washed, and resuspended in culture medium. The cell cultures were then divided into two parts: one part to be used for quantification of iron concentration and kinetic analysis of cell growth, and the other, for MR imaging. Viable trypan blue–stained smooth muscle cells were counted with a Mallas chamber. Control cells, which were not labeled with anionic magnetic nanoparticles but were incubated in $0.1$ mmol/L HEPES solution for the same time as labeled cells, were
Qualitative Visualization of Iron Oxide–labeled Cells

At Perls Prussian blue staining (J.F.D.), labeled cells were fixed with 4% glutaraldehyde, washed with phosphate-buffered saline, incubated for 15 minutes in a solution of 2% potassium ferrocyanide (Perls Prussian blue stain) in 2% HCl, washed, and counterstained with eosin.

In Vitro MR Image Acquisition

The contrast properties of labeled cells were investigated in agarose phantoms that contained a known number of dispersed fixed labeled cells with different durations of culture by using a clinical 1.5-T MR imager (Signa; GE Medical Systems, Milwaukee, Wis.). All samples were placed in a head coil and imaged with a section thickness of 4 mm, field of view of 17.9 × 17.9 cm, imaging matrix of 256 × 256, and two signals acquired. The imaging protocol consisted of coronal spin-echo sequences (J.F.D.). The signal intensity was measured by using a homogeneous 20-mm² circular region of interest. Proton relaxation times T1 and T2 were deduced from fitting of the experimental signal intensity values obtained with the spin-echo sequence as a function of repetition time (TR) (with TR of 25–3000 msec and with fixed echo time [TE]) and of TE (with TE of 10–600 msec and with fixed TR), respectively, by using an exponential law (22). In a similar way, T2* was deduced from fitting of the experimental signal intensity values obtained with the gradient-echo sequence as a function of TE.

T2 and T2* are well described with the following equation:

\[ \frac{1}{T2} = \frac{1}{T2_{Fe=0}} + r2* \cdot Fe, \]

where Fe is the concentration of intracellular iron in the agarose gel and r2* is the transverse (inhomogeneous) relaxivity measured in seconds per millimole per liter. 1/T2* is thus directly proportional to intracellular iron concentration (22).

Animal Study

Male Fisher adult rats (n = 7) with a mean weight of 300 g were used for experiments. Animals were anesthetized with intraperitoneal pentobarbital (0.1 mL per 100 g body weight). Hearts were extracted after lateral thoracotomy. Left ventricular infarction was produced by ligation of the left coronary artery, as previously described (26), in four rats (J.B.M.). Fourteen days after ligation, 10⁶ syngeneic anionic magnetic nanoparticle–labeled (incubated 2 hours with 1.0 mmol/L iron at 37°C, with 0 days of subsequent culture) smooth muscle cells in 0.5 mL of rat serum were injected in the infarcted area (three adjacent locations, 0.16 mL in each) after lateral thoracotomy (J.B.M.). Animals were sacrificed 2 hours (n = 2) or 48 hours (n = 2) after injection.

Three other rats without infarction received either an injection of 10⁶ anionic magnetic nanoparticle–labeled smooth muscle cells (n = 1; three locations in left ventricular myocardium) or no injection (n = 2). Hearts were excised, rinsed in phosphate-buffered saline, fixed for 2 hours with 4% paraformaldehyde, and stored at 4°C in 15-mL tubes.

Ex Vivo MR Image Acquisition

Ex vivo MR imaging of the hearts was performed with the clinical 1.5-T MR imager by using a wrist coil with a diameter of 120 mm. T2-weighted spin-echo sequences were applied along the short axis of the heart and T2*-weighted gradient-echo sequences were applied along the long axis (J.F.D., J.P.L.) by using the following imaging parameters: for the T2-weighted spin-echo sequence, 2000/56 (TR msec/TE msec), 60-mm field of view, matrix of 256 × 256, and section thickness of 2 mm; for the T2*-weighted gradient-echo sequence, 16/2.8, 80-mm field of view, flip angle of 30°, matrix of 256 × 256, and section thickness of 2 mm.

Signal intensity in regions with hypointense signal (S_{AMNP}) that indicated the presence of anionic magnetite nanoparticles on T2-weighted spin-echo images was measured with consensus by two authors (J.F.D., 4 years of experience with cardiac MR imaging; J.P.L., 20 years of experience with cardiac MR imaging) by using a circular 15-mm² region of interest and the software available on the MR imager. The percentage of change in signal intensity (\(S_{PC} \)) was normalized to the signal intensity of healthy myocardium (S_{HM}) by using the equation:\n
\[ \frac{S_{PC}}{S_{HM}} = \left( \frac{S_{AMNP} - S_{HM}}{S_{HM}} \right) \cdot 100. \]
Pathologic Analysis

After MR imaging, the hearts were embedded in paraffin. Sections with a thickness of 5 μm were sliced perpendicular to the long axis of the heart, from base to apex, and stained with eosin and with Perls Prussian blue stain (J.F.D.). Immunohistochemical analysis with staining for α-actin and macrophages was performed to determine the presence of smooth muscle cells. Identification of cells was confirmed by a pathologist (J.B.M.) with 10 years of experience in pathologic analysis of muscle.

Statistical Analysis

Quantitative results in study sample groups (with each group containing three to 15 data samples, unless otherwise stated) were expressed as the mean ± standard deviation. For cell growth capacity, linear regression analysis was performed for log(C) as a function of d, with C being the number of cells and d being the number of days of culture. Mean values for iron load per cell were fitted with the internalization model developed by Wilhelm et al (24). Correlations between the division factor of iron load and the cell amplification factor, as well as between 1/T2 or 1/T2* and the number of labeled cells or number of days in culture, were performed by using software for linear or first-order exponential decay regression analysis (Origin 6.0; Originlab, Northampton, Mass).

RESULTS

In Vitro Uptake and Intracellular Localization of Anionic Magnetic Nanoparticles

Perls Prussian blue staining of cultured smooth muscle cells after 2 hours of incubation at 37°C in the suspension of anionic magnetic nanoparticles (1.0 mmol/L iron) showed intracytoplasmic iron inclusions as dense blue-stained vesicles (Fig 1a). In contrast, no iron was detected in cells after 2 hours of incubation in a suspension of dextran-coated iron oxide nanoparticles with the same iron concentration (Fig 1b). Transmission electron microscopy of labeled cells (Fig 2) indicated the presence of the anionic magnetic nanoparticles exclusively in polydisperse vesicles in the cytoplasm. The average size of the magnetic vesicles increased with cell division over time (from 173 nm ± 72 at day 0 to 600 nm ± 119 at day 7), as did the iron content.

Cell Viability and Quantification of Magnetic Label

Smooth muscle cell viability in culture after labeling, as determined by absence of uptake of trypan blue stain, was greater than 94% ± 1. As shown in Figure 3, cell growth capacity was not affected by labeling with iron concentrations of 1.0 and 5.0 mmol/L, in the sense that the average daily rate of change in the cell count for labeled cells—with 1.0 mmol/L iron, log(C) = 5.85 + 0.12 · d; with 5.0 mmol/L iron, log(C) = 5.95 + 0.10 · d—was not significantly different from that in nonlabeled controls—log(C) = 5.87 + 0.13 · d (P > .05).

Cellular uptake of iron oxide nanoparticles was evaluated for different extracellular iron concentrations of 0.1–10.0 mmol/L (Fig 4). We found a maximum iron load of 12 pg per cell with use of a high extracellular iron concentration (10.0 mmol/L) during 2-hour incubation. Long-term studies were performed with an extracellular iron concentration of 1.0 mmol/L and 2 hours of incubation time. By means of magnetophoresis measurements, the distribution of iron load per cell was obtained for different durations of culture (Fig 5a). A shift toward lower values was observed in parallel with cell division. Intracellular mean iron load of 1.7 pg ± 0.12 per cell at day 0 of culture decreased continuously with cell division (Δln = 1.02 · Δt) and reached the detection threshold at day 14 (Fig 5b), as the electron spin resonance spectrum of anionic magnetic nanoparticle–labeled smooth muscle cells and background reached comparability. A strong correlation (r² = 0.99) between the cell amplification factor and the division factor of iron load was observed with electron spin resonance until day 7 of culture. After 7 days of culture, the magnetic load per cell did not decrease in parallel with cell growth.

Iron quantification in the culture medium indicated that from day 0 to day 14, iron release in the supernatant remained outside the threshold of electron spin resonance detection (4.5 x 10⁻¹¹ mol/L). Therefore, the iron concentration that resulted from release of nanoparticles into the culture medium was less than 10⁻¹² mol/L.

MR Imaging

In vitro MR imaging—T2 and T2* were measured by using spin-echo and gradient-echo sequences, with one echo for each sequence, while varying the TE. In all cases, monoexponential signal intensity decreases were found that allowed nonambiguous determination of T2 and T2* values.

T1 also was determined by using spin-echo sequences while varying the TR. T1 of anionic magnetic nanoparticle–labeled smooth muscle cells appeared to be iron concentration independent (data not shown).

Relaxation rates 1/T2 and 1/T2* were plotted (Fig 6a, 6b) as a function of global iron concentration for two groups of phantoms: the cell density group, with different numbers (ie, 10⁻³–10⁻⁶) of labeled cells at day 0; and the proliferating cell group, with a constant number of cells (ie, 10⁶) from day 0 to day 14 after label-
ing. As demonstrated in Figure 6a, 1/T2 was proportional to global iron concentration, with the same relaxivity ($r_2 = 95.8 \text{ mmol/L}\cdot\text{sec}^{-1}$) measured in both groups. 1/T2 thus was found to be directly proportional to the number of cells per milliliter in the agarose phantoms and to cell magnetic load ($r^2 = 0.97$), as expressed in the equation $1/T2 = 0.004 + [(1.74 \times 10^{-8}) \cdot C \cdot \text{mL}^{-1}]$. While 1/T2* remained almost constant as a function of the number of cells at day 0, it decreased linearly with iron concentration in the proliferating cell group (Fig 6b).

As shown in Figure 6c, for a constant number of cells (ie, $10^6$) from day 0 to day 14 in the phantoms, 1/T2 decreased linearly with the number of days of culture after magnetic labeling ($r^2 = 0.95$) until it reached the level in control samples at day 14, as expressed in the equation $1/T2 = 0.011 - [(5.66 \times 10^{-4}) \cdot d]$. By contrast, the decrease in 1/T2* that occurred with cell proliferation ($r^2 = 0.99$) is best described by the nonlinear model equation $1/T2^* = 0.019 + [0.07 \cdot \exp(-d/4.02)]$. Agarose phantoms with an increasing number of cells from day 0 to day 14, mimicking cell growth in a constant volume, demonstrated negligible changes in relaxation rate 1/T2 for all culture times (Fig 6d). Thus, if cells proliferate in a constant volume, 1/T2 remains constant because of iron conservation. For 1/T2*, a slight decrease in relaxivity was observed with cell proliferation in a constant volume for early days (days 0–3).

Fixation with glutaraldehyde had an effect on relaxivities: $r_2$ and $r_2^*$ in fixed cells were 96 and 520 sec$^{-1}$ mmol$^{-1}$ L$^{-1}$, respectively, whereas $r_2$ and $r_2^*$ in unfixed cells were 291 and 380 sec$^{-1}$ mmol$^{-1}$ L$^{-1}$, respectively.

In vivo model of ischemia.—Infarction of rat myocardium appeared as an area of high signal intensity in association with thinning of the lateral ventricular wall...
on T2-weighted spin-echo images and T2*-weighted gradient-echo images. At MR imaging after injection of labeled cells, lesions with hypointense signal were present in the apical region and left ventricular myocardium on T2-weighted images in both healthy hearts and hearts with ischemic injury (Fig 7). Lesions were observed on both spin-echo and gradient-echo images acquired at 2 and 48 hours after injection of labeled cells. Areas of signal intensity loss were better identified on T2*-weighted gradient-echo images than on T2-weighted spin-echo images. The percentages of change (decrease) in signal intensity on the spin-echo images for labeled cells were $-52.1\% \pm 0.8$ for healthy heart, $-63.7\% \pm 0.4$ for ischemic heart at 2 hours ($n = 2$), and $-54.6\% \pm 4.8$ for ischemic heart at 48 hours ($n = 2$). On long-axis T2-weighted gradient-echo images, sites of injected cells appeared as distinct areas of hypointense signal (Fig 7, B). Healthy excised hearts ($n = 2$) without labeled cells showed no areas of hypointense signal in the left myocardium (data not shown).

Histologic sections obtained at the same levels as MR images and double stained with Perls Prussian blue stain and eosin showed blue vesicles within cells, a finding that indicated the presence of anionic magnetic nanoparticles in cells at 2 and 48 hours after injection in healthy and ischemic myocardium (Fig 7, D–F).

Immunohistochemical analysis of sections obtained in adjacent sites was negative for macrophages and positive for α-actin, findings that confirmed that the labeled cells were those originally injected (Fig 7, G).

**DISCUSSION**

The most commonly used iron oxide particles (27) are neutral dextran-coated particles that require different transfection agents, such as poly-L-lysine, to facilitate cell internalization (28). Other approaches, such as the use of magnetodendrimers (29) or TAT peptide–modified nanoparticles (17), have also been used. In this article, we report the use of anionic magnetic nanoparticles and their uptake by smooth muscle cells. The in vitro MR signal intensity of proliferating cells was studied. We also validated the in vivo implantation of smooth muscle cells in myocardium with MR imaging.

The uptake of anionic magnetic nanoparticles was reported previously for different species and different cell types (19), but not for smooth muscle cells.

![Figure 5. Graphs show iron load distribution in smooth muscle cells after magnetic labeling (incubation at 37°C for 2 hours with iron concentration of 1.0 mmol/L) for days (D) 0–14 after magnetic labeling. (a) Stacked histogram shows iron load measured with magnetophoresis at days 0, 2, 3, 7, and 9. (b) Graph shows mean value of iron load per cell (black bars), measured with electron spin resonance, at days 0, 2, 3, 7, 9, and 14 after magnetic labeling. If the iron load per cell increased with the amplification factor, then iron mass remained constant until day 7 (iron mass conservation, white bars). Experiments were performed three times.](image)

![Figure 6. Graphs show 1/T2 and 1/T2* values measured in agarose phantoms containing anionic maghemite nanoparticle–labeled cells after incubation for 2 hours at 37°C with 1.0 mmol/L iron. (a, b) Relaxation rates are shown as a function of global iron concentration for two groups of phantoms: the cell density group, in which phantoms contained different numbers of labeled cells (from 10^3 to 10^6) measured at day 0 after labeling (iron load, 1.7 pg per cell), and the proliferating cell group, in which phantoms contained a constant number of cells (10^6) measured from day 0 (iron load, 1.7 pg per cell) to day 14 (iron load, 0.002 pg per cell) after labeling. (c, d) Graphs (top) and corresponding MR images (bottom) obtained with T2-weighted spin-echo sequence (2000/120) show relaxation rates as a function of days of culture after labeling for (c) a constant number of cells (10^6) and (d) an increasing number of cells. In c, 1/T2 is directly proportional to days of culture ($r^2 = 0.95$), whereas the decrease in 1/T2* with cell proliferation ($r^2 = 0.99$) is best described by the nonlinear model. In d, the number of cells increased proportionally with the amplification factor ($C = A1 \times 2.10^k$), and 1/T2 was observed to change, on average, at the negligible rate of $2.94 \times 10^{-5}$ msec ^{-1}/d. For 1/T2*, a slight decrease was observed with cell proliferation in a constant volume.](image)
Transmission electron microscopy of labeled smooth muscle cells showed that anionic magnetic nanoparticles were confined within intracytoplasmic vesicles. The number of these vesicles decreased with cell proliferation, while their size and particle load increased. This labeling agent, as demonstrated with trypan blue exclusion counting and kinetic analysis of cell growth, is nontoxic, and it did not affect cell growth capacity. In contrast to anionic magnetic nanoparticles, the uptake of dextran-coated iron oxide particles (1.0 mmol/L) in smooth muscle cells was not detectable with Perls Prussian blue staining after 2 hours of incubation. In fact, the dextran-coated labeling agent follows a low-efficiency fluid phase endocytosis pathway and requires longer incubation times (>24 hours) for substantial iron uptake (19).

The quantitative uptake of the anionic magnetic nanoparticles measured with magnetophoresis and electron spin resonance immediately after the labeling procedure, up to 12 pg per cell, was comparable with that in other cell lines previously studied (19). This level of uptake is 100-fold that obtained with monodomain dextran-coated superparamagnetic iron oxide nanoparticles (16) used without a transfection agent. By using a short incubation time and low extracellular iron concentration, we achieved an efficient iron load without affecting cell viability. With longer incubation time or higher extracellular iron concentration, one can obtain an iron load per cell that is comparable with that obtained with engineered iron oxide particles modified with TAT peptide or encapsulated in dendrimers (17,29).

During cell proliferation, the simple shift of the monomodal distribution of iron load per cell measured with magnetophoresis is consistent with sharing of the magnetic load between daughter cells. There was actually a direct correlation between cell amplification factor and division factor of iron mass per cell. This factor may vary as a function of iron load per cell that occurs during proliferation after labeling. More generally, 1/T2 depends linearly on the global iron concentration and does not vary according to differences in cell density or cell proliferation. Moreover, cells growing in a constant volume do not induce variations in MR signal intensity. Hence, longitudinal relaxation does not depend on changes in particle load and confinement that occur during cell proliferation. Thus, 1/T2 appears to be a robust parameter for in vivo cell quantification, because its value is directly proportional to the number of cells. If cells are proliferating in a constant volume, signal intensity should remain at a constant level for a long time.

1/T2* is one order of magnitude larger than 1/T2, but its dependence on cell numbers and cell proliferation is less linear. Slight differences in 1/T2* can be observed between agarose phantoms with different cell densities and different iron loads per cell. These differences in the relaxation rate are consistent with a nonconstant signal intensity for proliferating cells in a constant volume. Hence, T2* may vary as a function of iron load per cell, size and particle density of intracellular vesicles, and overall spatial distribution of labeled cells. This variation accords with the characteristics of the T2*-weighted sequence, among which magnetic susceptibility dominates. Thus, even if the initial 1/T2* of the tissue under consideration, without labeled cells,
is known, cell quantification based on $1/T^*$ measurement, although this technique is more sensitive than $1/T_2$ (30), must be performed with caution.

The transverse relaxivity ($r_2$) of 291 sec$^{-1}$·mmol$^{-1}$·L measured for unfixed anionic magnetic nanoparticle-labeled smooth muscle cells is comparable with the $r_2$ of 406 sec$^{-1}$·mmol$^{-1}$·L obtained for magnetodendrimers (29). Moreover, the $r_2$ value for glutaraldehyde-fixed anionic magnetic nanoparticle-labeled smooth muscle cells was 96 sec$^{-1}$·mmol$^{-1}$·L. This decrease was probably due to the reduced proton diffusion in the cell after the fixation process. On the other hand, $r_*^2$ was increased by the fixation process (from 380 to 520 sec$^{-1}$·mmol$^{-1}$·L for unfixed and fixed labeled smooth muscle cells, respectively). An explanation might be the presence of a more static local magnetic field within the fixed cells, with resultant increased field inhomogeneity.

In our in vivo experiments, we observed that efficient detection of magnetically labeled smooth muscle cells can be obtained with MR imaging by using these anionic nanoparticles. Transplanted magnetically labeled cells appeared on T2-weighted images as areas of hypointense signal both in healthy tissue and in ischemic myocardium. These results, which are in agreement with our histologic observations, suggest that MR imaging could be used to noninvasively assess cell transplantation and monitor cell engraftment in myocardium over time (31,32).

We detected anionic magnetic nanoparticle-labeled smooth muscle cells at MR imaging up to 48 hours after cell implantation. Kraitchman et al (31) observed hypointense signal areas in myocardium up to 3 weeks after cell engraftment by using magnetically labeled mesenchymal cells. Bulte et al (33) followed the migration of magnetically labeled rat neuronal progenitor cells in vivo up to 6 weeks in a rodent model of dysmyelination. One clinically relevant problem could be the uptake of magnetic particles by monocytes or macrophages if cell lysis occurs. This would lead to a failure in tracking the appropriate cells. Nevertheless, in our in vivo study, the colocalization of Perls Prussian blue stain and α-actin confirmed that the detected cells were the original injected cells. After 48 hours, labeled smooth muscle cells were still present in the area of injection. Longer times are needed to follow the fate of labeled smooth muscle cells, and experiments are currently under way with MR imaging for cell detection after 4 weeks.

Our study had several limitations. First, we did not achieve as high an iron load per cell as that (30 pg) obtained in another study with use of a low-generation heat-activated dendrimer as transfection agent (Superfect; Qiagen, Valencia, Calif) (28). The iron uptake that we documented, however, is generally sufficient for cell tracking with MR imaging.

Second, although we found a good correlation between $1/T_2$ and iron concentration in agarose gels in vitro, in vivo iron quantification in tissue by using MR imaging data was not performed, because we were able to determine the product of $r_2$ and iron concentration by using the equation, without knowing either the $r_2$ or the concentration of iron in labeled cells in the heart. In fact, for in vivo measurements, calibration of the MR signal intensity is required, with use of different numbers of labeled cells in the tissue of interest. To be accurate and of clinical use, this calibration should be performed ex vivo (to ensure a known number of labeled cells) and in fresh tissue.

In conclusion, anionic magnetic nanoparticles used for cell labeling demonstrated stable and long-lasting MR contrast properties, without cell toxicity, allowing detection at MR imaging of iron-loaded cells in myocardium. Quantitative in vitro analysis performed with MR imaging is an exciting first step toward the in vivo quantification of labeled cells. Future studies are required to follow cell engraftment over time and to fully correlate hypointense signal with the number of labeled cells. These results are promising for future in vivo study.

Although these particles are not yet clinically approved, their ease of use could be of interest for cell labeling and long-term in vivo cell tracking. The data presented here could be of interest for determining cell labeling conditions and MR imaging sequences.

Acknowledgments: The authors thank L. Louedec (INSERM U 460, Hôpital Bichat) for excellent technical assistance, S. Chillon (Hôpital Bichat) for ex vivo MR imaging experiments, P. Prud’homme and D. Glutron (Centre Inter-établissement de Résonance Magnétique [CIERM], Bicêtre) for in vitro MR imaging experiments, N. Carreau (Université de Technologie de Compiègne) for transmission electron microscopy preparation and analysis, P. Smirnov (Laboratoire de Recherche en Imagerie, Paris) for helpful discussions, and J. Bittoun (Ciem, Bicêtre) for experiments performed in his laboratory.

References


16. Josephson L, Tung CH, Moore A, Weissleder R. High-efficiency intracellular magnetic labeling with novel super-
In the preface, the editor-in-chief states that “this book will mitigate some of the difficulties we are all facing in learning musculoskeletal imaging. This book covers the essentials of musculoskeletal imaging. It is intended to be comprehensive, yet concise.”

_Essentials of Musculoskeletal Imaging_ is a multiauthored book (31 authors, mostly from the United States) that consists of 10 sections. The sections are partially disease oriented (eg, rheumatic diseases and neoplasms), and some sections focus on imaging modalities (eg, magnetic resonance imaging and ultrasonography) while others focus on procedures (eg, interventional radiology and diskography) or anatomy (eg, knee, spine, and hip). The sections are divided into 97 chapters of various lengths; the chapter on the knee, for example, is 42 pages, whereas the chapter on liposarcoma is less than one page.

The text follows a strict format; each chapter starts with an overview of the disease, followed by imaging findings and differential diagnosis. The chapters conclude with recommended readings and contain references as recent as 2001.

The book is lavishly illustrated (1900 illustrations). The quality of the images is sufficient, albeit some illustrations display distracting patient or imaging data. Some practical information could be added in future editions; for example, is the use of gadopentetate dimeglumine useful in differentiating liposarcoma from lipoma? A nice example of useful information is the chapter “Imaging of Total Hip Replacement,” which is written by an orthopedic surgeon. Likewise, the chapter titled “Imaging of Low Back Pain” provides useful information for daily practice.

This book fulfills its intended purpose and meets the needs of its audience. It is reasonably priced at $99.00. Although there are a large number of textbooks and atlases on musculoskeletal imaging, this book deserves a place in our libraries. I can recommend it to general radiologists and orthopedic surgeons.

Reviewed by Paul R. Algra, MD, PhD
Abductor Tendons and Muscles Assessed at MR Imaging after Total Hip Arthroplasty in Asymptomatic and Symptomatic Patients

**PURPOSE:** To prospectively evaluate magnetic resonance (MR) imaging findings of abductor tendons and muscles in asymptomatic and symptomatic patients after lateral transgluteal total hip arthroplasty (THA).

**MATERIALS AND METHODS:** The institutional review board approved the study, and all patients provided informed consent. Two musculoskeletal radiologists blinded to clinical information analyzed triplanar MR images of the greater trochanter obtained in 25 patients without and 39 patients with trochanteric pain and abductor weakness after THA. Tendon defects, diameter, signal intensity, and ossification; fatty atrophy; and bursal fluid collections were assessed. In 14 symptomatic patients, MR imaging and surgical findings were correlated. Differences in the frequencies of findings between the two groups were tested for significance by using $\chi^2$ analysis.

**RESULTS:** Tendon defects were uncommon in asymptomatic patients and significantly more frequent in symptomatic patients: Two asymptomatic versus 22 symptomatic patients had gluteus minimus defects ($P < .001$); four asymptomatic versus 24 symptomatic patients, lateral gluteus medius defects ($P < .001$); and no asymptomatic versus seven symptomatic patients, posterior gluteus medius defects ($P = .025$). In both patient groups, tendon signal intensity changes were frequent, with the exception of those in the posterior gluteus medius tendon, which demonstrated these changes more frequently in symptomatic patients (in 23 vs five asymptomatic patients, $P = .002$). Tendon diameter changes were frequent in both groups but significantly ($P = .001$ to $P = .009$) more frequent in symptomatic patients (all tendon parts). Fatty atrophy was evident in the anterior two-thirds of the gluteus minimus muscle in both groups, without significant differences. In the posterosuperior third of the gluteus minimus muscle, however, differences in fatty atrophy between the two groups were significant ($P = .026$). Fatty atrophy of the gluteus medius muscle was present in symptomatic patients only, with significant differences among all muscle parts. Bursal fluid collections were more frequent in symptomatic patients ($n = 24$) than in asymptomatic patients ($n = 8$, $P = .021$). The MR imaging–based diagnosis was confirmed in all 14 patients who underwent revision surgery.

**CONCLUSION:** Abductor tendon defects and fatty atrophy of the gluteus medius muscle and the posterior part of the gluteus minimus muscle are uncommon in asymptomatic patients after THA.

---

Primary total hip arthroplasty (THA) is the second most common joint replacement procedure performed in the United States after primary total knee replacement. In 2002, more than 200 000 THA procedures were performed (1). Complications and rates of
revisions associated with total hip replacement have declined significantly, despite an increasing number of patients at risk for these outcomes (2). Reasons for residual pain after total hip replacement include malalignment of the prosthesis, infection, joint instability, trochanteric bursitis, ectopic bone formation, and prosthesis loosening (3). The imaging performed at work-up is usually focused on hardware failure such as loosening or malalignment of the prosthesis. However, a soft-tissue abnormality such as tendon tear, muscle atrophy, or bursitis is often the underlying reason for trochanteric pain and limping, especially if a transgluteal approach has been used (3–7).

Magnetic resonance (MR) imaging is the modality of choice for imaging soft-tissue abnormalities of the musculoskeletal system. MR imaging was initially considered to be of limited value owing to the presence of artifacts caused by the metallic implants. Later, optimized sequences began to be used to image soft-tissue abnormalities in the presence of metallic implants (8–12). More recently, MR imaging has been shown to be a valuable diagnostic tool in patients who have undergone total hip replacement (10,13). However, these findings are based on the experiences in studies involving only a few symptomatic patients (13). To our knowledge, the normal MR imaging appearance of the greater trochanter after hip replacement in asymptomatic patients has not been described.

Thus, the purpose of our study was to prospectively evaluate MR imaging findings of the abductor tendons and muscles in asymptomatic and symptomatic patients after THA performed with a lateral transgluteal approach.

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Asymptomatic Patients (n = 25)</th>
<th>Symptomatic Patients (n = 39)</th>
<th>P Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absent</td>
<td>Present</td>
<td>Absent</td>
</tr>
<tr>
<td>A: Tendon Defect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gluteus minimus</td>
<td>23 (92)</td>
<td>2 (8)</td>
<td>17 (44)</td>
</tr>
<tr>
<td>Lateral gluteus medius</td>
<td>21 (84)</td>
<td>4 (16)</td>
<td>15 (38)</td>
</tr>
<tr>
<td>Posterior gluteus medius</td>
<td>25 (100)</td>
<td>0</td>
<td>32 (82)</td>
</tr>
<tr>
<td>B: Tendon Diameter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gluteus minimus</td>
<td>2 (8)</td>
<td>4 (16)</td>
<td>19 (76)</td>
</tr>
<tr>
<td>Lateral gluteus medius</td>
<td>2 (8)</td>
<td>1 (4)</td>
<td>22 (88)</td>
</tr>
<tr>
<td>Posterior gluteus medius</td>
<td>0</td>
<td>24 (96)</td>
<td>1 (4)</td>
</tr>
<tr>
<td>C: Tendon Signal Intensity‡</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gluteus minimus</td>
<td>9 (36)</td>
<td>16 (64)</td>
<td>9 (23)</td>
</tr>
<tr>
<td>Lateral gluteus medius</td>
<td>3 (12)</td>
<td>22 (88)</td>
<td>7 (18)</td>
</tr>
<tr>
<td>Posterior gluteus medius</td>
<td>20 (80)</td>
<td>5 (20)</td>
<td>16 (41)</td>
</tr>
<tr>
<td>D: Tendon Ossification</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gluteus minimus</td>
<td>18 (72)</td>
<td>7 (28)</td>
<td>20 (51)</td>
</tr>
<tr>
<td>Lateral gluteus medius</td>
<td>20 (80)</td>
<td>5 (20)</td>
<td>18 (46)</td>
</tr>
<tr>
<td>Posterior gluteus medius</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E: Bursal Fluid Collection and Fan Sign§</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bursal fluid collection</td>
<td>17 (68)</td>
<td>8 (32)</td>
<td>15 (38)</td>
</tr>
<tr>
<td>Fan sign</td>
<td>21 (84)</td>
<td>4 (16)</td>
<td>14 (36)</td>
</tr>
</tbody>
</table>

* Data are numbers of cases, with percentages in parentheses.
† P values calculated by using χ² test.
‡ Tendon signal intensity graded on T1-weighted MR images as normal (hypointense) or abnormal (increased signal intensity) compared with signal intensity of normal tendon.
§ Fan sign refers to the finding, on a sagittal T1-weighted MR image obtained through the most lateral part of the greater trochanter, of gluteus medius muscle tissue reaching the bone outline of the greater trochanter like a fan. The fan sign was considered to be positive in the presence of a muscle tissue defect and usually was present at the surgical site (Fig 1).
MATERIALS AND METHODS

The study protocol was approved by the institutional review board. Informed consent was obtained from all patients.

Asymptomatic Patients

Between November 2002 and May 2003, 25 consecutive asymptomatic patients (14 men aged 39–81 years [mean, 60.4 years], 11 women aged 33–88 years [mean, 60.2 years]) who were asymptomatic after THA were prospectively included in the study protocol. All of these patients had undergone surgery because of osteoarthritis of the hip. All asymptomatic patients were clinically examined by one of two experienced hip surgeons (C.D. or H.P.N., with 12 and 18 years experience, respectively) during the routinely scheduled 1-year follow-up examination. Only patients fulfilling the following criteria were included:

(a) It was 1 year since they had undergone standardized lateral transgluteal THA (14);
(b) they had no hip pain at rest, during physical activity, or with local palpation of the greater trochanter;
(d) they were capable of normal abduction force on the hip side on which surgery was performed compared with their abduction force capability on the contralateral side; and (c) they were not limping.

Symptomatic Patients

Thirty-nine consecutive symptomatic patients (19 men aged 30–80 years [mean, 62.7 years], 20 women aged 39–79 years [mean, 64.3 years]) were prospectively included in the study during a period of 28 months—January 2001 to May 2003. All 39 patients were referred by the same two surgeons (H.P.N., C.D.) from our outpatient hip clinic for MR imaging of the greater trochanter for the evaluation of trochanteric pain, limping, and/or abduction weakness after total hip replacement. Thirty-four (87%) of these 39 patients presented with trochanter pain, and 36 (92%) presented with abductor weakness. Total hip replacement had been performed because of osteoarthritis of the hip in all 39 patients.

MR Imaging

MR imaging was performed in the asymptomatic and symptomatic patient groups by using a 1.5-T system (Symphony; Siemens Medical Solutions, Erlangen, Germany) according to a standardized protocol. A flexible wraparound receive-only surface coil was used. Coronal T2-weighted fast spin-echo (3910/75 [repetition time msec/echo time msec], echo train length of nine, 4-mm section thickness, 220 × 100-mm field of view, 512 × 256 matrix), transverse short inversion time inversion-recovery (5550/34/150 [repetition time msec/echo time msec/inversion time msec], 4-mm section thickness, 180 × 100-mm field of view, 512 × 256 matrix), and transverse T1-weighted spin-echo (669/18, 3-mm section thickness, 140 × 80-mm field of view, 512 × 256 matrix) MR images were obtained. The frequency-encoding gradient was always parallel to the long axis of the prosthesis (cranio-caudal direction). Only postsurgical hips were imaged.

Image Analysis

All MR images were independently analyzed by two experienced musculoskeletal radiologists (C.W.A.P. and M.Z., with 6 and 10 years experience, respectively). The asymptomatic and symptomatic patient images were reviewed randomly. The readers were blinded to the clinical symptoms and as to whether the patients were in the asymptomatic or symptomatic group. After the independent readings, interobserver differences were resolved by consensus.

The gluteus minimus tendon (ie, anterior part of the abductor apparatus), the lateral part of the gluteus medius tendon, and the posterior part of the gluteus medius tendon were analyzed separately. Tendon tears or detachments were defined as...
hyperintense signals extending to both surfaces of the tendon on MR images obtained with fluid-sensitive (ie, T2-weighted and short inversion time inversion-recovery) sequences. Osseous detachment of the tendon was recorded. The tendon diameter was qualitatively rated as normal, thinned, or thickened with respect to the radiologist’s knowledge of normal tendon diameter. Tendon signal intensity was graded on the T1-weighted MR images as normal (hypointense) or abnormal (increased signal intensity) compared with the signal intensity of a normal tendon.

Bursal fluid collections were rated as absent or present at fluid-sensitive MR imaging. The extent of the bursal fluid collection was measured electronically in all three dimensions by using a picture archiving and communication system workstation. The volume \( V \) of the bursal fluid was calculated by using the formula for elliptical volume:

\[
V = 0.52 \times H \times W \times D,
\]

where \( H \) is the height, \( W \) is the width, and \( D \) is the depth. Ossifications within the abductor tendon were documented to be present or absent. An area of ossification within the abductor mechanism was diagnosed only when it was depicted (with all sequences) as having a hypointense cortical rim with a centrally located area of fat signal intensity representing bone marrow.

Fatty atrophy of the abductor muscles was assessed on the transverse T1-weighted MR images by using the following grading system: Grade 0 indicated that no intramuscular fat was present; grade 1, that some fat streaks were present; grade 2, that fat was evident, but there was less fat than muscle tissue; grade 3, that there were equal amounts of fat and muscle tissue; and grade 4, that there was more fat than muscle tissue. This grading system corresponds to a classification system that is commonly used to categorize the rotator cuff muscles (15). Fatty atrophy was assessed on transverse T1-weighted MR images at two different levels: at one-third and at two-thirds of the distance between the iliac crest and the tip of the greater trochanter. At each of these levels, the anterior, middle, and posterior portions of the gluteus medius and gluteus minimus muscles were evaluated separately.

On a sagittal T1-weighted MR image obtained through the most lateral part of the greater trochanter (+) of the hip in a symptomatic patient. Note the osseous detachment (curved arrows) of the gluteus minimus tendon (arrowheads). The short inversion time inversion-recovery image shows a high-signal-intensity area (straight white arrow) at the gap (black arrows) between the bone fragment and the greater trochanter.
Comparison of MR Imaging
Findings with Subsequent Surgical Revision

In 14 patients, revision of the abductor tendon was performed by the same two orthopedic surgeons (H.P.N., C.D.). The indication for surgery was based on clinical and MR imaging findings in all of these patients. In these cases, one of the radiologists (C.W.A.P.) and one orthopedic surgeon (H.P.N.) in consensus retrospectively compared the MR imaging analysis results with the surgical reports in terms of the integrity of the abductor tendons and which tendons were involved. MR imaging findings referred to the data described in the Image Analysis section.

Statistical Analyses

Differences in the frequencies of findings between the two patient groups were tested for significance by using \( \chi^2 \) analysis. The Mann-Whitney \( U \) test was used to assess significant differences in fat content among the abductor muscles in both groups.

\( k \) Statistics were calculated for interobserver agreement. According to the method of Landis and Koch (16), agreement was rated as follows: \( k \) Values of 0–0.20 indicated slight agreement; values of 0.21–0.40, fair agreement; values of 0.41–0.60, moderate agreement; values of 0.61–0.80, substantial agreement; and values of 0.81–0.99, excellent agreement. A \( k \) value of 1.00 indicated absolute agreement.

A computer software package (SPSS, version 10.0.7; SPSS, Chicago, Ill) was used to perform all statistical calculations. \( P < .05 \) was considered to indicate a significant difference.

RESULTS

Demographic Data

No significant differences in the numbers of male and female patients were observed between the symptomatic and asymptomatic patient groups: In the asymptomatic group, 14 hips of men and 11 hips of women were assessed. In the symptomatic group, 19 hips of men and 20 hips of women were assessed (\( P = .129, \chi^2 \) analysis). No significant differences in side were seen: In the asymptomatic group, 13 hips on the right side and 12 hips on the left side were assessed. In the symptomatic group, 21 hips on the right side and 18 hips on the left side were assessed (\( P = .129, \chi^2 \) analysis). Age distribution was not significantly different between the two groups: Mean ages for the asymptomatic and symptomatic groups were 60.3 and 63.5 years, respectively (\( P = .353, \) Mann-Whitney \( U \) test).

MR Imaging

The frequency of MR imaging findings of the abductor tendons in the asymptomatic and symptomatic patients are summarized in Table 1. Tendon defects were uncommon in asymptomatic patients and significantly more frequent in symptomatic patients: Gluteus minimus defects were identified in two (8%) of the 25 asymptomatic patients versus 22 (56%) of the 39 symptomatic patients (\( P < .001 \)). Lateral gluteus medius tendon defects were identified in 16% (four of 25) of asymptomatic patients versus 62% (24 of 39) of symptomatic patients (\( P < .001 \)) (Fig 2). Posterior gluteus medius tendon defects were identified in no asymptomatic patients versus 18% (seven of 39) of symptomatic patients (\( P = .025 \)). In one case, the gluteus minimus tendon showed an osseous detachment from the greater trochanter (Fig 3).

In both patient groups, signal intensity changes within tendons were frequent, with the exception of those in the posterior gluteus medius tendon (Fig 4), which were significantly more frequent in the symptomatic patients (in 59% [23 of 39] vs in 20% [five of 25] of asymptomatic patients, \( P = .002 \)). Changes in tendons (thickened and thinned tendons combined) were very frequent in both groups but significantly (\( P = .001 \) to \( P = .009 \)) more frequent in all tendon parts in the symptomatic patients. Bursal fluid collections (Fig 5) were frequent in both asymptomatic (32% [eight of 25]) and symptomatic (62% [24 of 39]) patients (\( P = .021 \)). The mean volume of fluid collection was 1.57 mL (range, 0.20–3.65 mL) in the asymptomatic patients and 10.91 mL (range, 0.04–64.06 mL) in the symptomatic patients.
Interobserver agreement for assessment of the various findings is summarized in Table 2. Interobserver agreement for assessment of the abductor tendon defects was moderate to excellent (κ, 0.577–0.862). Agreement was lower for assessment of tendon diameter (κ, 0.474–0.559) and signal intensity changes (κ, 0.341–0.581).

The distribution of fatty atrophy is illustrated in Figures 6 and 7. Fatty atrophy was evident in the anterior two-thirds of the gluteus minimus muscle in both the asymptomatic and the symptomatic patients, without significant differences (P = .106 to P = .770) (Fig 8). The difference in the degree of fatty atrophy in the posterosuperior third of the gluteus minimus muscle, however, was significant between the two groups (P = .026). Fatty atrophy of the gluteus medius muscle was almost exclusively present in the symptomatic patients (Fig 8). Differences were significant among all three parts of this muscle at both evaluated levels (P < .001 to P = .036) (Fig 6). The fan sign was significantly more frequently positive in the symptomatic patients (64% [25 of 39]) than in the asymptomatic patients (16% [four of 25], P < .001) (Fig 1).

Surgical Findings after Revision

Fourteen patients underwent revision hip surgery after MR imaging. In nine of these 14 patients, a gluteus minimus tendon tear was seen. In seven patients, a tear in the lateral part of the gluteus medius tendon was seen. No tear in the posterior part of the gluteus medius tendon was diagnosed. The MR imaging–based diagnosis of tendon tear was confirmed in all 14 patients at revision surgery.

DISCUSSION

The attachments of the abductor tendons around the greater trochanter of the hip can be divided into three parts: The main tendon of the gluteus medius muscle has a strong insertion covering the posterosuperior aspect of the greater trochanter. The lateral part of the gluteus medius tendon insertion is obliquely oriented. It extends from the posterior to the anterior end and inserts at the lateral aspect of the greater trochanter. Parts of the gluteus medius muscle extend anteriorly and cover the insertion of the gluteus minimus tendon. The lateral part of the gluteus medius tendon is usually thin, and it may be composed almost entirely of muscle. The main tendon of the gluteus minimus attaches to the anterior part of
the trochanter. A part of the gluteus minimus insertion is muscular and inserts into the ventral-superior capsule of the hip joint (17).

A wide spectrum of abnormalities affecting the soft tissues of the hip has been noted in patients with trochanteric pain. Besides the frequently encountered trochanteric bursitis, there may also be so-called rotator cuff tears of the hip (18,19). Bunker and co-workers (19) described the typical appearance of this tear as a circular or oval gluteus minimus tendon defect that extends posteriorly into the lateral part of the gluteus medius tendon. MR imaging has been reported to be useful for the diagnosis of either tendinosis or tendon tears of the abductors (20,21). So far, little is known about MR imaging changes in the abductor tendons after total hip replacement (13).

The three basic surgical approaches used most commonly for THA are the lateral, posterior, and transtrochanteric methods (22). Complications related to each of these approaches have been reported and include joint dislocation, heterotopic ossification, neurovascular damage, postoperative limp, implant malalignment, and trochanteric nonunion (with the transtrochanteric approach) (3,4,7). The lateral approach necessitates splitting the middle of the gluteus medius muscle and releasing its anterior tendinous portion from the greater trochanter and has been reported to enable ease of access into the hip joint, optimal joint visualization, protection of the neurovascular structures of the hip, and achievement of predictable results in terms of postoperative hip function (14,23,24).

However, in a study to assess the postoperative integrity of the gluteus medius tendon after 97 consecutive lateral THAs, metal markers were placed in the gluteal aponeurosis—one on each side of the surgical line (4). Radiographs were obtained immediately, 2 weeks, 2 months, and 1 year after surgery. A division between the markers developed in about half of the patients, but gross divisions were rare. Trendelenburg gait was significantly increased in only those patients who had divisions larger than 2.5 cm; this finding indicates that a moderate gluteal elongation may be readily compensated for (4). In our series, 8% of the asymptomatic patients had a defect in the gluteus minimus tendon and 16% of these patients had a defect in the gluteus medius tendon. None of these patients were limping at presentation. The frequency of tendon defects was significantly higher in the symptomatic patients: 56% had gluteus minimus tendon defects, and 62% had gluteus medius tendon defects.

Fatty atrophy of the gluteus minimus muscle seems to be very frequent after THA, even in asymptomatic patients. In contrast, fatty atrophy of the gluteus medius muscle was seen almost exclusively in the symptomatic patients. Fatty atrophy of the muscle is an important predictor of the success of a tendon reconstruction. It is known from experiences with shoulder joints that once fatty degeneration of the muscle tissue has developed owing to a tendon tear, fatty atrophy rarely resolves after tendon reconstruction (25). Muscle atrophy is an important differential diagnosis relative to simple tendon tear in patients with limping, and, therefore, careful attention should be paid to muscle atrophy.

Another cause of Trendelenburg gait or limping is damage to the superior gluteal nerve after lateral-approach hip surgery. In a prospective study involving 81 consecutive patients who underwent lateral-approach surgery, the abductor muscles of the hip were assessed electromyologically and clinically. In nine patients, complete denervation occurred. Persistent damage to the nerve was associated with a positive Trendelenburg test result (26).

The results of repeat surgery for repair of abductor tendon tears that occurred as a complication of primary THA seem to be promising. In a study (27) to analyze the results obtained for nine patients after they underwent revision of the abductor tendons, limping was markedly decreased in five of these patients. The need for ambulatory aids also was reduced in five of these patients. However, with regard to pain reduction, the procedure was less successful; this result suggests that the best indications for repair are symptoms of marked abductor weakness. It seems that substantial preoperative pain is less likely to be decreased (27).

Fluid collections around the greater trochanter were frequently found in both the asymptomatic and the symptomatic patients in this study. Fluid collections around other joints after surgery seem to be a common finding in asymptomatic patients also and should not be misinterpreted as a sign of bursitis (28). However, in our series, fluid collections larger than 4 ml were seen only in the symptomatic patients.

Traditionally, MR imaging has had a very limited role in the examination of patients after arthroplasty, primarily because of susceptibility artifacts related to the metallic implants. Imaging evaluation of the painful hip after arthroplasty typically has been limited to conventional radiography, arthrography, and bone scintigraphy; however, the findings of such examinations are often focused on the bones and the orthopedic hardware. More recently, ultrasonography has been shown to be promising for evaluating the soft tissue around the greater trochanter (29). Modifications of conventional MR imaging sequences can be used to reduce the artifacts generated by implants, including total knee arthroplasty (11), THA (10), and shoulder arthroplasty (12) prostheses. Optimized
image quality can be achieved with spin-echo MR imaging by using a high bandwidth (at least 130 Hz per pixel), a high-spatial-resolution matrix (512 × 512), sequences with multiple refocusing pulses, and a frequency-encoding axis parallel to the long axis of the prosthesis. The degree of distortion is reduced by using this optimized technique (30). Imaging of the greater trochanter and the abductors with acceptable image quality in the setting of total hip replacement was possible in all patients in the symptomatic and asymptomatic groups. We believe that MR imaging can yield important information for the differential diagnosis of trochanteric pain and limping in patients after hip arthroplasty (13).

There were limitations to our study that should be considered. No preoperative MR imaging of the hip was performed in our study; therefore, some of the findings may have been present before surgery. For example, fatty atrophy of the abductor muscles may have been present before surgery. All surgical revisions were based on clinical and MR imaging findings. Only those patients who received a positive diagnosis of abductor tendon tear at MR imaging underwent surgery; therefore, the number of false-negative cases could not be assessed.

It is important to recognize that although many MR imaging findings such as altered signal intensity and abductor tendon diameter, bursal fluid collections, and fatty atrophy of the anterior gluteus minimus muscle are more frequent in symptomatic patients, they are also frequently found in asymptomatic patients after lateral transgluteal THA. However, defects of the abductor tendons and fatty atrophy of the gluteus medius muscle and the posterior part of the gluteus minimus muscle are uncommon in asymptomatic patients after THA and therefore appear to be clinically relevant.

References

Anterior Tibial Tendon Abnormalities: MR Imaging Findings

In the orthopedic literature, only about 50 cases of complete ATT tears have been reported. The ATT is usually only exposed to minor mechanical stress owing to its straight course. Consequently, abnormalities are less common in this tendon than in other tendons. Nevertheless, hypoxic degenerative tendinosis or mucoid degeneration occurs and may lead to a partial or complete tear of the ATT. Most tears occur without a trauma. Patients often present with slight foot drop preceded by a long history of swelling and pain at the dorsomedial aspect of the midfoot. Discontinuity of the ATT and the longitudinal extent of signal intensity abnormalities were measured (Mann-Whitney U test). Signal intensity abnormalities of the ATT and irregularities of the underlying tarsal bones were analyzed in consensus by two blinded radiologists ($\chi^2$ test).

RESULTS: In the symptomatic group, three cases of tendinosis and 13 partial and 12 complete ATT tears were diagnosed. In 11 cases (one case of tendinosis and two cases of partial and eight cases of complete ATT tear), surgical correlation was available and the MR imaging diagnosis was confirmed. In the asymptomatic group, four cases of tendinosis of the ATT were seen. The ATT diameter was significantly thicker in symptomatic patients at 1 cm (5.1 vs 3.1 mm in control group, $P = .001$), 3 cm (5.8 vs 3.4 mm, $P < .001$), and 6 cm (5.4 vs 4.3 mm, $P = .006$) proximal to the distal point of insertion. Most ATT abnormalities (in 23 [82%] of 28 patients) were located within the first 3 cm proximal to the insertion. Signal intensity abnormalities were seen in the anterior portion of the ATT in two (7%) of the 28 symptomatic patients and in the posterior portion in 11 (39%); diffuse involvement was seen in 15 (54%). Bone spurs on the navicular surface (nine [32%] patients vs no [0%] control subjects, $P = .001$), a ridged shape of the medial surface of the medial cuneiform bone (13 [46%] vs one [4%], $P < .001$), and osteophyte formation at the first tarsometatarsal joint (eight [29%] vs two [7%], $P = .056$) were significantly more common in the symptomatic patient group.

CONCLUSION: Characteristic findings of ATT abnormalities include tendon thickening ($\geq 5$ mm) and diffuse or posterior signal intensity abnormalities of the tendon within 3 cm from the distal point of insertion.

© RSNA, 2005

The anterior tibial tendon (ATT) is usually only exposed to minor mechanical stress owing to its straight course (1). Consequently, abnormalities are less common in this tendon than in other tendons. Nevertheless, hypoxic degenerative tendinosis or mucoid degeneration occurs and may lead to a partial or complete tear of the ATT (2). Most tears occur without a trauma. Patients often present with slight foot drop preceded by a long history of swelling and pain at the dorsomedial aspect of the midfoot. Discontinuity of the ATT and, occasionally, a mass at the anterior part of the ankle are palpable (3). Whereas a complete tear usually can be easily recognized, the clinical diagnosis of tendinosis or partial tear can be challenging.

In the orthopedic literature, only about 50 cases of complete ATT tears have been
The patients ranged in age from 30 to 82 years (mean age, 63.2 years). The female patients ranged in age from 41 to 82 years (mean age, 64.8 years), and the male patients ranged in age from 30 to 82 years (mean age, 59.3 years). One patient had diabetes mellitus. No patient had systemic inflammatory disease. Only one patient had a history of trauma. The onset of symptoms was acute in four patients. Twenty-four patients had a slow onset of symptoms with a chronic course (mean course, 65 weeks; range, 10–300 weeks). Six patients with a chronic course reported acute pain exacerbation.

Physical examination revealed a palpatable ATT defect (n = 9), swelling along the ATT (n = 17), and/or reduced dorsiflexion (n = 15). Indications for MR imaging were as follows: (a) probable complete ATT tear (n = 8), (b) probable other ATT abnormality (n = 12), and (c) possible ATT abnormality (n = 8). For 11 patients who underwent open reconstruction of the ATT, surgical confirmation of MR imaging findings was available. Seventeen patients were treated conservatively. After conservative treatment, 14 patients reported a reduction in their symptoms, and two patients were free of symptoms. In one patient, a complete tear occurred 5 months after the initial presentation, and surgery was performed.

Control Group

The control group consisted of 28 age- and sex-matched asymptomatic volunteers. Criteria for inclusion were as follows: (a) no foot pain, (b) no trauma to the ankle or foot in the previous 2 years, (c) no previous visit to a physician because of foot complaints, (d) no prior foot surgery, and (e) no systemic inflammatory disease. The volunteers ranged in age from 33 to 83 years (mean age, 62.9 years). The 20 female volunteers ranged in age from 45 to 83 years (mean age, 64.3 years), and the eight male volunteers ranged in age from 33 to 83 years (mean age, 59.3 years).
MR Imaging Protocol

MR imaging was performed with a 1.0- or 1.5-T system (Expert or Symphony; Siemens Medical Solutions, Erlangen, Germany). Patients and volunteers were examined in the supine position, with one ankle placed in a neutral position in the extremity coil. T1-weighted spin-echo MR images were obtained in the coronal plane (repetition time msec/echo time msec, 450–722/14–20; section thickness, 3–4 mm; field of view, 16 cm) and in the oblique transverse plane (45° between the coronal and transverse planes) (435–722/14–20; section thickness, 3–4 mm; field of view, 16 cm). T2-weighted fast spin-echo MR images were obtained in the coronal plane (4000/91–96; section thickness, 3–4 mm; field of view, 16 cm) and in the transverse plane (4000/86–96; section thickness, 4 mm; field of view, 15–16 cm). A fast spin-echo short inversion time inversion-recovery sequence (repetition time msec/echo time msec/inversion time msec, 4000–4800/30/150; section thickness, 3–4 mm; field of view, 17–18 cm) was performed in the sagittal plane.

MR Image Analysis

Qualitative evaluation.—MR imaging results were analyzed in consensus by two experienced musculoskeletal radiologists (two staff radiologists with 12 [M.Z.] and 7 years [C.W.A.P.] of experience in musculoskeletal radiology) who were blinded to the diagnosis. The review of the MR imaging examination results for the patients and the volunteers was performed in a randomized fashion. The following qualitative criteria were evaluated: signal intensity abnormalities of the ATT substance (defined as elevated signal intensity compared with the low signal intensity of normal tendon) on T1- and T2-weighted images and the location of signal intensity abnormalities within the cross section of the ATT. In terms of their locations, the signal intensity abnormalities were recorded as being in the anterior or posterior portion of the ATT or as being diffuse. In addition, the presence of a longitudinal split of the ATT prior to its point of insertion was recorded.

The readers were then asked to render one of the following diagnoses for each ATT: (a) a normal ATT, (b) a normal ATT that had signal intensity changes within its substance owing to a magic-angle effect (which was defined as a signal intensity abnormality in the ATT where the tendon was oriented 50°–60° to the main magnetic field that was seen only on T1-weighted images in the absence of circumscribed thinning or thickening of the tendon [11,12]), (c) tendinosis (which was diagnosed when there was a signal intensity change that was predominantly visible on T1-weighted images—a change that was not limited to the magic-angle region and/or was associated with changes in the tendon diameter), (d) a partial tear (which was diagnosed when there were signal intensity abnormalities on both T1- and T2-weighted images, with or without changes in the tendon diameter), or (e) a complete tear (discontinuity of the ATT). As part of evaluating the magic-angle effect, the angle between the main magnetic field and the tendon at the location of the signal intensity changes was measured on sagittal images.

The images were also evaluated for the following associated findings: fluid within the ATT sheath (whether present or absent); edema-like bone marrow abnormalities at the distal point of insertion; dorsal osteophytes at the talonavicular, cuneonavicular, and medial tarsometatarsal joints; and bone spurs on the surface of the navicular bone adjacent to the ATT. In addition, the shape of the medial surface of the medial cuneiform bone was characterized as convex, smooth concave, or ridged.

Measurements.—Measurements were performed by a fellow in musculoskeletal radiology (B.M., with 2 years of experience in musculoskeletal radiology) who used a picture archiving and communication system workstation (Image Devices, Idstein, Germany) and were obtained to the nearest tenth of a millimeter and then rounded to the nearest millimeter (Fig 1).

The short-axis diameter of the normal, the tendinopathic, and the partially torn ATTs was measured at 1, 3, and 6 cm proximal to the point of insertion on oblique transverse T1-weighted MR images. These measurements were performed without knowledge of the diag-
nosis. In the tendinopathic, the partially torn, and the completely torn ATTs, the signal intensity abnormalities were measured in terms of their longitudinal extent and distance from the distal point of insertion. In addition, the maximal short-axis diameter of the portion of the tendon with signal intensity abnormalities (Fig 1), the distance between the point of this maximal short-axis diameter and the point of insertion, and the relationship of this point to the underlying tarsal bones were determined. In cases of complete tear, the distance between the point of insertion and the location of the tear, as well as the size of the gap between the tendon stumps, was measured. The ATT has multiple attachment sites; it attaches to the first metatarsal as well as to the medial cuneiform bone. For simplification, the joint space of the medial tarsometatarsal joint was chosen as the site of the distal attachment so that we could perform reproducible measurements.

Statistical Analysis

Surgical reports (available for 11 patients) or results of clinical follow-up examinations at which a diagnosis was confirmed by the clinician were the standard of reference for the diagnosis of an ATT abnormality. These reports and results were reviewed in consensus by an experienced musculoskeletal radiologist (M.Z.) and a fellow in musculoskeletal radiology (B.M.). In cases for which no surgical correlation was available, the criteria used by the clinician for the diagnosis of ATT abnormality were swelling and pain at the insertion point of the ATT, with or without reduced dorsiflexion force, and reduction in the severity of symptoms with conservative treatment. Qualitative criteria were compared by using the $\chi^2$ test. Continuous data were analyzed with the two-tailed Mann-Whitney $U$ test. Using results for the 16 symptomatic patients with tendinosis or partial tears and the 28 asymptomatic volunteers, we calculated the sensitivity and specificity of MR imaging in the diagnosis of abnormal ATTs for different cutoff values for the short-axis diameter of the tendon as measured within the first 3 cm from the distal point of insertion. Completely torn ATTs were not included in this calculation because in such situations the tendon diameter can no longer be measured. A $P$ value of less than .05 was considered to indicate a statistically significant difference.

RESULTS

Qualitative Evaluation

In the symptomatic patient group ($n = 28$), three (11%) patients had tendinosis (Fig 2, B), 13 (46%) had partial tears (Fig 2, C), and 12 (43%) had complete tears (Fig 3) of the ATT (Table 1). In 11 cases (39%) (one case of tendinosis and two cases of partial and eight cases of complete ATT tears), surgical correlation was available and the MR imaging diagnosis was confirmed. In one patient in whom a partial tear was seen on MR images (Fig 4a, 4b), acute pain exacerbation associated with increased foot drop occurred abruptly 5 months later. Surgery was then performed, and a complete tear was found (Fig 4c). In the control group ($n = 28$), MR images of three (11%) ATTs showed a magic-angle effect (the angles between the ATT and the main magnetic field were 52°, 53°, and 58°). In four (14%) volunteers in the control group, tendinosis was diagnosed. A longitudinal split close to the point of insertion was present in five ATTs in the control group and in one ATT in the symptomatic patient group (Fig 5). Signal intensity abnormalities were seen in the anterior portion of the ATT in two (7%) of the symptomatic patients and in the posterior (adjacent to bone) portion of the ATT in 11 (39%) such patients; diffuse involvement was seen in 15 (54%) patients.

<table>
<thead>
<tr>
<th>TABLE 1 Results of Qualitative Evaluation of MR Images of ATT in Two Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Diagnosis</td>
</tr>
<tr>
<td>Normal ATT</td>
</tr>
<tr>
<td>Normal ATT with magic-angle effect</td>
</tr>
<tr>
<td>Tendinosis</td>
</tr>
<tr>
<td>Partial tear</td>
</tr>
<tr>
<td>Complete tear</td>
</tr>
<tr>
<td>Longitudinal split of ATT close to point of insertion</td>
</tr>
<tr>
<td>Location of signal intensity abnormalities within abnormal tendon</td>
</tr>
<tr>
<td>Anterior half</td>
</tr>
<tr>
<td>Posterior half</td>
</tr>
<tr>
<td>Diffuse</td>
</tr>
<tr>
<td>Associated findings</td>
</tr>
<tr>
<td>Edema-like bone marrow abnormalities at point of insertion</td>
</tr>
<tr>
<td>Dorsal osteophytes</td>
</tr>
<tr>
<td>Cuneonavicular joint</td>
</tr>
<tr>
<td>Medial tarsometatarsal joint</td>
</tr>
<tr>
<td>Bone spur with contact to tendon on navicular surface</td>
</tr>
<tr>
<td>Shape of medial surface of medial cuneiform bone Convex</td>
</tr>
<tr>
<td>Smooth concave</td>
</tr>
<tr>
<td>Rigid</td>
</tr>
</tbody>
</table>

* Data are numbers of subjects, with percentages in parentheses.
† NC = not calculated. (A report of abnormalities of the ATT at MR imaging was one of three criteria for inclusion for the symptomatic group; hence, statistical comparison of this criterion was not performed.)
Fluid in the ATT sheath was a common finding in the symptomatic group, manifesting in 82% of patients but in only 7% of volunteers ($P < .001$); 12 of 12 ATTs with a complete tear (Fig 3), nine of 13 ATTs with a partial tear, and two of three ATTs with tendinosis had fluid within the tendon sheath. Edema-like bone marrow abnormalities at the distal point of insertion of the ATT were seen in nine (32%) of the 28 patients in the symptomatic group but in no subjects in the asymptomatic control group ($P < .001$). None of the three ATTs with tendinosis in the patient group had bone marrow abnormalities at the point of insertion.

Bone spurs with ATT contact on the surface of the navicular bone (32% of patients vs 0% of volunteers, $P = .001$) (Fig 6) and osteophytes at the medial tarsometatarsal joint (29% of patients vs 7% of volunteers, $P = .036$) were significantly more frequent in the symptomatic patients than in the control group. There was a significant difference between patients and volunteers in the shape of the medial surface of the medial cuneiform bone (Fig 7): More symptomatic patients had a medial cuneiform bone with a ridged shape (46% of patients vs 4% of volunteers, $P < .001$) than one with a convex shape, which was more common in asymptomatic subjects (21% of patients vs 61% of volunteers, $P = .003$).

Dorsal osteophytes of the cuneonavicular joint (Fig 6) were more frequent in the symptomatic patients, occurring in 29% of patients and 14% of volunteers, but this difference was not significant ($P = .19$).

**Measurements**

The short-axis diameter of the ATT was significantly larger in the symptomatic group than in the control group at 1 cm (5.1 mm vs 3.1 mm, $P < .001$), 3 cm (5.8 mm vs 3.4 mm, $P < .001$), and 6 cm (5.4 mm vs 4.3 mm, $P = .006$) proximal to the point of insertion (Table 2). A short-axis diameter of 5 mm or greater within the first 3 cm proximal to the point of insertion had a sensitivity of 94% and a specificity of 98% for the diagnosis of tendinosis or partial tear of the ATT. Signal intensity abnormalities of the ATT substance began 0–51 mm from the distal point of insertion (mean distance, 8.8 mm) and had a mean longitudinal extent of 42 mm (range, 6–102 mm) (Table 3). In cases of tendinosis ($n = 3$) or partial tear ($n = 13$), the maximal short tendon diameter was 4–11 mm (mean, 7.1 mm), and this point of maximal diameter was located 3–76 mm (mean, 23.8 mm) proximal to the point of insertion (Table 3). Complete tears ($n = 12$) were found within the distal 0–35 mm (mean, 8.8 mm) from the point of insertion, with a gap of 12–77 mm (mean, 48.8 mm) (Table 3).

ATT abnormalities were most commonly found within the distal 3 cm of the ATT (in 23 [82%] of the 28 symptomatic patients). Nine (32%) of the 28 ATT lesions in the symptomatic group were located at the level of the cuneonavicular joint, and 14 (50%) were located at the level of the medial cuneiform bone.

**DISCUSSION**

There are two different clinical manifestations of ATT tears. The more common manifestation is that when a patient between the ages of 60 and 70 years reports the development of chronic symptoms (3) that accompany a swelling at the dorsomedial aspect of the midfoot, possibly in combination with an acute exacerbation of pain (13). Acute ATT tears are uncommon, occur at any age, and are the result of massive trauma associated with fractures or soft-tissue lacerations. According to the literature, ATT tears in...
general are more commonly seen in men than in women (7,14). Although the age distribution of our study population was in accordance with that in other studies, we observed a female predominance of 71%. A spontaneous tear of the ATT occurs mainly in the presence of hypoxic or mucoid degeneration (2). In cases of tendinosis or partial tear, the patient experiences pain and swelling along the distal part of the ATT that become worse during dorsiflexion of the foot. Although a complete tear of the ATT is usually diagnosed clinically, compensation by the extensor hallucis longus and extensor digitorum longus muscles for the loss of function may confuse the issue, and the diagnosis of an ATT tear may be missed (15). Late diagnosis may hamper primary surgical repair (13).

The anterior tibial muscle originates from the lateral tibial condyle and from the proximal half to proximal two-thirds of the tibia. The musculotendinous junction is located at the middle third of the tibia. The superior extensor retinaculum above the ankle retains the ATT. The tendon crosses the anteromedial aspect of the ankle and runs toward the medial border of the foot. It inserts at a tubercle on the anteromedial aspect of the medial cuneiform bone and on the inferomedial aspect of the base of the first metatarsal bone. The tendon is retained close to the talar head by the superomedial band and close to the medial cuneiform bone by the inferomedial band of the inferior extensor retinaculum, as well as by the transverse retinaculum band (16). In 84%–96% of individuals, the tendon inserts on both the medial cuneiform bone and first metatarsal bone, with a division just prior to insertion (16–18). In our series, a longitudinal split of the distal part of the ATT was visible in five asymptomatic volunteers in the control group and in one patient who had a partial tear. Although Khoury et al (8) reported a case of longitudinal tear in the midportion of the ATT with surgical correlation, a longitudinal split of the distal portion of the ATT may represent a normal variant.

At the ankle joint, the cross section of the ATT is round or oval and becomes flat distally. Our study results revealed that, within 3 cm of the distal point of insertion, an ATT thickness of 5 mm or less should be considered normal. This threshold value had a sensitivity of 94% and a specificity of 98% for the differentiation of normal ATTs from those with tendinosis or partial tears. In cases in which there is a normal diameter and signal intensity abnormalities are seen only on T1-weighted images, diagnosis of a clinically relevant ATT lesion should be made with caution. Signal intensity abnormalities may be caused by a magic-angle effect, as seen in three ATTs in the control group. Also, four volunteers in the control group had MR imaging signs of tendinosis but no clinical symptoms.

According to the orthopedic literature, ATT tears are typically located within 5–30 mm from the point of insertion (19–22). This is consistent with the data in our population, in which all ATT tears occurred within 35 mm of the distal point of insertion. In addition, signal intensity abnormalities associated with tendinosis and partial tears began a mean of 8.8 mm from the insertion point, and maximal tendon thickening was ob-

**TABLE 2**

Results of Measurement of Short-Axis Diameter of ATTs

<table>
<thead>
<tr>
<th>Measurement Location*</th>
<th>Symptomatic Group (n = 28)</th>
<th>Asymptomatic Control Group (n = 28)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of Subjects</td>
<td>Mean Value (mm)</td>
</tr>
<tr>
<td>1 cm from distal point of insertion</td>
<td>16</td>
<td>5.1</td>
</tr>
<tr>
<td>3 cm from distal point of insertion</td>
<td>16</td>
<td>5.8</td>
</tr>
<tr>
<td>6 cm from distal point of insertion</td>
<td>16</td>
<td>5.4</td>
</tr>
</tbody>
</table>

Note.—In terms of value in the diagnosis of tendinosis or partial tear of the ATT, the sensitivity of the finding of a short-axis diameter of 5 mm or greater within the first 3 cm from the distal point of insertion was 94% (15 of 16 symptomatic patients); the specificity of this finding was 98% (27 of 28 symptomatic patients).

* For an illustration of the method used to obtain these measurements, see Figure 1, D.
† These measurements were not performed in subjects with complete ATT tears (n = 12).
served a mean of 23.8 mm from the insertion point. In terms of their relationship to the tarsal bones, 82% of all ATT abnormalities were found at the level of the cuneonavicular joint and the medial cuneiform bone. At this location, the ATT is slightly bent over the tarsal bones and retained by the supero- and infero-medial bands of the inferior extensor retinaculum and the transverse retinaculum band.

The etiology of ATT tears is still being debated. Although some investigators have reported homogeneous vascularity throughout the ATT (21), others consider an anterior area of hypovascularity in the region of the retinacula to be a possible cause of ATT disorders. Petersen et al (23) have reported the presence of an avascular zone in the anterior half of the ATT that begins 5–16 mm from the insertion point and extends 45–67 mm. Our data do not support this theory: The signal intensity abnormalities in our study were most commonly found either in the posterior portion (39%) of the tendon—close to the bone—or in a diffuse distribution (54%); they were found in the anterior portion of the ATT in only 7% of cases. The significant association of bone morphologic features ($P < .001$ for a bone spur on the navicular surface, $P < .001$ for a ridged shape of the medial cuneiform bone, and $P = .036$ for osteophytes at the medial tarsometatarsal joint) with ATT abnormalities indicates that a mechanical irritation of the ATT by bone structures may be pathogenetically relevant.

Markarian et al (13) recommend surgical reconstruction of ATT tears in younger active patients, whereas nonsurgical forms of management like bracing or ankle-foot orthosis are appropriate treatments in less active elderly patients. In cases of partial tear, surgical management must also be considered.

The fact that surgical correlation was available for only 11 of our 28 patients represents an important study limitation. However, this ratio between surgically and conservatively treated patients corresponds to clinical experiences with ATT tears. Our use of reported MR imaging abnormalities of the ATT as an inclusion criterion introduced a selection bias because subtle ATT abnormalities may therefore have been underrepresented in our investigation. Owing to the retrospective design of this study, there may have been an influence of the MR imaging interpretations on the clinical diagnosis; this can be another source of bias. However, initial clinical findings and follow-up examination results that revealed

### TABLE 3

<table>
<thead>
<tr>
<th>Measurement*</th>
<th>No. of Patients</th>
<th>Mean Value (mm)</th>
<th>Standard Deviation</th>
<th>Range (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance between beginning of signal intensity abnormalities and distal point of insertion</td>
<td>28</td>
<td>8.8</td>
<td>12.0</td>
<td>0–51</td>
</tr>
<tr>
<td>Longitudinal extent of signal intensity abnormalities†</td>
<td>16</td>
<td>42.3</td>
<td>25.5</td>
<td>6–102</td>
</tr>
<tr>
<td>Maximal short-axis diameter of portion of tendon with signal intensity abnormality†</td>
<td>16</td>
<td>7.1</td>
<td>1.8</td>
<td>4–11</td>
</tr>
<tr>
<td>Distance between point of maximal short-axis diameter and point of insertion†</td>
<td>16</td>
<td>23.8</td>
<td>17.6</td>
<td>3–76</td>
</tr>
<tr>
<td>Distance between complete tear and distal point of insertion†</td>
<td>12</td>
<td>8.8</td>
<td>10.7</td>
<td>0–35</td>
</tr>
<tr>
<td>Length of gap between tendon stumps in cases of complete tear†</td>
<td>12</td>
<td>48.8</td>
<td>15.7</td>
<td>12–77</td>
</tr>
</tbody>
</table>

* For an illustration of the methods used to obtain these measurements, see Figure 1, B and C.
† Not measured in cases of complete tear ($n = 12$).
In conclusion, characteristic findings of ATT abnormalities include tendon thickening (≥ 5 mm) and diffuse or posterior signal intensity abnormalities of the tendon within 3 cm from the distal point of insertion. The significant association between irregularities of the underlying tarsal bones and ATT lesions may indicate that mechanical irritation is a pathogenetic factor in ATT lesions.

References
**Neuroradiology**

Fumiaki Yamashita, MD, PhD
Kazuo Kurisu, MD, PhD
Kenichi Satoh, PhD
Kazunori Arita, MD, PhD
Kazuhiko Sugiyama, MD, PhD
Megu Ohtaki, PhD
Junko Takaba, RT
Atushi Tominaga, MD, PhD
Ryosuke Hanaya, MD, PhD
Hiroyuki Yoshioka, MD, PhD
Seiji Hama, MD, PhD
Yoko Ito, MD
Yoshinori Kajiwara, MD, PhD
Kaita Yahara, MD
Taiichi Saito, MD
Muhamad A. Thohar, MD

Published online before print 10.1148/radiol.2353031338
Radiology 2005; 235:985–991

Abbreviations:
ADC = apparent diffusion coefficient
DNT = dysembryoplastic neuroepithelial tumor
DW = diffusion weighted
PNET = primitive neuroectodermal tumor
WHO = World Health Organization

**Tumors at MR Imaging**

1 From Depts of Neurosurgery (F.Y., K.K., K.A., K.Sugiyama, A.T., R.H., S.H., Y.K., K.Y., T.S., M.A.T.) and Radiology (J.T.), Graduate School of Biomedical Sciences, Hiroshima Univ, 1–2–3 Kasumi, Minami-ku, Hiroshima 734-8551, Japan; Dept of Environmetrics and Biometrics, Research Inst for Radiation Biology and Medicine, Hiroshima Univ, Japan (K. Satoh, M.O.); and Dept of Neurosurgery, National Hosp, Kure Med Ctr, Kure, Japan (H.Y., Y.I.). Received Aug 21, 2003; revision requested Oct 31; final revision received Aug 19, 2004; accepted Sep 8. Address correspondence to K.K. (e-mail: kuku422@hiroshima-u.ac.jp).

Authors stated no financial relationship to disclose.

**Purpose:** To determine if apparent diffusion coefficient (ADC) can be used to differentiate brain tumors at magnetic resonance (MR) imaging.

**Materials and Methods:** Institutional review board approval or informed patient consent was not required. MR images were reviewed retrospectively in 275 patients with brain tumors: 147 males and 128 females 1–81 years old, treated between September 1997 and July 2003. Regions of interest were placed manually in tumor regions on MR images, and ADC was calculated with a five-point regression method at b values of 0, 250, 500, 750, and 1000 sec/mm². ADC values were average values in tumor. All brain tumor subgroups were analyzed. Logistic discriminant analysis was performed by using ADC, age, and patient sex as independent variables to discriminate among tumor groups.

**Results:** A significant negative correlation existed between ADC and astrocytic tumors of World Health Organization grades 2–4 (grade 2 vs grades 3 and 4, accuracy of 91.3% [P < .01]; grade 3 vs 4, accuracy of 82.4% [P < .01]). ADC of dysembryoplastic neuroepithelial tumors (DNTs) was higher than that of astrocytic grade 2 tumors (accuracy, 100%) and other glioneuronal tumors. ADC of malignant lymphomas was lower than that of glioblastomas and metastatic tumors (accuracy, 83.6%; P < .01). ADC of primitive neuroectodermal tumors (PNETs) was lower than that of ependymomas (accuracy, 100%). ADC of meningiomas was lower than that of schwannomas (accuracy, 92.4%; P < .01). ADC of craniopharyngiomas was higher than that of pituitary adenomas (accuracy, 85.2%; P < .05). ADC of epidermoid tumors was lower than that of chordomas (accuracy, 100%). In meningiomas, ADC was not indicative of malignancy grade or histologic subtype.

**Conclusion:** ADC is useful for differentiation of some human brain tumors, particularly DNT, malignant lymphomas versus glioblastomas and metastatic tumors, and ependymomas versus PNETs.

© RSNA, 2005

Diffusion-weighted (DW) magnetic resonance (MR) imaging, currently the only MR imaging technique that provides information on water diffusion, involves the use of phase-defocusing and phase-refocusing gradients to allow evaluation of the rate of microscopic water diffusion within tissues. DW MR imaging has been used to study brain tumors and response to treatment (1), and its diagnostic potential and usefulness for obtaining the apparent diffusion coefficient (ADC) have been reported (2–4). There appears to be a correlation between the ADC on the one hand and tumor cellularity and tumor grade on the other (5–15). At present, however, no large overview studies regarding ADC of brain tumors are available, and to our knowledge, the ADCs of dysembryoplastic neuroepithelial tumors (DNTs) and other rare tumors remain unknown. In addition, the relationship between the ADCs of gliomas and World Health Organization (WHO) grade (9,14,16) and between ADCs of meningiomas and their histologic subtype await illumination (3,14). Thus, the purpose of our study was to determine if ADC values could be used to differentiate brain tumors.
MATERIALS AND METHODS

Patients

We retrospectively studied MR images of pathologically proved brain tumors. The study population consisted of 275 patients (275 brain tumors) ranging in age from 1 to 81 years (mean age, 45.9 years ± 21.8 [standard deviation]; median age, 50 years) treated between September 1997 and July 2003. There were 147 males ranging in age from 1 to 79 years (mean age, 44.9 years ± 21.8; median age, 48 years) and 128 females ranging in age from 1 to 81 years (mean age, 47.0 years ± 21.3; median age, 50.5 years). Although at Hiroshima University Hospital, institutional review board approval or informed patient consent is not required for retrospective review of MR images and patient records, we obtained prior informed consent from all patients or from members of their families before entering them into this study. In addition, to protect patient privacy, we removed all identifiers from our records at the completion of our analyses.

MR Imaging and Image Processing

All MR studies were performed with a 1.5-T superconducting system (Signa Horizon; GE Medical Systems, Milwaukee, Wis) and a circularly polarized head coil. All patients underwent MR imaging, which included, at the minimum, unenhanced and contrast material–enhanced transverse T1-weighted images, unenhanced transverse T2-weighted images, unenhanced transverse fluid-attenuated inversion-recovery (FLAIR) images, and unenhanced transverse DW images. We (J.T. and F.Y., with 15 and 10 years of experience with brain MR imaging, respectively) performed these imaging studies by using the same section orientations for all examinations; selections were made by means of consensus.

The transverse T1-weighted spin-echo MR sequence was performed with the following parameters: 3500/100; field of view, 22 × 16 cm; matrix size, 256 × 192; echo train length, 12; section thickness, 5 mm; repetition time msec/ echo time msec, 400/8; field of view, 22 × 16 cm; matrix size, 256 (frequency) × 192 (phase); section thickness, 5 mm; section gap, 2.5 mm; and two signals acquired. The contrast-enhanced T1-weighted sequences were performed after the administration of a gadolinium compound (0.1 mmol gadodiamide per kilogram of body weight). The transverse fast spin-echo T2-weighted sequence was performed with the following parameters: 3500/100; field of view, 22 × 16 cm; matrix size, 256 × 192; echo train length, 12; section thickness, 5 mm; section gap, 2.5 mm; and two signals acquired. Transverse FLAIR images were obtained by using fast and interleaved multisection sequences with the following parameters: 10 000/150; inversion time, 2200 msec; field of view, 22 × 22 cm; matrix size, 256 × 192; echo train length, 16; section thickness, 5 mm; section gap, 2.5 mm; and one signal acquired.

Transverse DW imaging was performed by using a single-shot T2-weighted echo-planar spin-echo sequence with the following parameters: 1600/107; diffusion gradient encoding in three (x, y, z) orthogonal directions; b values of 250, 500, 750, and 1000 sec/mm²; field of view, 24 × 24 cm; matrix size, 128 × 128; section thickness, 7.5 mm; section gap, 0 mm; and one signal acquired. At each b value, x, y, and z single-direction DW images and a baseline image (b = 0 sec/mm²) were acquired; combined [(x + y + z)/3] DW imaging was calculated and performed automatically by the MR instrument. We obtained 10 sections with 50 images at each b value in 13 seconds (10 images of combined [(x + y + z)/3] DW imaging, 10 images of the baseline image, and 10 images each of the x-, y-, and z-direction DW images). Therefore, each DW imaging study yielded a total of 200 images.

All DW imaging data were transferred to a computer workstation (SUN Sparc 20; Sun Microsystems, Mountain View, Calif) for determination of the signal intensity and ADC. MR Vision (version 1.5.5; L.A. Systems, Oyama, Japan) was used for the creation of the ADC maps that were generated and produced automatically by the MR instrument. We calculated the mean and standard deviation of the ADC value and imaging information. The ADC of individual tumors was determined using a commercially available software package (Statview, version 5.0; SAS Institute, Cary, NC). For each brain tumor classified according to the WHO system in our series, we calculated the mean and standard deviation of the ADC and patient age. Logistic discriminant (regression) analysis was performed by using ADC, age, and patient sex as independent variables, to discriminate among the tumor groups. With two groups, the method assumes that the misclassification rate is zero, maximum log likelihood estimators and the significance of the estimated coefficients are not available (17). The histologic results for two compared tumor groups were defined as both areas of hyperintensity on T2-weighted MR images and areas of hypointensity on FLAIR MR images. Necrotic components were differentiated on contrast-enhanced T1-weighted images as the interior of enhanced lesions. Hemorrhagic lesions were differentiated on unenhanced T1-weighted MR images as areas of hyperintensity. We (F.Y., J.T., and Y.K., with 10, 15, and 9 years of brain MR imaging experience, respectively) compared the ADC maps and other MR images carefully and placed the regions of interest only in the solid tumor components by means of consensus. We excluded cystic, necrotic, and hemorrhagic tumor areas. We chose regions of interest as central as possible within the tumor area at random and averaged the ADC of each tumor.


RESULTS

The histologic tumor types, male-female distribution, patient age, and tumor ADCs are presented in Table 1. The histologic subtype of meningiomas and their distribution are presented in Table 2. The results of logistic discriminant analysis are presented in Table 3.

Neuroepithelial Tumors

Astrocytic tumors, oligodendrogial tumors, and ependymal tumors.—Among astrocytic tumors in our series, the ADC of diffuse astrocytomas (WHO grade 2) was significantly higher than that of anaplastic astrocytomas (WHO grade 3) and glioblastomas (P < .01). The accuracy of logistic discriminant analysis was more than 90%. The ADC of anaplastic astrocytomas was higher than that of glioblastomas (P < .01). The higher the astrocytic tumor WHO grade, the lower the ADC.

Our results show that we could discriminate without misclassification between grade 2 oligodendrogial tumors and grade 3 glioblastomas. The ADC of grade 2 astrocytic tumors was higher than that of glioblastomas, and between grade 2 astrocytic tumors and ependymal tumors, and between grade 3 astrocytic tumors and ependymal tumors, although the number of these cases in our series was not large. On the other hand, we could not discriminate between astrocytic tumors and oligodendrogial tumors.

DNTs.—The ADC of DNTs is higher than that of diffuse astrocytomas, and the ADC of DNTs is higher than that of other grade 1 glioneuronal tumors. In addition, the ADC of DNTs is higher than that of any other neuroepithelial tumor, and there is no overlapping of ADC values.

Ependymoma versus PNET, Ventricle Tumors

We found that the ADC of ependymomas is higher than that of PNETs, and there is no overlapping. The ADC of ependymomas is consistently higher than 1.00 × 10⁻³ mm²/sec. The ADC of PNETs is consistently lower than 1.00 × 10⁻³ mm²/sec. The ADC of central neurocytomas is as low as that of glioblastomas and lower than that of subependymomas.

Glioblastomas, Metastatic Tumors, and Malignant Lymphomas

Malignant lymphomas manifest lower ADCs than those in glioblastomas and metastatic tumors (P < .01). In our study, however, glioblastomas and metastatic tumors could not be discriminated.

Meningiomas and Schwannomas

The ADC does not identify the histologically benign subtypes of meningioma. There was no statistical difference between the benign (WHO grade 1) and malignant (WHO grades 2 and 3) subtypes. In our series, there was only one microcystic meningioma. Its very high ADC may be characteristic of these tumors.

The ADC of schwannomas is significantly higher than that of meningiomas (P < .01).

Tumors Developing in Parasellar Lesions

Pituitary adenomas, craniopharyngiomas, germ cell tumors, and parasellar meningiomas that frequently develop in parasellar regions are sometimes difficult to differentiate from each other. We included only large adenomas with supratentorial extension because the microadenomas and small adenomas in our series were not sufficiently large to assign a region of interest. The ADC of pituitary adenomas is not significantly different from that of meningiomas and germ cell tumors. The ADC of craniopharyngiomas is significantly higher than that of pituitary adenomas (P < .05) and meningiomas (P < .05). We could discriminate without misclassification between craniopharyngiomas and germ cell tumors, and while the ADC distribution of craniopharyngiomas was higher than that of germ cell tumors, our series contained only a few cases. The ADC of germ cell tumors is not significantly different from that of meningiomas.

Tumors Developing Pineal Lesions and Other Analyses

The accuracy of logistic discriminant analysis is more than 90% between PNET and germ cell tumors, between PNET and glioblastomas, and between PNET and meningiomas. While the ADC distribution of PNET is lower than that in other tumors, a statistical difference was found only between PNET and glioblastomas (P < .05). We can discriminate without misclassification between PNET and malignant lymphomas; however, there is no difference in the ADC distribution between PNET and malignant lymphomas.

The ADC of epidermoid tumors is lower than that of chordomas; however, only a small number of cases were available for analysis.

DISCUSSION

According to our results, the higher the tumor WHO grade, the lower the ADC in glioneuronal tumors, especially astrocytic tumors. Our observations suggest that the ADC may be useful for predicting the degree of malignancy of astrocytic tumors. To our knowledge, the ADCs of astrocytic tumors, oligodendrogial tumors, and ependymal tumors have not been compared to date. We rec-
## TABLE 1
Summary of Histologic Diagnoses and ADC

<table>
<thead>
<tr>
<th>Tumor Type</th>
<th>No. of Patients</th>
<th>Mean Age (y) ± Standard Deviation</th>
<th>ADC Range (×10⁻³ mm²/sec) ± Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Neuroepithelial tumors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WHO grade 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pilocytic astrocytoma</td>
<td>6 (3/3)</td>
<td>18.5 ± 12.8 (16.3/20.7)</td>
<td>1.302–1.921 (1.302–1.841/1.425–1.921)</td>
</tr>
<tr>
<td>WHO grade 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diffuse astrocytoma</td>
<td>17 (13/4)</td>
<td>34.6 ± 15.6 (34.7/34.5)</td>
<td>1.270–1.776 (1.270–1.746/1.528–1.776)</td>
</tr>
<tr>
<td>WHO grade 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaplastic astrocytoma</td>
<td>16 (11/5)</td>
<td>32.9 ± 15.8 (33.4/31.8)</td>
<td>1.045–1.576 (1.045–1.466/1.068–1.576)</td>
</tr>
<tr>
<td>WHO grade 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glioblastoma</td>
<td>36 (24/12)</td>
<td>51.7 ± 20.1 (55.5/43.2)</td>
<td>0.769–1.422 (0.769–1.389/0.962–1.422)</td>
</tr>
<tr>
<td>McGillstroma (medulloblastoma, PNET)</td>
<td>9 (5/4)</td>
<td>16.8 ± 17.8 (17.4/16.0)</td>
<td>0.676–0.994 (0.676–0.994/0.818–0.962)</td>
</tr>
<tr>
<td><strong>Other tumors</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chordoma</td>
<td>2 (2/0)</td>
<td>58.5 ± 5.0</td>
<td>1.425–1.501</td>
</tr>
<tr>
<td>Cranioopharyngioma</td>
<td>8 (4/4)</td>
<td>41.3 ± 22.0 (40.4/38.5)</td>
<td>1.182–1.801 (1.182–1.800/1.182–1.801)</td>
</tr>
<tr>
<td>Epidermoid</td>
<td>9 (3/6)</td>
<td>59.2 ± 11.6 (51.7/63.0)</td>
<td>0.840–1.395 (0.840–1.372/0.840–1.395)</td>
</tr>
<tr>
<td>Germ cell tumor</td>
<td>11 (8/3)</td>
<td>19.8 ± 9.3 (19.3/21.3)</td>
<td>0.953–1.504 (0.953–1.504/0.953–1.503)</td>
</tr>
<tr>
<td>Hemangioblastoma</td>
<td>2 (2/0)</td>
<td>61.0 ± 21.2</td>
<td>1.379–1.532</td>
</tr>
<tr>
<td>Hemanangiopericytoma</td>
<td>1 (1/0)</td>
<td>56</td>
<td>1.305</td>
</tr>
<tr>
<td>Histiocytoma</td>
<td>1 (0/1)</td>
<td>2</td>
<td>0.627</td>
</tr>
<tr>
<td>Hypothalamic hamartoma</td>
<td>2 (1/1)</td>
<td>24.0 ± 25.5 (6/42)</td>
<td>1.540–1.676 (1.540–1.676)</td>
</tr>
<tr>
<td>Malignant lymphoma</td>
<td>8 (4/4)</td>
<td>58.1 ± 13.0 (59.5/56.8)</td>
<td>0.504–0.1067 (0.504–0.908/0.504–1.067)</td>
</tr>
<tr>
<td>Malignant nerve sheath tumor</td>
<td>1 (0/1)</td>
<td>33</td>
<td>0.876</td>
</tr>
<tr>
<td>Meningioma</td>
<td>55 (15/40)</td>
<td>57.3 ± 14.6 (60.0/56.3)</td>
<td>0.796–2.677 (0.796–2.677)</td>
</tr>
<tr>
<td>Metastatic tumor</td>
<td>29 (20/9)</td>
<td>61.7 ± 10.7 (64.4/55.8)</td>
<td>0.887–1.581 (0.887–1.581/0.887–1.537)</td>
</tr>
<tr>
<td>Olfactory neuroblastoma</td>
<td>1 (1/0)</td>
<td>55</td>
<td>0.968</td>
</tr>
<tr>
<td>Pituitary adenoma</td>
<td>19 (10/9)</td>
<td>53.8 ± 17.2 (51.7/55.7)</td>
<td>0.742–1.612 (0.742–1.612/0.742–1.319)</td>
</tr>
<tr>
<td>Schwannoma</td>
<td>12 (3/9)</td>
<td>49.5 ± 14.3 (43.7/51.4)</td>
<td>1.192–1.661 (1.201–1.486/1.192–1.661)</td>
</tr>
</tbody>
</table>

Note.—Numbers in parentheses are presented as male/female data.
* Numbers in parentheses are means only.
Recognized no difference between astrocytic, oligodendrogial, and mixed oligoastrocytic tumors. Further studies are necessary, because our series contained only one oligodendrogial tumor and one mixed oligoastrocytic tumor. Our finding that the ADC of ependymomas was significantly lower than that of diffuse astrocytomas may help in the currently difficult differentiation of supratentorial ependymomas and grade 2 astrocytomas. Further studies are underway to elucidate the relationship between the ADCs of oligodendrogial and ependymal tumors and their WHO grades.

To our knowledge, the ADCs of ependymomas, medulloblastomas, and PNETs have not been compared to date. We found that the ADC of ependymomas was higher than that of PNETs, and there was no overlapping. Since the ADC of ependymomas was consistently higher than $1 \times 10^{-3}$ mm²/sec and that of PNET was consistently lower than $1 \times 10^{-3}$ mm²/sec, we suggest that preoperative determination of the ADC of fourth-ventricle tumors makes possible the differential diagnosis between ependymomas and medulloblastomas.

In our search of the literature, we were unable to find reports on the ADC of subependymomas and central neurocytomas. Since we found the ADC of central neurocytomas to be much lower than that of subependymomas, we posit that preoperative evaluation of the ADC of tumors at the lateral ventricle may provide useful information for differentiating between central neurocytomas and subependymomas. However, a study of patient populations larger than ours is necessary to determine the ADCs of rare tumors, such as central neurocytomas, pleomorphic xanthoastrocytomas, gangliogliomas, subependymomas, and other rare neuroepithelial tumors.

**DNTs**

DNTs may develop in any part of the supratentorial cortex and may arise in the area of the caudate nucleus, cerebellum, third ventricles and pons, and septum pelliculum. The preoperative diagnosis of DNT tends to be difficult because these tumors exhibit nonspecific features on conventional MR and computed tomographic images. The ADC of DNT was higher than that of other WHO grade 1 and grade 2 gliomas. In addition, the ADC of DNTs was higher than that of any other glioneuronal tumors, and there was no overlapping of ADC values. The high ADC of DNTs may be attributable to the presence of large extracellular spaces and their cellularity, which is much lower than that of other human brain tumors.

We suggest that preoperative ADC determination may facilitate a differential diagnosis of DNTs. The preoperative identification of DNTs and the distinction of DNTs from other gliomas have important treatment implications. During long-term follow-up, patients who had undergone complete or incomplete surgical removal of DNTs did not manifest clinical or radiologic evidence of tumor recurrence (18–21).

**Glioblastomas, Metastatic Tumors, and Malignant Lymphomas**

These tumors are sometimes difficult to differentiate from each other when only conventional MR imaging studies are available because they usually manifest as enhanced masses. The ADC of glioblastomas is reportedly lower than that of metastatic tumors (2). However, our results indicate that the ADC is not useful for differentiating between glioblastomas and metastatic tumors.

Reportedly, malignant lymphomas manifested lower ADCs than did high-grade astrocytomas (15). We restricted our analysis to glioblastomas and excluded anaplastic astrocytomas because the latter sometimes show weak or no enhancement. We found that malignant lymphomas manifested lower ADCs than those of glioblastomas and metastatic tumors. Our results suggest that preoperative evaluation of the ADC may make it possible to obtain a differential diagnosis of malignant lymphoma.

**Meningiomas and Schwannomas**

According to Kono et al (14), the ADC is not indicative of the histologic subtype of meningiomas. Others (3) reported that atypical (WHO grade 2) and anaplastic (WHO grade 3) meningiomas exhibited lower ADC than that of benign meningioma subtypes. We found that the ADC may not be predictive of the degree of malignancy in meningiomas or of their histologic subtype. Studies in large populations are needed to draw definitive conclusions.

To our knowledge, the ADCs of meningiomas and schwannomas—tumors that are sometimes difficult to differentiate—have not been compared to date. We found that the ADC of schwannomas was significantly higher than that of meningiomas. Histologically, schwannomas comprise Antoni type A and type B neurinomas, and their higher ADC may reflect the lower cell density of Antoni type B neurinomas. We posit that preoperative evaluation of the ADC of tumors at the cerebellopontine angle and the middle cranial fossa may provide useful information for differentiating between meningiomas and schwannomas. Some meningioma subtypes, however, such as microcystic meningiomas, may have higher ADC values than those of schwannomas.

**Tumors Developing Parasellar and Pineal Lesions and Other Analyses**

The preoperative differentiation of craniohypophyngiomas and pituitary adeno-
Radiology

mas can be difficult. The squamous-papillary type of craniopharyngioma may contain only solid components, while pituitary adenomas may be complicated by cystic formation. In addition, it can be difficult to differentiate sellar meningiomas and pituitary adenomas. We found that the ADC is a useful parameter for differentiating between craniopharyngiomas and pituitary adenomas and between craniopharyngioma and meningioma but not for differentiating between pituitary adenomas and meningiomas.

Since both epidermoid tumors and chordomas exhibit high signal intensity on DW images, probably as a result of the T2 shine-through effects, which contrib-

### TABLE 3
Logistic Discriminant Analyses of ADC

<table>
<thead>
<tr>
<th>Tumor Type and Observation 0 vs Observation 1</th>
<th>Coefficient of Score</th>
<th>Accuracy (%)</th>
<th>Discrimination Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADC</td>
<td></td>
<td>(Observation, Prediction)</td>
</tr>
<tr>
<td></td>
<td>Constant</td>
<td>Age (y)</td>
<td>Sex*</td>
</tr>
<tr>
<td>Glioneuronal tumors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low grade (grades 1 and 2) vs high grade</td>
<td>8.772</td>
<td>-7.398†</td>
<td>0.031‡</td>
</tr>
<tr>
<td>(grades 3 and 4)</td>
<td></td>
<td>0.332</td>
<td>84.2</td>
</tr>
<tr>
<td>Grade 1 vs grade 2</td>
<td>6.148</td>
<td>-4.363‡</td>
<td>0.079‡</td>
</tr>
<tr>
<td>Grade 3 vs grade 4</td>
<td>8.711</td>
<td>-8.969‡</td>
<td>0.038‡</td>
</tr>
<tr>
<td>Astrocytic tumor grade 2 vs astrocytic tumor grade 3-4</td>
<td>28.873</td>
<td>-20.245‡</td>
<td>-0.029‡</td>
</tr>
<tr>
<td>Astrocytic tumor grade 3 vs astrocytic tumor grade 4</td>
<td>6.849</td>
<td>-7.491‡</td>
<td>0.052‡</td>
</tr>
<tr>
<td>Oligodendrogial tumor grade 2 vs oligodendrogial tumor grade 3</td>
<td>251.015</td>
<td>-193.568</td>
<td>0.425</td>
</tr>
<tr>
<td>Ependymal tumor grade 2 vs ependymal tumor grade 3</td>
<td>581.148</td>
<td>-460.851</td>
<td>-4.627</td>
</tr>
<tr>
<td>Astrocytic and oligodendrogial tumor grade 2 vs astrocytic and oligodendrogial tumor grade 3</td>
<td>34.007</td>
<td>-22.742‡</td>
<td>-0.076</td>
</tr>
<tr>
<td>Astrocytic tumor grade 2 vs oligodendrogial tumor grade 2</td>
<td>11.617</td>
<td>-10.052</td>
<td>0.005</td>
</tr>
<tr>
<td>Astrocytic tumor grade 3 vs oligodendrogial tumor grade 3</td>
<td>0.199</td>
<td>-3.813‡</td>
<td>0.088‡</td>
</tr>
<tr>
<td>Astrocytic and ependymal tumor grade 2 vs astrocytic and ependymal tumor grade 3</td>
<td>9.539</td>
<td>-7.737‡</td>
<td>0.014</td>
</tr>
<tr>
<td>Astrocytic tumor grade 2 vs ependymal tumor grade 3</td>
<td>905.765</td>
<td>-592.633</td>
<td>-3.395</td>
</tr>
<tr>
<td>Astrocytic tumor grade 3 vs ependymal tumor grade 3</td>
<td>30.307</td>
<td>-9.477</td>
<td>-3.867</td>
</tr>
<tr>
<td>Astrocytic tumor grade 2 vs DNT</td>
<td>-123.010</td>
<td>50.984</td>
<td>0.373</td>
</tr>
<tr>
<td>Glioneuronal tumors grade 1 (except DNT) vs DNT</td>
<td>-63.543</td>
<td>-6.808</td>
<td>NA</td>
</tr>
<tr>
<td>Ependymal tumor vs PNET</td>
<td>-103.47</td>
<td>50.482</td>
<td>-0.092</td>
</tr>
<tr>
<td>Glioblastoma, metastatic tumor, and malignant lymphoma</td>
<td>377.076</td>
<td>-364.558</td>
<td>-0.288</td>
</tr>
<tr>
<td>Malignant lymphoma vs glioblastoma and metastatic tumors</td>
<td>-13.219</td>
<td>19.114‡</td>
<td>-0.028</td>
</tr>
<tr>
<td>Glioblastoma vs metastatic tumors</td>
<td>-5.273</td>
<td>2.374</td>
<td>0.041‡</td>
</tr>
<tr>
<td>Meningioma and schwannoma</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benign subtype of meningioma vs malignant subtype of meningioma</td>
<td>-15.297</td>
<td>-15.981</td>
<td>-0.015</td>
</tr>
<tr>
<td>Meningioma vs schwannoma</td>
<td>-16.981</td>
<td>14.992†</td>
<td>-0.049</td>
</tr>
<tr>
<td>Paraseellar developing tumors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pituitary adenoma vs craniopharyngioma</td>
<td>-24.148</td>
<td>15.285‡</td>
<td>0.012</td>
</tr>
<tr>
<td>Pituitary adenoma vs meningioma</td>
<td>1.026</td>
<td>-0.956</td>
<td>0.009</td>
</tr>
<tr>
<td>Pituitary adenoma vs germ cell tumor</td>
<td>15.293</td>
<td>-7.476</td>
<td>-0.178‡</td>
</tr>
<tr>
<td>Craniopharyngioma vs meningioma</td>
<td>4.806</td>
<td>-4.188‡</td>
<td>0.034</td>
</tr>
<tr>
<td>Germ cell tumor vs craniopharyngioma</td>
<td>-190.718</td>
<td>113.801</td>
<td>0.823</td>
</tr>
<tr>
<td>Meningioma vs germ cell tumor</td>
<td>-5.512</td>
<td>-0.156</td>
<td>0.137</td>
</tr>
<tr>
<td>Pineal developing tumors, other analyses</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNET vs germ cell tumor</td>
<td>-46.876</td>
<td>46.643</td>
<td>0.037</td>
</tr>
<tr>
<td>PNET vs glioblastoma</td>
<td>-13.159</td>
<td>13.446‡</td>
<td>0.069‡</td>
</tr>
<tr>
<td>PNET vs meningioma</td>
<td>-23.709</td>
<td>20.782</td>
<td>0.131‡</td>
</tr>
<tr>
<td>PNET vs malignant lymphoma</td>
<td>90.345</td>
<td>-600.571</td>
<td>8.333</td>
</tr>
<tr>
<td>Glioblastoma vs germ cell tumor</td>
<td>-1.214</td>
<td>3.500</td>
<td>-0.103‡</td>
</tr>
<tr>
<td>Glioblastoma vs meningioma</td>
<td>-1.142</td>
<td>-0.977</td>
<td>0.028</td>
</tr>
<tr>
<td>Epidermoid tumor vs chordoma</td>
<td>-714.320</td>
<td>520.8</td>
<td>-0.226</td>
</tr>
</tbody>
</table>

Note.—NA = not applicable; P value could not be calculated because of complete separation.

* Male = 0, female = 1. P < .01.
† P < .01.
‡ P < .05.
ute to the DW images, the differentiation of these two tumors can be difficult. In our study, the ADC of epidermoid tumors was lower than that of chordomas. When compared with the ADC of normal brain, the ADC of epidermoids is somewhat higher, and that of chordomas is much higher (22,23). We posit that determination of the ADC may aid in differentiating between epidermoid tumors and chordomas. Since we cannot rule out that our single-shot echo-planar imaging-based diffusion techniques resulted in degradation because of the diamagnetic susceptibility effects of the skull base, we are in the process of performing fast spin-echo–based diffusion and ADC studies.

Large patient populations must be studied to evaluate the applicability of the ADC for obtaining a differential diagnosis of germ cell tumors, cranioopharyngiomas, pituitary adenomas, epidermoid tumors, and chordomas.

Our study has some limitations. First, our results were compromised by the partial volume effect. Our section thickness at DW MR imaging was 7.5 mm, and thinner section thickness could reduce the partial volume effect. Second, our results may be affected by the spatial distortion of DW images. Multishot echo-planar imaging–based or fast spin-echo–based diffusion techniques will reduce this spatial distortion. In addition, studies with repetition times longer than the 1600 msec we used must be performed to obtain unequivocal results. Third, non-enhanced and enhanced tumors and grade 3 glioneuronal tumors must be addressed; however, only a few cases were available for analysis. Similarly, our series included only a few patients with rare tumors. Further studies are necessary to determine the usefulness of the ADC for differentiating these tumors.

Despite these caveats, we conclude that the ADC is useful for the differentiation of some human brain tumors. Our study revealed a good inverse correlation between ADC and WHO grade 2–4 astrocytic tumors. The ADC of malignant lymphomas was lower than that of glioblastomas and metastatic tumors. Interestingly, the ADC of DNT was higher than that of any other glioneuronal tumors. The ADC of medulloblastomas, PNETs, and pineoblastomas was lower than that of ependymomas; it was lower in meningiomas than in schwannomas, lower in pituitary adenomas than in cranioopharyngiomas, and lower in epidermoid tumors than in chordomas. Although the preoperative assessment of the ADC can be useful for the differentiation of some tumors, because it overlaps in some tumor types, additional evaluation parameters are needed for unequivocal differentiation among different kinds of human brain tumors.

References
CT Sign of Brain Swelling without Concomitant Parenchymal Hypoattenuation: Comparison with Diffusion- and Perfusion-weighted MR Imaging

PURPOSE: To retrospectively evaluate the apparent diffusion coefficient (ADC) on magnetic resonance (MR) images and the perfusion parameters of lesions that show brain swelling without concomitant parenchymal hypoattenuation on computed tomographic (CT) scans.

MATERIALS AND METHODS: Review board approval was obtained, and informed consent was waived. A total of 14 patients (seven men and seven women; mean age, 64 years ± 11) were retrospectively selected from the consecutive 172 patients with acute cerebral ischemia who underwent CT within 6 hours of symptom onset. All patients had brain swelling without parenchymal hypoattenuation, including loss of gray-white matter distinction on CT scans, and they underwent diffusion- and perfusion-weighted MR imaging shortly after CT. CT attenuation, ADC, and perfusion parameters of relative cerebral blood volume (CBV), time to peak (TTP), and relative cerebral blood flow (CBF) were calculated for gray and white matter of the lesion. The measured values were compared with those of the contralateral hemisphere by using the paired t test; comparison of values of perfusion parameters among three subgroups was performed with the Kruskal-Wallis test. Arterial occlusions were determined with MR angiography or conventional angiography.

RESULTS: The mean interval between initial CT and MR imaging was 2.4 hours ± 0.9 (range, 0.4–3.4 hours). The ADC of lesions was similar to that of contralateral normal tissue (mean ADC ratio for gray matter and white matter, 0.99 and 0.97, respectively) (P > .05). Lesions had an increased relative CBV (P < .001), a mild to moderate TTP delay (P < .001), and a variable but not statistically significant reduction of relative CBF. The mean relative CBF of gray matter was less in patients who had complete infarction (0.81 ± 0.16) than that in patients with partial infarction (0.99 ± 0.16) or those with a normal radiologic outcome (1.12 ± 0.22), but this difference was not statistically significant (P > .05). Proximal cerebral artery occlusions were found in all patients. In five (36%) patients, the lesion did not progress to infarction at follow-up.

CONCLUSION: The CT sign of brain swelling without concomitant parenchymal hypoattenuation in patients with acute cerebral ischemia does not represent severe ischemic damage and may suggest ischemic penumbral or oligemic tissue.

Brain swelling with sulcal effacement is one of the early ischemic changes noted with computed tomography (CT), and such early ischemic changes have been used as exclusion criteria (the so-called one-third rule) in large clinical trials of both intravenous and intraarterial thrombolytic therapy (1–4). The one-third rule is defined as the presence of
early ischemic changes on unenhanced CT scans obtained in one-third or more of the middle cerebral artery territory in the patients with acute ischemic stroke. Early ischemic changes on CT scans include parenchymal hypoattenuation, including loss of gray-white matter distinction, and diffuse or focal brain swelling with effacement of cerebral sulci (1,2,5,6). Although the importance of early ischemic changes remains controversial in patients within 3 hours of symptom onset (6,7), the one-third rule has been used widely and recommended as an exclusion criterion for determining thrombolytic therapy in patients within 6 hours of symptom onset (1–3,8). Early ischemic change of parenchymal hypoattenuation indicates severe ischemic edema of brain tissue and irreversible ischemic injury (9). Patients showing parenchymal hypoattenuation of more than one-third of the middle cerebral artery territory receive no benefit and have higher risk of symptomatic hemorrhage after tissue plasminogen activator treatment, as shown in the European Cooperative Acute Stroke Study I trial (5). However, the prognostic importance of a CT sign of brain swelling without concomitant hypoattenuation is questionable (10), although it is rarely documented. It is not clear whether this CT sign represents severe ischemic change of the brain tissue, which may increase the risk of hemorrhage or a poor outcome after thrombolytic therapy. We hypothesized that brain swelling without concomitant hypoattenuation on CT scans might not represent severe ischemic change. Thus, the purpose of our study was to retrospectively evaluate the apparent diffusion coefficient (ADC) on magnetic resonance (MR) imaging, including both diffusion-weighted (DW) and perfusion-weighted (PW) MR imaging shortly after CT scanning; and (c) patients who underwent follow-up CT scanning or MR imaging within 10 days of the initial CT examination. We excluded patients who underwent intravenous or intraarterial thrombolytic therapy before MR imaging. The study was approved by the review board of Sungkyunkwan University School of Medicine (Seoul, Korea), and informed consent was waived for our retrospective study.

The Baseline National Institutes of Health Stroke Scale score was obtained for all patients. Two neuroradiologists (D.G.N. and E.Y.K., with 10 and 6 years, respectively, of experience with CT) independently evaluated CT scans to determine the presence of a CT sign of brain swelling without concomitant parenchymal hypoattenuation. Disagreements between interpreters were decided by consensus. The interpreters were blind to the initial MR images, follow-up images, and clinical information, except for information about which hemisphere of the brain was affected. The presence of this CT sign was determined when obvious asymmetric sulcal effacement was observed in more than one section and when there was no concomitant cortical hypoattenuation, including loss of gray-white matter distinction. The 60-HU window width with a 30-HU center level was generally used, and variable window width was also applied in the evaluation of the lesion on unenhanced CT scans.

CT and MR Imaging
All CT scans were obtained with a helical CT scanner (High-Speed Advantage or LightSpeed Ultra; GE Medical Systems, Milwaukie, Wis). The scanning parameters of unenhanced CT were 120 kV and 240 mA, with an image matrix of 512 × 512, a 23- or 24-cm field of view, and a 5-mm section thickness. MR imaging was performed with a 1.5-T unit (Signa or CV/i; GE Medical Systems) with echo-planar imaging sequences, including DW and PW imaging. The typical stroke MR imaging protocol consisted of DW imaging, PW imaging, T2-weighted gradient-echo imaging, gadolinium-enhanced T1-weighted imaging, and three-dimensional time-of-flight MR angiography. Follow-up MR imaging included fast spin-echo T2-weighted and fast fluid-attenuated inversion-recovery (FLAIR) imaging. DW MR images were obtained in 20 sections with b values of 0 and 1000 sec/mm². Averaged DW MR images were generated online by averaging three orthogonal-axis images. Imaging parameters of DW imaging were as follows: repetition time msec/echo time msec, 6500/96.8; matrix, 128 × 128; field of view, 24 or 28 cm; section thickness, 5 mm; and intersection gap, 2 mm. Perfusion MR imaging was performed with gradient-echo echo-planar sequences (2000/60) during the injection of 0.2 mmol per kilogram of body weight gadopentetate dimeglumine (Magnevist; Schering, Berlin, Germany) at a rate of 4 mL/sec with an MR-compatible power injector (Spectris; Medrad, Pittsburgh, Pa). The bolus of contrast material was followed by a 15-mL bolus of saline that was administered at the same injection rate. The imaging parameters of T2-weighted and FLAIR images included a matrix of 256 × 192, a field of view of either 23 or 24 cm, a section thickness of 5 mm, and an intersection gap of 2 mm. Two signals were acquired with T2-weighted imaging, and one signal was acquired with FLAIR imaging.

Averaged DW images were processed to generate a trace ADC map based on pixel-by-pixel calculation of signal intensity. Perfusion maps of relative cerebral blood volume (CBV), time to peak (TTP), and relative cerebral blood flow (CBF) were generated off-line at a workstation. After eliminating the recirculation of contrast agent with γ-variate curve fitting, the relative CBV was computed by means of a numeric integration of the curve. The TTP was generated by computing the arrival time of contrast material to maximal concentration. The relative CBF map was obtained with the singular value decomposition deconvolution method (11,12). The shape of the arterial input function was determined from the proximal middle cerebral artery contralateral to the affected hemisphere, and the relative CBF was determined as the height of the deconvoluted tissue curve.

Data Processing and Analysis
All DW and PW images were spatially co-registered to the first volume of CT scans to superimpose the regions of interest (ROIs) delineated on CT scans by use of SPM2 software (Wellcome Department of Cognitive Neuroscience, London, England). All CT scans and MR images were co-registered into the volume data with the same 256 × 256 matrix. After selection of patients with CT evidence of brain swelling without hypoattenuation, two interpreters (D.G.N. and E.Y.K.) selected one CT image in each patient that showed the most obvious
brain swelling, by means of consensus. They then independently determined ROIs and manually drew an ROI for the gray matter of the lesion showing gyral swelling without hypoattenuation and another ROI for the subcortical white matter of the lesion. These ROIs were transferred to the corresponding coregistered DW and PW images. The ROI values were averaged from the measurements obtained by the two interpreters. For comparison, mirror ROIs were manually redrawn on CT scans for the gray and white matter of the contralateral hemisphere. An ADC threshold of 1.2 × 10⁻³ mm²/sec was used to minimize partial volume effect with cerebrospinal fluid. The relative ratios of ADC, signal intensity on DW images, relative CBF, and relative CBF of gray and white matter were calculated by dividing the lesion values by the mirror ROI value of the contralateral hemisphere. TTP delay was defined as the difference between the TTP of a lesion and that of the contralateral hemisphere.

The time to CT and MR imaging after symptom onset, the time between initial CT and MR imaging, and the time to follow-up imaging were assessed. The location of gyral swelling without concomitant hypoattenuation was evaluated on CT scans in each patient, and arterial occlusions were assessed on MR angiograms or conventional angiograms by means of consensus of the two neuroradiologists (D.G.N. and E.Y.K.).

The radiologic outcome of the initial lesion was determined by the same two interpreters in consensus on the follow-up CT scan or FLAIR MR image after completion of evaluation of the initial CT scans and MR images. The radiologic outcome was categorized as normal, partial, complete infarct, or complete infarct according to the presence of cortical infarction, which had gyral swelling on the initial CT scan.

**Statistical Analysis**

Statistical analysis was performed using commercially available software (SPSS-PC, version 10.0; SPSS, Chicago, Ill.). The normality of each variable was assessed with the Kolmogorov-Smirnov test. Comparisons of the mean values of perfusion parameters between the lesion and contralateral hemisphere were performed with the paired t test. Comparison of mean values of perfusion parameters among three subgroups (normal, partial, and complete infarction) categorized by the radiologic outcome was performed with the Kruskal-Wallis test because normality of the variable was rejected in one subgroup. The values of perfusion parameters of a subgroup with normal or small infarction at follow-up were compared with those of other patients by using an unpaired t test. A P value of less than .05 was considered to indicate a statistically significant difference. The statistical power was assessed for the statistical test of ADC value by using computer software (13). Interobserver agreement concerning the presence of the CT sign of brain swelling was assessed with κ statistics. Interobserver reliability of measured ROI values was assessed by using intraclass correlation coefficients of the ROI sizes determined by the two raters. The coefficients of lesion size on the CT images were 0.730 for gray matter, 0.932 for white matter, and 0.885 for the overall size of the lesion.

**RESULTS**

Early ischemic change of brain swelling without concomitant hypoattenuation or loss of gray-white matter distinction was observed in 22 (13%) of the 172 patients. Of these 22 patients, eight were excluded because they did not meet inclusion criteria (n = 6) or had poor PW image quality (n = 2); therefore, 14 patients were included for analysis in this study.

Thrombolytic therapy was performed after MR imaging in two of 14 patients (intraarterial thrombolysis in one and intravenous and intraarterial combined therapy in the other). The agreement rate for the presence of a CT sign of brain swelling was 87%, and the κ value was 0.50 (moderate) (14). Table 1 shows the demographic data of 14 patients (seven men and seven women; mean age, 64 years ± 11). The mean time to CT and MR imaging ± standard deviation was 2.8 hours ± 1.5 and 5.2 hours ± 1.8, respectively, after symptom onset. The mean interval between initial CT and MR imaging was 2.4 hours ± 0.9, and the mean time to follow-up imaging was 5.7 days ± 1.6. The initial National Institutes of Health Stroke Scale score was 15.0 ± 5.6. The occlusion of proximal cerebral artery (internal carotid artery, 5; M1 segment, 6; middle cerebral artery bifurcation, 2; and M2 segment, 1) was

**TABLE 1**

Demographic Data of the 14 Patients

<table>
<thead>
<tr>
<th>Patient No./Age (y)/Sex</th>
<th>Location of Arterial Occlusion</th>
<th>Time to CT (h)</th>
<th>Time to MR Imaging (h)</th>
<th>Time to Follow-up (d)</th>
<th>Outcome of the Lesion</th>
<th>Location of Brain Swelling</th>
<th>Baseline National Institutes of Health Stroke Scale Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/60/F</td>
<td>Left ICA</td>
<td>1.8</td>
<td>4.0</td>
<td>7</td>
<td>N</td>
<td>Frontoparietal</td>
<td>14</td>
</tr>
<tr>
<td>2/84/F</td>
<td>Left M1</td>
<td>1.0</td>
<td>3.1</td>
<td>7</td>
<td>N</td>
<td>Parietal</td>
<td>17</td>
</tr>
<tr>
<td>3/54/F</td>
<td>Left M1</td>
<td>2.1</td>
<td>5.0</td>
<td>7</td>
<td>N</td>
<td>Parietal</td>
<td>17</td>
</tr>
<tr>
<td>4/49/F</td>
<td>Left ICA</td>
<td>4.0</td>
<td>7.3</td>
<td>7</td>
<td>N</td>
<td>Parietal</td>
<td>14</td>
</tr>
<tr>
<td>5/55/M</td>
<td>Left M1</td>
<td>4.8</td>
<td>8.2</td>
<td>4</td>
<td>N</td>
<td>Frontoparietal</td>
<td>11</td>
</tr>
<tr>
<td>6/56/M</td>
<td>Right M1</td>
<td>2.0</td>
<td>4.6</td>
<td>7</td>
<td>P</td>
<td>Parietal</td>
<td>6</td>
</tr>
<tr>
<td>7/68/M</td>
<td>Right ICA</td>
<td>5.4</td>
<td>6.2</td>
<td>5</td>
<td>P</td>
<td>Parietal</td>
<td>12</td>
</tr>
<tr>
<td>8/55/M</td>
<td>Right M1</td>
<td>1.7</td>
<td>4.4</td>
<td>4</td>
<td>P</td>
<td>Frontal</td>
<td>12</td>
</tr>
<tr>
<td>9/60/F</td>
<td>Right M1</td>
<td>1.6</td>
<td>4.6</td>
<td>7</td>
<td>P</td>
<td>Parietal</td>
<td>13</td>
</tr>
<tr>
<td>10/66/F</td>
<td>Right M2</td>
<td>1.0</td>
<td>1.4</td>
<td>7</td>
<td>P</td>
<td>Frontoparietal</td>
<td>15</td>
</tr>
<tr>
<td>11/60/M</td>
<td>Left M1</td>
<td>2.3</td>
<td>5.0</td>
<td>6</td>
<td>P</td>
<td>Parietal</td>
<td>22</td>
</tr>
<tr>
<td>12/80/M</td>
<td>Left ICA</td>
<td>4.0</td>
<td>6.2</td>
<td>3</td>
<td>C</td>
<td>Frontoparietal</td>
<td>22</td>
</tr>
<tr>
<td>13/79/F</td>
<td>Right ICA</td>
<td>4.9</td>
<td>7.8</td>
<td>6</td>
<td>C</td>
<td>Parietal</td>
<td>17</td>
</tr>
<tr>
<td>14/80/M</td>
<td>Left M1</td>
<td>2.4</td>
<td>4.9</td>
<td>3</td>
<td>C</td>
<td>Frontotemporal</td>
<td>22</td>
</tr>
</tbody>
</table>

Note.—C = complete infarction, ICA = internal carotid artery, M1 = M1 segment, M2 = M2 segment, MB = middle cerebral artery bifurcation, N = normal, P = partial infarction.
found in all patients who underwent MR angiography (n = 13) or conventional angiography (n = 1). Follow-up CT (n = 5) and MR (n = 9) images showed that lesions with a CT sign of brain swelling without parenchymal hypoattenuation were normal (n = 5) and progressed to partial (n = 6) or complete (n = 3) infarctions (Figs 1–3). In three of six patients with partial infarction at follow-up, there were only small cortical or subcortical infarctions.

The mean CT attenuation of the gray matter showing gyral swelling was slightly greater than that of the contralateral gray matter, but this difference was statistically significant (31.09 ± 3.55 and 30.61 ± 3.30, respectively) (P = .031) (Table 2). The mean values of ADC and signal intensity on DW images of the lesion were similar to those of the contralateral normal tissue and were not statistically significant (P > .12) (Figs 1, 2). Although not statistically significant, the mean ADC of the white matter (ADC ratio, 0.97 ± 0.10) was slightly less than the control value (Fig 3).

Lesion relative CBF was significantly greater for gray matter (1.24 ± 0.19; range, 0.93–1.59) (P < .001) and white matter (1.39 ± 0.35; range, 0.92–2.07) (P = .001) (Table 3) than for lesion relative CBF in the contralateral hemisphere. The mean TTP value was slightly or moderately delayed in most patients (P < .001). Lesion relative CBVs were variable, and the difference between lesion relative CBVs and control values was not statistically significant (P > .05). In two patients with no infarction and small partial infarction at follow-up, the relative CBVs of lesions were significantly greater (>1.2) than those of the contralateral hemisphere. The mean relative CBV and TTP delay of the gray matter were similar among patient subgroups with normal radiologic outcome or partial or complete infarction (P > .05). Although the mean relative CBF value of gray matter was less in patients with complete infarction (0.81 ± 0.16) than in those with partial infarction (0.99 ± 0.16) or normal radiologic outcome.

### Table 2

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Gray Matter</th>
<th>White Matter</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT attenuation (HU)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lesion</td>
<td>31.09 ± 3.55</td>
<td>25.54 ± 2.08</td>
</tr>
<tr>
<td>Control</td>
<td>30.61 ± 3.30</td>
<td>25.38 ± 2.28</td>
</tr>
<tr>
<td>Difference between attenuation of lesion and control</td>
<td>0.48 ± 0.74*</td>
<td>0.17 ± 1.33</td>
</tr>
<tr>
<td>ADC (10−6 mm²/sec)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lesion</td>
<td>867 ± 49</td>
<td>771 ± 84</td>
</tr>
<tr>
<td>Control</td>
<td>875 ± 46</td>
<td>804 ± 96</td>
</tr>
<tr>
<td>ADC ratio</td>
<td>0.99 ± 0.05</td>
<td>0.97 ± 0.10</td>
</tr>
<tr>
<td>Ratio of signal intensity on DW image</td>
<td>1.05 ± 0.08</td>
<td>1.05 ± 0.09</td>
</tr>
</tbody>
</table>

Note.—Data are mean ± standard deviation.

* Significantly different from control value (P = .031).
Brain swelling may be classified into two major groups according to the pathophysiology: It may be classified as either cerebral hyperemia due to increased CBV or cerebral edema due to increased tissue water (15). In cases of traumatic brain injury, the brain swelling may be mainly due to cerebral hyperemia in early hours or days, but brain edema is the major cause of brain swelling several days after head injury (16–18). In acute cerebral ischemia, brain swelling is usually associated with early ischemic edema accompanied by an increase in tissue water. An experimental study (19) showed that brain swelling gradually increases after middle cerebral artery occlusion and is accompanied by ischemic edema and increase of tissue water when the CBF rate falls to less than 10–15 mL·100 g⁻¹·min⁻¹. Ischemic brain swelling without edema may be explained by cerebral hyperemia, which can develop due to either compensatory vasodilation after a mild decrease of CBF (20,21) or early posts ischemic hyperperfusion (22).

Early ischemic change of parenchymal hypoattenuation, including loss of gray-white matter distinction, results from ischemic edema. Attenuation seen on CT scans is linearly proportional to tissue water content (23), and a 1% increase in tissue water content causes a 2–3 HU decrease in attenuation on CT scans (24). Thus, the CT attenuation of brain swelling caused by ischemic edema will decrease, but the CT attenuation of brain swelling due to cerebral hyperemia does not decrease; rather, it theoretically increases due to a relative increase in the blood content of tissue. Our study suggests that brain swelling without hypoattenuation is predominantly caused by cerebral hyperemia and not by ischemic edema. This result also supports the suggestion of von Kummer (10) and von Kummer et al (5) that brain swelling without hypoattenuation is an effect of compensatory vasodilation and is not closely associated with brain tissue damage. In our study, the relative CBF of the lesion was variable, whereas an increase in CBF was found consistently in most patients. This may be explained by the biphasic hemodynamics of CBV (21) and early posts ischemic hyperperfusion (22). Maximal vasodilation occurs before a significant CBF reduction to approximately 50% of that of normal tissue, and cerebral hyperemia can occur in both oligemic and penumbral tissue (20,21), which has also been shown with perfusion MR studies (25–30). Increased CBF in the lesion might be explained by early posts ischemic hyperperfusion.
hyperperfusion, including spontaneous reperfusion or uncommon hyperperfusion during permanent artery occlusion, which may be caused by the opening up or widening of collateral circulation (22). According to the results of previous studies (29,30), the mean lesion volume on DW images was close to the lesion volume of ischemic tissue with TTP delay of more than 6–8 seconds, and TTP delay of 4 or more seconds was found to be associated with functional impairment (29). Thus, moderate TTP delay (<6–8 seconds) of the lesion with a CT sign of only gyral swelling may be consistent with TTP delay of oligemic or penumbral tissue.

The CT finding of brain swelling was found in 14% of patients during a National Institute of Neurological Disorders and Stroke trial and in 79% of patients during a European Cooperative Acute Stroke Study I trial (5,6). Although the incidence of brain swelling without concomitant hypoattenuation was not separately documented in those trials, the incidence of a CT sign showing brain swelling without hypoattenuation was low (occurring in 13% of patients) in our study. Although the reason for low incidence is unclear, brain swelling due to cerebral hyperemia might develop only during a short period of hemodynamic change selectively in the patients with proximal vessel occlusion and relatively good collateral flow. In this study, we included patients with obvious brain swelling, and some patients with mild brain swelling might have been missed because mild brain swelling is subtle on visual inspection.

Our study has limitations. First, the correlation between CT and MR imaging may have been influenced by the time interval between CT and MR imaging. Second, the manually drawn ROIs of gray matter may have included a portion of subcortical white matter. Third, the patient population is relatively small, and further investigations may be needed for verification and to determine the clinical importance.

In this study, we evaluated lesions only by using one image that showed the most obvious brain swelling to avoid the possible inclusion of false-positive lesions, as the CT finding investigated was subtle, and interobserver agreement was not high.

In conclusion, brain swelling without hypoattenuation on CT scans had no acute cytotoxic edema and was associated with a high CBV and no significant CBF decrease. This CT sign of early ischemic change does not represent severe ischemic damage, but it does suggest ischemic penumbral or oligemic tissue. Thus, brain swelling without concomitant hypoattenuation on CT scans should not be used as a CT exclusion criterion for the one-third rule when determining thrombolytic therapy in patients with acute ischemic stroke.

References


The Pituitary Gland: Changes on MR Images During the 1st Year after Delivery

PURPOSE: To longitudinally and prospectively investigate changes in the volume and signal intensity on T1-weighted magnetic resonance (MR) images of the pituitary gland up to 1 year after delivery and evaluate whether termination of lactation has an effect on these parameters.

MATERIALS AND METHODS: All participants provided informed consent for participation in the study, which was approved by the institutional review board. Thirteen volunteers (mean age, 28 years; age range, 26–32 years) underwent MR imaging 2 and 4 weeks after delivery and then at intervals of 0.5–2.0 months until 1 year after delivery. Eight participants terminated lactation during the study period. Sagittal and coronal T1-weighted images were obtained. Signal intensities of the anterior and posterior lobes of the pituitary were calculated relative to that of the pons. The volume of the pituitary was also calculated. Two-tailed paired Student t tests and separate simple linear regression analyses were used to test for statistically significant differences.

RESULTS: The mean pituitary volume was 544 mm³ at 2 weeks, 523 mm³ at 4 months, 512 mm³ at 8 months, and 511 mm³ at 12 months after delivery, with significant differences between 2 weeks and 4 months (P = .002) and between 4 and 8 months (P = .003) after delivery. The mean ratio of the signal intensity of the anterior lobe of the pituitary to the signal intensity of the pons was 1.11 at 2 weeks, 1.07 at 4 months, 1.03 at 8 months, and 1.00 at 12 months after delivery, with significant differences between 2 weeks and 4 months (P = .004) and between 4 and 8 months (P = .0001) after delivery. Termination of lactation had no statistically significant effect on pituitary volume or the ratio of the signal intensity of the anterior or posterior lobe of the pituitary to the signal intensity of the pons.

CONCLUSION: The volume of the pituitary gland decreases up to 8 months after delivery, and the T1-weighted signal intensity of the anterior lobe of the pituitary decreases; termination of lactation has no statistically significant effect on these parameters.

Results of several autopsy studies have revealed that the pituitary gland may be enlarged in pregnant and postpartum women (1–4). Pituitary gland enlargement in pregnancy has also been evaluated in vivo with magnetic resonance (MR) imaging (5,6). The signal intensity of the anterior lobe of the pituitary on T1-weighted MR images may also be increased during pregnancy (7).

Elster et al (6) and Dinc et al (8) analyzed the MR imaging appearance of the pituitary in postpartum women up to 6 months after delivery. Elster et al (6) reported that the volume of the pituitary gland rapidly returns to normal beyond the 1st week after delivery. Dinc et al (8) reported no significant differences in pituitary volume between subjects studied 2–6 months after delivery and nonpregnant control subjects. These studies, however, are cross sectional; thus, a small change in volume might have been masked by individual variation in pituitary volume. To overcome the flaw of cross-sectional analysis in detecting such a small change in pituitary volume, longitudinal analysis is mandatory. However, to the best of our knowledge, no investigators have longitudinally investigated changes in the pituitary after delivery.
Moreover, results of several animal studies have revealed regression of the pituitary gland after termination of lactation (9–11). The effect of lactation on the human pituitary gland has not yet been fully studied (4).

In view of the foregoing, the purpose of our study was to longitudinally and prospectively investigate changes in volume and signal intensity on T1-weighted images of the pituitary gland up to 1 year after delivery and determine whether the termination of lactation has an effect on these parameters.

MATERIALS AND METHODS

Subjects

Thirteen healthy female volunteers who demonstrated no evidence of pituitary disorder participated in this prospective study after delivering a baby. Subject age at delivery ranged from 26 to 32 years (mean, 28 years). Because previous investigators (4,5) have demonstrated differences in pituitary weights between multigravid and primigravid women, our study was limited to primigravid subjects. All subjects underwent neurologic examination by one of the coauthors (M.L.K., with 1 year of experience as a neurology resident and 9 years of experience as a radiologist) and were free from any neurologic signs or symptoms. Subjects taking any medications were excluded from the study because medication might affect pituitary function. Owing to potential risks to fetuses from MR imaging (12), subjects with a possible second pregnancy were either excluded from the study or left the study. As a consequence, one participant left the study 8 months after delivery. The remaining 12 participants participated in the study until 1 year after delivery. No participants underwent MR imaging within 2 weeks after delivery (0–13 days after delivery) because it was not practical for us to ask the subject to come to our institution for the study (it is widely believed in our country that women are supposed to be at rest for at least 2 weeks after delivery). All participants provided informed consent for participation in the study, which was approved by the institutional review board of Kyoto University Medical School.

After delivering a baby, the subjects underwent MR imaging at 2 and 4 weeks after delivery and then at intervals of 0.5–2.0 months until 1 year after delivery. A total of 132 MR images were obtained. All 13 participants breast-fed their babies until at least 6 months after delivery; eight subjects terminated lactation during the study period. After termination of lactation, MR imaging was performed at intervals of 2 weeks for up to 1 month to identify any rapid changes in pituitary volume or signal intensity.

MR Imaging Technique

MR imaging was performed by using a 1.5-T unit (Signa Horizon; GE Medical Systems, Milwaukee, Wis.). Sagittal and coronal T1-weighted images of the pituitary were obtained by using a spin-echo sequence with the following parameters: repetition time msec/echo time msec, 400/14; two signals acquired; 18-cm field of view; image acquisition matrix, 256 × 256; and 3-mm-thick sections without an intersection gap. The polarity of the read-out gradient on the sagittal and coronal images was set so that fat was moved posteriorly and inferiorly, respectively, by chemical shift misregistration artifacts. This technique is important for observing the pituitary without overlapping by fatty marrow in the dorsum sellae or clivus (13–17).

Pituitary Gland Measurements

Images were electronically transferred to a personal computer workstation (Personal Station DP500; Plat’Home, Tokyo, Japan). ExaVision LITE (version 1.02d) software (ZIO Software, Tokyo, Japan) was used for image display and quantitative measurements. Measurements were performed by one of the authors (T.L.H., with 4 years of experience in MR imaging) and confirmed by two experienced neuroradiologists (Y.M. and T.L.H.), who were blinded to subject information. Pituitary shape was assessed on sagittal T1-weighted images by using the five-point scoring system proposed by Elster et al (6). In this scoring system, grade 1 represents a gland with a markedly concave superior surface; grade 2, a gland with a mildly concave superior surface and less than 2 mm of depression centrally; grade 3, a gland with a flat superior surface; grade 4, a gland with a mildly (<2 mm) convex superior surface; and grade 5, a gland with a markedly convex superior surface that appears nearly spherical on sagittal images (6). The superior surface of the pituitary was defined as the surface that is in contact with cerebrospinal fluid. Initial evaluations were made independently, and any disagreements over final conclusions (which occurred with five of the 132 MR images [3.8%]) were resolved by consensus.

Statistical Analyses

The statistical analyses were performed by a statistician (M.R.). Data are presented as means ± standard deviations. Two-tailed paired Student t tests were used to evaluate differences in pituitary volumes and relative signal intensity ratios of the anterior and posterior lobes between MR images obtained at 2 weeks and those obtained at 4 months (n = 13), between those obtained at 4 and those obtained at 8 months (n = 13), and between those obtained at 8 months and
those obtained at 1 year \((n = 12)\) after delivery. The Wilcoxon signed rank test was used to evaluate if there was a difference in pituitary shape score between MR images obtained at 2 weeks and those obtained at 4 months \((n = 13)\), between those obtained at 4 and those obtained at 8 months \((n = 13)\), and between those obtained at 8 months and those obtained at 1 year \((n = 12)\) after delivery. The Bonferroni adjustment procedure was applied for multiple comparisons to keep the type I error less than .05. Thus, \(P\) values less than .0167 (.05 divided by three, the number of comparison groups) were considered to indicate statistically significant differences.

Simple linear regression analyses (one model for one patient) were performed by using the following variables: the time after delivery (independent variable), pituitary volume, and relative signal intensity ratios of the anterior and posterior lobes (dependent variables). In addition, time-series analysis was performed by using forced (because the time of measurement was not equally spaced) generalized least-squares regression to determine whether pituitary volume or signal intensity ratios of the anterior and posterior lobes changed significantly after delivery. Separate simple linear regression analyses were conducted to examine the relationship between the time of termination of lactation (independent variable) and pituitary volume or relative signal intensity ratios of the anterior and posterior lobes of the pituitary (dependent variables). All statistical analyses were performed by using either Stata Intercooled version 7 (Stata, College Station, Tex) or JMP statistical software version 5.0.1 (SAS Institute, Cary, NC).

**RESULTS**

**Volume**

The mean pituitary volume was 544 \(\text{mm}^3\) \(\pm\) 41 at 2 weeks, 523 \(\text{mm}^3\) \(\pm\) 40 at 4 months, 512 \(\text{mm}^3\) \(\pm\) 41 at 8 months, and 511 \(\text{mm}^3\) \(\pm\) 41 at 12 months after delivery. Volumes differed significantly between 2 weeks and 4 months \((P = .002)\) and between 4 and 8 months \((P = .003)\) after delivery but not between 8 and 12 months after delivery. Changes in pituitary volume showed good fits with the time after delivery for all participants (Fig 1).

**Signal Intensity**

The mean ratio of the signal intensity of the anterior lobe of the pituitary relative to the signal intensity of the pons was 1.11 \(\pm\) 0.06 at 2 weeks, 1.07 \(\pm\) 0.07 at 4 months, 1.03 \(\pm\) 0.08 at 8 months, and 1.00 \(\pm\) 0.09 at 12 months after delivery. The signal intensity of the anterior lobe of the pituitary relative to the signal intensity of the anterior lobes changed significantly after delivery. The Wilcoxon signed rank test was used to evaluate if there was a difference in pituitary shape score between MR images obtained at 2 weeks and those obtained at 4 months \((n = 13)\), between those obtained at 4 and those obtained at 8 months \((n = 13)\), and between those obtained at 8 months and those obtained at 1 year \((n = 12)\) after delivery. The Bonferroni adjustment procedure was applied for multiple comparisons to keep the type I error less than .05. Thus, \(P\) values less than .0167 (.05 divided by three, the number of comparison groups) were considered to indicate statistically significant differences.

Simple linear regression analyses (one model for one patient) were performed by using the following variables: the time after delivery (independent variable), pituitary volume, and relative signal intensity ratios of the anterior and posterior lobes (dependent variables). In addition, time-series analysis was performed by using forced (because the time of measurement was not equally spaced) generalized least-squares regression to determine whether pituitary volume or signal intensity ratios of the anterior and posterior lobes changed significantly after delivery. Separate simple linear regression analyses were conducted to examine the relationship between the time of termination of lactation (independent variable) and pituitary volume or relative signal intensity ratios of the anterior and posterior lobes of the pituitary (dependent variables). All statistical analyses were performed by using either Stata Intercooled version 7 (Stata, College Station, Tex) or JMP statistical software version 5.0.1 (SAS Institute, Cary, NC).

**RESULTS**

**Volume**

The mean pituitary volume was 544 \(\text{mm}^3\) \(\pm\) 41 at 2 weeks, 523 \(\text{mm}^3\) \(\pm\) 40 at 4 months, 512 \(\text{mm}^3\) \(\pm\) 41 at 8 months, and 511 \(\text{mm}^3\) \(\pm\) 41 at 12 months after delivery. Volumes differed significantly between 2 weeks and 4 months \((P = .002)\) and between 4 and 8 months \((P = .003)\) after delivery but not between 8 and 12 months after delivery. Changes in pituitary volume showed good fits with the time after delivery for all participants (Fig 1).

**Signal Intensity**

The mean ratio of the signal intensity of the anterior lobe of the pituitary relative to the signal intensity of the pons was 1.11 \(\pm\) 0.06 at 2 weeks, 1.07 \(\pm\) 0.07 at 4 months, 1.03 \(\pm\) 0.08 at 8 months, and 1.00 \(\pm\) 0.09 at 12 months after delivery. The signal intensity of the anterior lobe of the pituitary relative to the signal intensity of the anterior lobes changed significantly after delivery. The Wilcoxon signed rank test was used to evaluate if there was a difference in pituitary shape score between MR images obtained at 2 weeks and those obtained at 4 months \((n = 13)\), between those obtained at 4 and those obtained at 8 months \((n = 13)\), and between those obtained at 8 months and those obtained at 1 year \((n = 12)\) after delivery. The Bonferroni adjustment procedure was applied for multiple comparisons to keep the type I error less than .05. Thus, \(P\) values less than .0167 (.05 divided by three, the number of comparison groups) were considered to indicate statistically significant differences.

Simple linear regression analyses (one model for one patient) were performed by using the following variables: the time after delivery (independent variable), pituitary volume, and relative signal intensity ratios of the anterior and posterior lobes (dependent variables). In addition, time-series analysis was performed by using forced (because the time of measurement was not equally spaced) generalized least-squares regression to determine whether pituitary volume or signal intensity ratios of the anterior and posterior lobes changed significantly after delivery. Separate simple linear regression analyses were conducted to examine the relationship between the time of termination of lactation (independent variable) and pituitary volume or relative signal intensity ratios of the anterior and posterior lobes of the pituitary (dependent variables). All statistical analyses were performed by using either Stata Intercooled version 7 (Stata, College Station, Tex) or JMP statistical software version 5.0.1 (SAS Institute, Cary, NC).
The signal intensity of the posterior lobe of the pituitary relative to the signal intensity of the pons did not differ significantly between any of the compared periods. Relative signal intensities of the posterior lobe of the pituitary were within 5% of each other (Fig 3). Results of linear regression, however, showed good fits in some subjects.

### Relationship between Time after Delivery and Volume and Signal Intensity Ratio

Results of time-series analysis with generalized least-squares regression showed that pituitary volume decreased during the time (in months) after delivery (regression coefficient, −2.0; \( P = .001 \)), as did the ratios of the signal intensity of the anterior lobe of the pituitary to the signal intensity of the pons (regression coefficient, −0.007; \( P = .001 \)), although the changes in the relative signal intensities of the posterior lobes were not statistically significant (regression coefficient, −0.002; \( P = .72 \)).

### Shape

At 2 weeks after delivery, the superior surface of the pituitary gland was classified as grade 3 (\( n = 3 \)), grade 4 (\( n = 3 \)), or grade 5 (\( n = 7 \)). At 4 months after delivery, the surface was classified as grade 2 (\( n = 2 \)), grade 3 (\( n = 4 \)), grade 4 (\( n = 4 \)), or grade 5 (\( n = 3 \)). At 8 months after delivery, the surface was classified as grade 2 (\( n = 2 \)), grade 3 (\( n = 4 \)), grade 4 (\( n = 4 \)), or grade 5 (\( n = 3 \)). At 1 year after delivery, the surface was classified as grade 2 (\( n = 2 \)), grade 3 (\( n = 3 \)), grade 4 (\( n = 5 \)), or grade 5 (\( n = 2 \)) (Table). The Wilcoxon signed rank test revealed a statistically significant difference in pituitary gland shape score between 2 weeks and 4 months after delivery (\( P = .004 \)). However, this score did not differ significantly between 4 and 8 months (\( P > .99 \)) or between 8 and 12 months (\( P = .50 \)) after delivery (Figs 4, 5).

### Relationship to Termination of Lactation

Results of simple linear regression analyses showed no statistically significant relationship between termination of lactation and pituitary volume or relative signal intensity ratios of the anterior or posterior lobes of the pituitary gland in separate models.

### DISCUSSION

To the best of our knowledge, this is the first study in which the changes in the pituitary gland after delivery were longitudinally investigated, revealing that the volume of the pituitary gland decreases after delivery up to 8 months after delivery. Thus, the volume of the pituitary gland decreases for a longer period than previously thought. Analysis of MR imaging findings led Elster et al (6) to report
that the volume of the pituitary gland rapidly returns to normal beyond the 1st week after delivery. Dinc et al (8) also reported no significant differences in pituitary volume between subjects studied 2–6 months after delivery and control subjects. Differences between the results of the present study and those reported by Elster et al (6) and Dinc et al (8) are probably attributable to the cross-sectional nature of those earlier studies. Because there is substantial interindividual variation in pituitary gland volume (19), cross-sectional studies may fail to reveal subtle changes in volume.

In the present study, the pituitary shape scores showed significant differences between 2 weeks and 4 months after delivery; however, these scores did not differ significantly between 4 and 8 months or between 8 and 12 months. This may be contradictory with our finding that pituitary volume significantly differed between 4 and 8 months after delivery; however, we believe the reason the pituitary shape scores did not significantly differ between 4 and 8 months after delivery is that the volume change in this period is too small to be noticed visually.

The results of our study also demonstrated that the signal intensity of the anterior lobe of the pituitary on T1-weighted MR images also decreases up to 8 months after delivery. Miki et al (7) reported that the anterior lobe of the pituitary may be hyperintense on T1-weighted MR images during pregnancy or in the postpartum period. Although the mechanisms of the association between hyperintensity of the anterior pituitary gland on T1-weighted images and pregnancy are not fully understood, they are believed to be related to histologic changes within the lobe (7). The most striking histologic change in the anterior lobe of the pituitary that is associated with pregnancy is hyperplasia of prolactin cells (4,20,21). If it is assumed that the hyperintensity of the anterior lobe of the pituitary on T1-weighted images is related to an increase in the number of prolactin cells, our results concur with those from the pathology study by Scheithauer et al (4), who reported that hyperplasia of prolactin cells gradually disappears within several months after delivery. Because we used T1-weighted images to measure signal intensities, the T2 effect may have partially caused signal intensity changes; in the future, it may be useful to measure true T1 and T2 relaxation times. Other methods, including fat-suppression imaging, T2*-weighted susceptibility imaging, or MR spectroscopy, may also be useful for evaluating the characteristics of the neurochemical changes in the anterior lobe of the pituitary.

Figure 4. Subject 1. Sagittal T1-weighted spin-echo MR images (400/14) of pituitary gland in 26-year-old woman who delivered a baby. Arrowhead indicates the pons. (a) Image obtained 2 weeks after delivery shows that the superior surface of the pituitary gland (arrow) is markedly convex (grade 5). (b–d) Images obtained 4, 8, and 12 months after delivery, respectively, show that the superior surface (arrow) displays mild convexity (grade 4). The anterior lobe of the pituitary is hyperintense to the pons at 2 weeks (a), slightly hypointense to the pons at 4 months (b), and markedly hypointense to the pons at 8 (c) and 12 (d) months after delivery.

Figure 5. Subject 7. Sagittal T1-weighted spin-echo MR images (400/14) of pituitary gland in 32-year-old woman who delivered a baby. Arrowhead indicates the pons. (a) Image obtained 2 weeks after delivery shows that the superior surface of the pituitary gland (arrow) is essentially flat (grade 3). (b–d) Images obtained 4, 8, and 12 months after delivery, respectively, show that the superior surface is mildly concave (grade 2). The anterior lobe of the pituitary appears hyperintense to the pons 2 weeks after delivery (a) and appears almost isointense to the pons 4 (b), 8 (c), and 12 (d) months after delivery.
The signal intensity of the posterior lobe of the pituitary did not differ significantly between 2 weeks and 4, 8, and 12 months after delivery. The posterior lobe of the pituitary is usually hyperintense to the pons on T1-weighted images, and this high signal intensity is well known to be due to neurosecretory granules that contain antidiuretic hormone (13,15, 22,23). Our results suggest that the number of neurosecretory granules that contain antidiuretic hormone does not change significantly after delivery. To the best of our knowledge, there are no reports in the literature in which the relationship between oxytocin and high signal intensity in the posterior lobe of the pituitary has been described; oxytocin might not have a substantial effect on signal intensity in the posterior lobe of the pituitary.

In this study, we failed to identify any statistically significant effect of termination of lactation on volume and T1-weighted signal intensity of the pituitary gland. This finding agrees with those of Elster et al (6), who found no significant differences in height or convexity of the pituitary between lactating and nonlactating women. Further investigation, however, is needed to determine the effect of lactation on the volume of the pituitary gland and the signal intensity of the anterior lobe of the pituitary because the subject population in the current study may have been too small to enable us to detect minor variations.

There were a few drawbacks to this study. Owing to ethical considerations and practical limitations, MR imaging was not performed during pregnancy or within 2 weeks after delivery. The volume of the pituitary gland is reportedly maximal at term or within 1 week after delivery (6,8,24). Therefore, the current investigation does not reveal the percentage decrease in pituitary volume compared to maximum volume. Because we did not measure the volume of the anterior and posterior lobes of the pituitary separately, the change in the volume of the posterior lobe may partly have affected the total pituitary volume, although we believe that this change (if any) is small. Prolactin levels were not measured in this study because we wished to make this study minimally invasive. However, future studies are needed to reveal any correlation between a decrease in gland volume or T1-weighted signal intensity and a decrease in prolactin levels.

The results of our longitudinal in vivo study revealed that both the volume of the pituitary gland and the signal intensity on T1-weighted MR images of the anterior lobe decrease until 8 months after delivery; the termination of lactation had no statistically significant effect on volume and T1-weighted signal intensity. It may be useful to be aware of this physiologic change in the pituitary gland when interpreting MR images in postpartum patients.

References


Fetal Lung-to-Liver Signal Intensity Ratio at MR Imaging: Development of a Normal Scale and Possible Role in Predicting Pulmonary Hypoplasia in Utero\textsuperscript{1}

**PURPOSE:** To define retrospectively a normal range for lung-to-liver signal intensity ratio (LLSIR) in fetuses of 16–40 weeks gestation by using half-Fourier single-shot turbo spin-echo magnetic resonance (MR) imaging.

**MATERIALS AND METHODS:** Approval from the regional ethics review board for retrospective evaluation was obtained, and informed consent was waived. Retrospective analysis and follow-up of 157 pregnant women who underwent MR imaging over the past 4 years were performed. Seventy-four fetuses were subsequently identified as having clinically normal lung function or normal lung morphologic features at autopsy. A total of 141 normal lungs were analyzed, and the LLSIR was calculated from images on an MR workstation. A mixed-effects statistical model was applied, and 95% prediction intervals were calculated. Ten fetuses with hypoplastic lungs at autopsy were also evaluated.

**RESULTS:** Plotting LLSIR against gestational age demonstrated that, according to the fitted mean curve, the signal intensity ratio was higher with more advanced gestational age. Statistical modeling suggests a quadratic relationship between gestational age and LLSIR. For fetuses in the normal population, the LLSIR ranged from 1.52 at 21 weeks gestation to 4.31 at 34 weeks gestation. For all hypoplastic lungs in fetuses at or beyond 25 weeks gestation, the LLSIR was outside the lower bound of the 95% prediction interval for the normal population. The distinction between hypoplastic lungs and normal lungs at less than 25 weeks gestation is less definitive.

**CONCLUSION:** This study provides a normal scale with a 95% prediction interval for LLSIR.

© RSNA, 2005

Sufficient pulmonary development in utero is an important determinant of neonatal viability. A number of conditions, such as prolonged oligohydramnios and intrathoracic mass, put a fetus at risk for pulmonary hypoplasia. Lung hypoplasia results in substantial morbidity and mortality after birth. Unfortunately, an effective method of accurately diagnosing pulmonary hypoplasia prenatally has yet to be determined (1). The usefulness of magnetic resonance (MR) imaging in the diagnosis of this condition has been evaluated by a number of groups. MR volume calculations for fetuses with normal lungs have been compared with volume calculations for fetuses with pulmonary hypoplasia and associated conditions, such as congenital diaphragmatic hernia (2–7). In a previous study, investigators concluded that fetal lung volume could be predicted with moderate accuracy by using MR imaging (8).

While the use of MR imaging–derived lung volumes may be useful in identifying some cases of fetal pulmonary hypoplasia, the neonatal and clinical importance of lung volume
is still unclear. In one study, the majority of fetuses with oligohydramnios had lung volume ranges that were within the normal lung volume range (5). Paek et al (6) determined that relative lung volume was predictive of outcome in cases of left-sided congenital diaphragmatic hernia, while other investigators determined that lung volume differences in surviving and nonsurviving fetuses were statistically insignificant in cases of congenital diaphragmatic hernia (3). In one study, however, researchers noted two cases of congenital diaphragmatic hernia in which the fetuses had similar lung volumes; in these cases, the surviving fetus had a higher lung-to-liver signal intensity ratio (LLSIR) than the nonsurviving fetus (5).

Signal intensity of the lung at MR imaging has been evaluated in several studies and may be useful in diagnosing pulmonary hypoplasia that could be of neonatal importance (9–11). However, Levine et al (9) performed a qualitative examination of signal intensity and suggested that the use of signal intensity in the diagnosis of pulmonary hypoplasia may not be consistent in the second trimester. Kuwashima et al (10) calculated an LLSIR in 13 fetuses with normal lung function and in 10 fetuses with hypoplastic lungs. They suggested that LLSIR may allow prediction of pulmonary hypoplasia. The signal intensity of a tissue is dependent on its distance from the receiver coil. We believe that a ratio comparing the signal intensity of the lungs with that of another structure of similar depth could provide a correction factor. The liver is adjacent to the lung, has relatively homogeneous signal intensity, and, thus, may serve as a control for depth.

To our knowledge, the validity of MR imaging signal intensity and, specifically, LLSIR in the diagnosis of pulmonary hypoplasia has not been determined because of the small number of normal lungs evaluated. The LLSIR could provide an objective noninvasive method for quantifying fetal lung development that can be applied to image analysis post hoc. A normal range for LLSIR in the diagnosis of pulmonary hypoplasia has not been determined because of the small number of normal lungs evaluated. The LLSIR could provide an objective noninvasive method for quantifying fetal lung development that can be applied to image analysis post hoc. A normal range for LLSIR could also enable future application of this scale in the assessment of fetuses at risk for pulmonary hypoplasia. The identification of fetuses with pulmonary hypoplasia would be of considerable clinical importance in the antenatal care of fetuses at risk for this condition. Standard prenatal screening for fetal anomalies by using ultrasonography (US) begins at 16 weeks in our referral base, with subsequent follow-up US performed as clinically indicated. Therefore, the purpose of this study was to define retrospectively a normal range for LLSIR in fetuses of 16–40 weeks gestation by using half-Fourier single-shot turbo spin-echo MR imaging.

MATERIALS AND METHODS

Patients

Between July 1999 and July 2003, 157 pregnant women were referred from the Northern Alberta Perinatal Program for MR imaging evaluation of suspected fetal or uterine anomalies. Approval from the regional ethics review board for retrospective evaluation of these patients was obtained, and informed consent was waived. Follow-up regarding the clinical outcome of all fetuses imaged was achieved through chart review and/or telephone interview (L.J.B., R.B.). A normal population was defined as those fetuses in which clinical assessment showed normal pulmonary function after birth or normal lung morphologic and histologic features at autopsy. A total of 74 fetuses with normal lungs were identified in 73 of 157 pregnancies (71 singleton pregnancies and two dichorionic twin gestations, including one twin gestation in which one fetus died, and only the surviving fetus was available for analysis). The remaining 84 pregnancies were excluded because of fetal renal abnormality (n = 26), fetal lung and/or diaphragm abnormality (n = 25), fetal demise without autopsy (n = 8), fetal cardiac anomaly (n = 6), non-renal-associated oligohydramnios (n = 4), fetal liver abnormality (n = 3), fetus not born at time of analysis (n = 3), trisomy 18 syndrome (n = 2), MR images not available (n = 3), mother less than 18 years old (n = 2), fetal lungs not visible at MR imaging (n = 1), and fetus less than 16 weeks gestation (n = 1).

Among the 84 excluded pregnancies, 22 fetuses underwent autopsy subsequent to demise. Review of the autopsy reports identified 10 fetuses (all from singleton pregnancies) with hypoplastic lung(s) (L.J.B., R.B.). Pathologic criteria for hypoplastic lung diagnosis included a lung-to-body weight ratio of less than 0.015 for fetuses of less than 28 weeks gestation and a lung-to-body ratio of less than 0.012 for fetuses of 28 weeks gestation or more (11). On the basis of the classification criteria previously described (12), hypoplastic lungs were categorized as either associated with oligohydramnios or not associated with oligohydramnios because these two conditions are thought to be structurally different.

MR Imaging and Analysis

At the time of MR imaging, all mothers included in this study were older than 18 years of age, were carrying fetuses of 16 weeks gestation or more, were mentally capable of signing informed consent, and did not have any contraindication for MR imaging (eg, pacemaker). Gestational age at the time of MR imaging for fetuses with normal lungs was determined from the mother’s recollection of last menstrual period and was confirmed at US (n = 40) or was adjusted for last menstrual period at US (n = 33). Gestational age at the time of MR imaging for fetuses with hypoplastic lungs was determined from the mother’s recollection of last menstrual period and was confirmed at US (n = 7) or was adjusted for last menstrual period at US (n = 3) (R.S.C.). MR imaging was performed in all fetuses by using a 1.5-T imager (Symphony; Siemens, Erlangen, Germany) with a phased-array surface coil. Half-Fourier single-shot turbo spin-echo sequences (1100/68 [repetition time msec/echo time msec], 149° flip angle, 19 sections, 4- or 5-mm section thickness, one signal acquired, acquisition time of 21 seconds) that were oriented in the transverse, sagittal, and coronal planes relative to the fetal position were performed. No sedative was administered during imaging.

MR images were analyzed on an MR workstation (Syngo Leonardo; Siemens). Because images of only one lung were available for seven of the 74 fetuses, a total of 141 normal lungs were examined. Six of the fetuses were imaged multiple times at different gestational ages. Our analysis includes only the final antenatal imaging examination in these six fetuses so that results from only one examination were used for all 74 fetuses. Eighteen hypoplastic lungs were reviewed in 10 fetuses because one fetus with hypoplasia had only a single lung that was visible on MR images, and another fetus had unilateral lung hypoplasia at autopsy. MR images were interpreted by one author (R.B.) with 5 years experience reading MR images of the fetus. LLSIR calculations were performed by one author (L.J.B.) with less than 1 year of experience reading MR images of the fetus. A region-of-interest tool was placed within a homogeneous portion of lung and liver from the same series of images by using a method modified from one previously described (10). The por-
tion of the lung and liver that was chosen was visibly free of intraparenchymal vessels and adjacent structures. The region-of-interest tool was placed without consideration of the distance from the receiver coil, thereby providing a quantitative value of signal intensity (Fig 1). The area of the tool ranged from 0.3 to 1.3 cm². For each lung examined, three lung-to-liver ratio readings were used to obtain an average ratio. Each of the three ratios calculated per lung was taken from a different image and/or section and, when possible, from a different orientation plane.

Statistical Analysis

One-way repeated-measures analysis of variance F tests were used to assess differences among the three LLSIR values averaged from the same lung. Numerical summaries, such as mean and standard deviation, summarize the LLSIR and gestational ages. The correlation between the right and left lung of the same fetus was calculated. A paired t test was used to determine whether the difference in mean LLSIR between right and left lungs of the entire population was statistically significant. A mixed-effects model was developed (Y.L.) to capture the relationship between gestational age and LLSIR by using commercially available statistical software packages (SAS, version 8; SAS Institute, Cary, NC; Splus, Mathsoft, Seattle, Wash). A backward model building approach was used. A mixed model is a generalization of the standard linear model in the sense that predictors are entered into the model as an ordinary linear regression, but an additional term, called random effect, is also added. Because repeated measurements were taken in the same fetus, these measurements were correlated. A mixed model is needed to capture the dependence of the left and right lungs in the same fetus through a fetus random effect. The fetus random effect provides a way of modeling the mean, as well as variances and correlations, of the data. The mixed model can be calculated by using statistical software (SAS PROC MIXED; SAS Institute). A 95% prediction interval for normal lungs was determined on the basis of results from the mixed model. These prediction intervals provide limits for the prediction of a new LLSIR score for a new lung at a specified gestational age. A P value of less than .05 was considered to indicate a statistically significant difference.

RESULTS

LLSIR Measurements

The three left lung LLSIR measurements had means of 2.435 ± 0.765 (± standard deviation), 2.488 ± 0.850, and 2.366 ± 0.720. Right lung measurements had means of 2.357 ± 0.576, 2.390 ± 0.635, and 2.360 ± 0.684. No significant differences between the three averaged LLSIR measurements per lung were found in the left (F₂,132 = 2.314, P = .103) or right (F₂,146 = 0.280, P = .756) lungs. Averaging the repeated measure is appropriate given the lack of evidence of a significant difference. The LLSIR of normal lungs is illustrated according to gestational age in Figure 2. Fetuses with normal lung function had a mean gestational age of 28.0 weeks ± 5.4 (median age, 29 weeks; age range, 16–40 weeks). LLSIR was obtained for 74 right lungs and 67 left lungs. The left lungs had a mean LLSIR of 2.437 ± 0.745 (median, 2.16; range, 1.52–4.31), and the right lungs had a mean LLSIR of 2.369 ± 0.587 (median, 2.22; range, 1.57–3.89). The mean LLSIR for left lungs was not significantly different from the mean LLSIR for right lungs (paired t test, t = 1.799; df = 66; P = .077). The left and right lungs from the same fetus were highly linearly related, with a correlation of 0.73 (95% confidence intervals: 0.68–0.78). In the normal population, LLSIR varied from 1.52 at 21 weeks gestation to 4.31 at 34 weeks gestation. The LLSIR was higher in fetuses of more advanced gestational age.

Hypoplastic Lungs

Eighteen hypoplastic lungs were identified at autopsy in 10 fetuses. Information regarding these cases is summarized in Table 1, and the data are plotted in Figure 3. The mean gestational age of fetuses with hypoplastic lungs was 22.7 weeks ± 4.3; these fetuses had a mean LLSIR of 1.500 ± 0.180. Hypoplastic lungs associated with oligohydramnios had a mean signal intensity ratio of 1.370 ± 0.090 (range, 1.22–1.50), whereas hypoplastic lungs not associated with oligohydramnios had a mean signal intensity ratio of 1.660 ± 0.110 (range, 1.55–1.86). The LLSIR scores for hypoplastic lungs were usually less than the LLSIR scores for normal lungs in fetuses of similar gestational age.

Mixed-Model Results

To capture the relationship between LLSIR and gestational age, a mixed model was fit to adjust for the dependence of lung measurements from the same fetus (Table 2). The fitted model suggests a quadratic relationship between gestational age and LLSIR. Although the P value for the quadratic term in gestational age is greater than .05, the quadratic term is the most significant, and our approach dictates that lower order terms be included if higher order terms are in the model. The P value for the quadratic age term is greater than .05 because of the inclusion of age in the model. Left and/or right lung effect was not included in the final model because, when incorporated, the lung effect was not significant (P = .102). The model assumes that the correlation between the left and right lungs from the same fetus is the same for each fetus, and, with this assumption, the estimated correlation was 0.51. The within-fetus variance of 0.1368 estimates the dependence of lungs from the same fetus.

The 95% prediction intervals for a new fetus at a given gestational age can be calculated on the basis of the fitted model. Because the mixed-effect model allows for variability within subjects, multiple LLSIR predictions can be obtained at a specified gestational age. These intervals provide limits for the prediction of LLSIR in a new fetus at a spec-

Figure 1. Coronal half-Fourier single-shot turbo spin-echo MR image (1100/68, 4-mm section thickness, 30 × 30-cm field of view, 218 × 256 matrix) of fetus at 28 weeks gestation shows regions of interest in lung (upper region of interest) and liver (lower region of interest) from which LLSIR was calculated. Areas chosen in each organ were homogeneous and free of organ borders and vascular structures.
ified gestational age. Figures 2 and 3 provide the estimated average LLSIR value at each gestational age for normal lungs, as well as the lower and upper 95% prediction intervals.

All the abnormal lungs we assessed in fetuses of 25 weeks gestation or more had LLSIR scores that were lower than the 95% prediction lower bound (Fig 3). When assessed at 18 weeks gestation, the lungs of fetuses with hypoplasia and associated oligohydramnios were significantly different from those of fetuses in the normal population. Several normal lungs exceeded the upper 95% prediction value.

**DISCUSSION**

Normal lungs that were assessed in this study showed a higher LLSIR at more advanced gestational ages. This finding agrees with previous reports of increasing signal intensity and relaxation times in fetal lungs on T2-weighted MR images through gestation (9,14). Fetal lung fluid is an essential component of lung development. It is likely that the amount of lung fluid that is present in the lungs contributes to the increase in signal intensity ratio with time. During the lung maturation process through the canalicular and alveolar stage, bronchiolitis lumina become larger, and the multiplication of respiratory bronchioles, terminal sacs, and alveoli occurs (15). From 29 weeks to term, alveoli multiply from 29 million to approximately 150 million, which is associated with an increase in surface area and decreasing interstitium thickness between alveoli (16). This maturation process increases the volume available for lung fluid. Accompanying this, the secretion rate of lung fluid increases with gestational age as a result of the developing pulmonary microvascular and epithelial surface area (17). Hypoplastic lungs have fewer and smaller peripheral airspaces and a reduced number of airway branches, arteries, and veins (18). This pathologic condition would result in reduced capacity to contain fetal lung fluid, which is likely to contribute to the lower lung-to-liver signal intensities that were seen in the cases of hypoplasia we evaluated. We did not evaluate normal lungs that exceeded the upper 95% prediction value because our interest focused on the lower limit as a means to separate normal lungs from hypoplastic lungs.

Our data for hypoplastic lungs are in agreement with the data of Kuwashima et al (10). Together, both sets of data suggest that the LLSIR remains low in hypoplastic lungs over time, thereby amplifying the difference between normal lungs and hypoplastic lungs at more advanced gestational ages. Our data suggest that there is a marked difference between normal lungs and hypoplastic lungs at 25 weeks gestation or more. The distinction between hypoplastic lungs and normal lungs at less than 25 weeks gestation is less clear. When considering the normal evolution of fetal lung development, it is possible that normal lungs and hypoplastic lungs share similarities prior to 25 weeks gestation.

Hypoplastic lungs have been divided into two major groups. Wigglesworth and Desai (12) reported that hypoplastic lungs associated with anything other than oligohydramnios have reduced cell number but are structurally normal for gestational age, whereas hypoplastic lungs associated with oligohydramnios are biochemically and structurally immature for gestational age, having reduced cell number, low lung phospholipid content, and impaired development of epithelial and interstitial tissues. Wigglesworth and Desai (12) also suggested that the structural maturation arrest seen in cases of hypoplasia with oligohydramnios is a consequence of a total failure of liquid retention in the lungs. This loss of liquid retention is likely the result of pressure effects and spinal flexion (19). In cases of lung hypoplasia without associated oligohydramnios, such as a thoracic volume reduction, liquid retention in the lungs could still occur (13). Considering

---

**Figure 2.** Scatter plot of LLSIR for normal lungs. Mean curve (solid line) and 95% prediction interval (dotted lines) for each gestational age are shown and are based on mixed-effects statistical model. LLSIR increases with age, likely because of increased lung fluid retained within the lung as lung spaces develop.

**Figure 3.** Graph shows LLSIR for hypoplastic lungs compared with normal LLSIR range. O = hypoplastic lungs associated with oligohydramnios. + = hypoplastic lungs not associated with oligohydramnios. Data on hypoplastic lungs from Kuwashima et al (10) (△) are included for comparison. At 25 weeks and beyond, all LLSIRs for hypoplastic lungs are outside the lower bound.
our limited data, there may be a difference with regard to LLSIR between hypoplastic lungs associated with oligohydramnios and those that are not associated with oligohydramnios. At less than 25 weeks, three (50%) of six hypoplastic lungs associated with oligohydramnios were outside the 95% confidence interval. For hypoplastic lungs not associated with oligohydramnios in this age category, all six lungs were within normal LLSIR range. The structural variation between lungs associated with oligohydramnios and those that are not associated with oligohydramnios may explain this difference.

There are several limitations to consider with our study. First, the use of MR imaging as a screening tool for fetuses in the normal population is not currently feasible, and, therefore, we were dependent on the retrospective review of a patient population that was referred for suspected uterine or fetal anomalies. While we excluded all fetuses with abnormalities believed to have a potential influence on the normality of the lung, we cannot be completely certain that such exclusion occurred. Second, there is a limited amount of data for normal lungs in some gestational age groups. Third, the LLSIR was not obtained at the same distance from the surface coil for each fetus. The signal drop-off associated with this distance may alter the ratios in an unknown manner.

Data from this study have derived a normal range with a 95% prediction interval for LLSIR that will serve as a reference for future studies and aid in the assessment of fetal lung appearance. Our work supports a potential role for MR imaging-derived LLSIR in the antenatal diagnosis of lung hypoplasia, especially after 25 weeks gestation. This information could assist prenatal counseling and optimize patient care. Future studies will include a prospective look at cases of pulmonary hypoplasia and correlate LLSIR with outcome morbidity and mortality.

Acknowledgment: The authors thank B. F. (Peter) Mitchell, MD, Perinatal Research Centre, Edmonton, Alberta, Canada, for reviewing the manuscript.

References
13. Greenough A. Pulmonary hypoplasia. In:


Cerebral MR Venography in Children: Comparison of 2D Time-of-Flight and Gadolinium-enhanced 3D Gradient-Echo Techniques

PURPOSE: To prospectively compare two-dimensional (2D) time-of-flight cerebral magnetic resonance (MR) venography with gadolinium-enhanced three-dimensional (3D) gradient-echo cerebral MR venography in children.

MATERIALS AND METHODS: This investigation had investigational review board approval and was Health Insurance Portability and Accountability Act compliant; parental informed consent was obtained. Thirty-seven patients (20 boys, 17 girls) who ranged in age from 4 days to 15 years underwent 2D and 3D MR venography. Two pediatric neuroradiologists compared the visibility of the superior sagittal, straight, transverse, and sigmoid sinuses and the internal jugular veins on images obtained with the two sequences.

RESULTS: In 17 (46%) of the 37 patients, the sequences were equivalent in terms of their depiction of venous anatomy. In 19 (51%) of the 37 patients, 3D MR venography was superior to 2D MR venography. Suboptimal enhancement of veins occurred in one (3%) patient at 3D MR venography. Venous anomalies suggested at 2D MR venography but not present at 3D MR venography included flow gaps in the nondominant transverse sinuses of four patients, unilateral transverse sinus atresia in eight, and a narrowed superior sagittal sinus in two. Two-dimensional MR venography results failed to reveal a persistent falcine sinus associated with straight sinus atresia in one patient and suggested transverse sinus thrombosis in two patients in whom 3D MR venography results were normal. Additionally, the extent of dural thrombosis was overestimated at 2D MR venography in one patient. As compared with 3D MR venography, 2D MR venography failed to reveal sigmoid sinus stenosis in one patient and poorly depicted posterior fossa dural sinus anatomy in two patients with dural arteriovenous fistula.

CONCLUSION: Three-dimensional MR venography is often superior to 2D MR venography in the delineation of major cerebral venous structures in children. Most of the artifactual loss of vascular signal seen with the use of 2D MR venography occurred in nondominant transverse sinuses.

The major veins draining the brain are often studied by using a two-dimensional (2D) time-of-flight magnetic resonance (MR) venography sequence (1,2). However, the depiction of smaller venous structures and venous structures that have slower flow may be limited by saturation effects (1–3). Relatively recent reports (4–7) suggest that gadolinium-enhanced three-dimensional (3D) MR venography may be superior to 2D MR venography in the depiction of normal dural venous anatomy, thrombotic disease, and nonthrombotic venous stenoses in adults. Thus, the purpose of our study was to prospectively compare 2D time-of-flight cerebral MR venography with gadolinium-enhanced 3D gradient-echo MR venography in children.
MATERIALS AND METHODS

This prospective study was approved by the investigational review board of the University of Texas Southwestern Medical School. Parental informed consent was obtained, and the study conformed to Health Insurance Portability and Accountability Act guidelines. Although technical support for optimization of the protocols used to perform MR venography was provided by an author (J.C.) who is an employee of Philips Medical Systems, the other authors controlled the inclusion of all data and the conclusions reached. There was no financial support for this investigation. Thirty-seven patients (20 boys and 17 girls) who ranged in age from 4 days to 15 years (mean age, 4.6 years) and who had been referred for MR imaging underwent both 2D and 3D MR venography. There was a myriad of clinical indications for MR imaging; no patient underwent MR imaging solely for the purpose of obtaining the MR venography data. The most common indications were seizures (in eight patients) and chronic recurrent headache (in seven patients); four of the seven patients with headache had pseudotumor cerebri. Five patients had macrocephaly or known hydrocephalus, five had an intracranial tumor or had previously undergone resection of an intracranial tumor, and four had an intracranial vascular malformation. Closed head injury and mastoiditis were clinical indications in three patients each, one patient had a tongue mass, and there were clinical indications in three patients.

MR Imaging

MR imaging was performed with superconducting 1.5-T MR units (Intera Version 9.0; Philips, Best, the Netherlands) and standard head coils. Routine MR imaging included the following sequences: sagittal and transverse T1-weighted spin echo (repetition time msec/echo time msec, 450/11; number of signals acquired, two), transverse and/or coronal T2-weighted fast spin echo (3200–4500/90–120; number of signals acquired, three to four), and transverse fluid-attenuated inversion recovery (repetition time msec/echo time msec/inversion time msec, 8000/120/2300). Other pulse sequences were performed as clinically indicated. Two-dimensional MR venography was performed in the coronal plane by using the following parameters: 23/5.1; flip angle, 50°; section thickness, 2 mm; gap, −0.6 mm; matrix, 320 × 512; and field of view, 24 cm with a caudal saturation slab. Image acquisition time was 4 minutes 12 seconds.

Three-dimensional MR venography was timed to achieve maximum enhancement of major venous structures. The dose of gadopentetate dimeglumine (Magnevist; Schering, Berlin, Germany) was 0.2 milliliters per kilogram of body weight. In patients with an indwelling inserted central line or peripheral intravenous line smaller than 21 g, the contrast agent was manually injected as fast as possible. When feasible, the contrast agent was injected with a power injector (Spectris; Medrad, Indianola, Pa) at a rate of 2 mL/sec. A 10–20-mL flush of normal saline followed the contrast agent injection. A fluoroscopically triggered 2D sequence (3.7/1.04; flip angle, 40°; section thickness, 8 cm; field of view, 30 cm) was performed in the sagittal plane during injection of the gadolinium chelate. When the gadolinium chelate was visually detected in the sagittal sinus by the technologist (T.R.), the coronal 3D gradient-echo sequence was initiated with the following parameters: 6.0/2.0; flip angle, 35°; matrix, 320 × 512; field of view, 24 cm; 0.8-mm overcontiguous sections with segmented central k-space ordering; and image acquisition time, 2 minutes 11 seconds.

Image Interpretation

Maximum intensity projections (MIPs) were created at the MR operator console for 2D and 3D MR venography data sets. The MIP images were viewed in the sagittal, transverse, and coronal planes by two pediatric neuroradiologists (N.R. and C.L., working in consensus) with 18 and 2 years of experience, respectively. Images were interpreted on commercially available Sun Microsystems (Santa Clara, Calif) software-based picture archiving and communication system workstations (Magic View Version 42B; Siemens, Erlangen, Germany). Source data from 3D MR venography were transferred to a commercially available 3D workstation (Easy Vision, Philips; or Vitrea Version 3.2.4, Vital Images, Plymouth, Minn) for creation of shaded surface renderings.

The continuity and visibility of the superior sagittal, transverse, sigmoid, and straight sinuses and the internal jugular veins (IJVs) were compared between the 2D and the 3D MR venograms. When an area of signal loss or luminal narrowing was seen equally well with both sequences, the 2D and 3D MR venograms were judged as equal. When 2D MR venograms suggested an area of signal loss or absence of a venous structure but 3D MR venograms showed a patent and continuous corresponding venous structure, the 3D MR venograms were judged to be superior. When narrowing of a venous structure was qualitatively more severe on the 2D MR venograms than on the 3D MR venograms, the 3D MR venograms were judged to be superior. When enhancement of regional vascular structures impeded visualization of a major cerebral vein on the 3D MR venograms but the venous structure was clearly seen on the 2D MR venograms, the 2D MR venograms were considered to be superior.

RESULTS

In the Table, the findings at 2D MR venography are compared with those at 3D MR venography. In 17 (46%) of 37 patients, venous anatomy was normal and depicted equally well with both techniques. The superior sagittal sinus, the straight sinus, the transverse sinus, and the sigmoid sinus were patent and continuous, as were the IJVs. In one patient, there was suboptimal enhancement of the venous structures, and 2D MR venography was considered to be superior to 3D MR venography in this case. In 19 patients (51%), 2D MR venography results suggested variations in venous anatomy or venous disease that were not confirmed with 3D MR venography. Areas of signal loss referred to as flow gaps (3) were seen in a transverse sinus at 2D MR venography in four patients with otherwise normal venous anatomy (Fig 1a); such flow gaps were within the nondominant transverse sinus in three of the four patients. In these four patients, 3D MR venography results indicated that the transverse sinuses were patent and continuous (Fig 1b).

One patient had congenital absence of both transverse sinuses with small sigmoid sinuses; venous drainage occurred via a large occipital sinus (Fig 2). In this patient, 2D MR venography and 3D MR venography were considered to be equivalent in depicting venous anatomy, although the level of confidence with which the diagnosis of congenital absence of the transverse sinuses was made was considerably higher at 3D MR venography. Moreover, fenestration of the straight sinuses was most conspicuous on the shaded surface rendering from the 3D MR venography data set, and the possibility of thrombus within the torcula was excluded.
In two other patients, 2D MR venography failed to reveal the persistent occipital sinuses that were seen at 3D MR venography. In eight patients, 2D MR venography results suggested absence of a transverse sinus but 3D MR venography results indicated that these sinuses were hypoplastic but patent (Fig 3). In addition to suggesting absence of a transverse sinus in two of the four neonates evaluated (ie, patients 2 and 3), 2D MR venography results did not fully depict the caliber of the posterior third of the superior sagittal sinus in these two neonates (Fig 4). In another neonate (patient 21), no straight sinus was seen at 2D MR venography; 3D MR venography revealed a persistent falcine sinus and hypoplasia of the straight sinus (Fig 5).

Four patients (patients 19, 24, 35, and 37) were suspected of having thrombosis.
of a transverse sinus on the basis of increased intravascular signal intensity on fluid-attenuated inversion recovery images (Fig 6a). In two of these patients, 2D MR venography revealed patent transverse sinuses, while in the other two patients, 2D MR venography results suggested transverse sinus thrombus (Fig 6b). Three-dimensional MR venography revealed patent transverse sinuses in three of these four patients (i.e., patients 19, 24, and 35) (Fig 6c), while the remaining patient (patient 37) was found to have thrombosis of the transverse and sigmoid sinuses at both 2D and 3D MR venography. However, 3D MR venography better delineated the caudal extent of the thrombus in that patient (Fig 7).

Three-dimensional MR venography revealed segmental occlusion of both IJVs in two of the four patients; both IJVs in the other two patients were patent and continuous. The four patients with pseudotumor cerebri had no venous disease or disorder that was identifiable at 2D MR venography. In one of these patients, 3D MR venography revealed a stenosis in the sigmoid sinus that drained the dominant sinus (Fig 8). Two of the 37 study patients had occlusive disease of the posterior fossa dural sinuses that was associated with intracranial vascular malformations; 3D MR venography was judged to be superior to 2D MR venography in depicting restricted venous drainage in these patients. In three patients, suboptimal timing of the gadolinium chelate bolus resulted in enhancement of overlapping arterial structures or of prominent posterior condylar veins that obscured the junction between the sigmoid sinus and the IJV; these problems were solved by electronically removing the overlapping structures at the 3D workstations.

**DISCUSSION**

The larger venous structures draining the brain may be studied noninvasively at MR imaging by using phase-contrast techniques, time-of-flight techniques, and contrast material–enhanced 3D sequences (1–7). Phase-contrast MR angiography is limited by gradient imperfections, eddy currents, and long acquisition times and potential lack of sensitivity to slow flow if the selected velocity encoding is incorrect (2,6). Two-dimensional time-of-flight MR venography suffers from progressive signal loss caused by slow-flowing protons and by the flow of protons parallel to rather than perpendicular to the imaging plane; both of these phenomena result in spin saturation (3). Gadolinium chelate reduces the spin saturation and is best administered as a bolus to avoid enhancement of chronically thrombosed venous structures; enhancement of such structures may simulate the appearance of a patent venous sinus (6). Techniques
for bolus injection of gadolinium chelate with subsequent acquisition of images with a 3D rapid gradient-echo sequence have been described previously for the imaging of intracranial venous anatomy in adults (5,6).

However, the use of a power injector in the pediatric population is often limited by the presence of a surgically implanted or percutaneously inserted central line or small peripheral intravenous cannulas, through which power injections are contraindicated at our institution. In children who had one of these lines, we did not insert an additional intravenous catheter but chose instead to manually inject the contrast agent through the indwelling venous catheter. The slower injection of the contrast agent may result in suboptimal enhancement of the venous structures, a problem we observed most often in small infants with small-gauge intravenous lines in the extremities. However, diagnostic 3D MR venograms were acquired even in the smallest infants when the contrast agent was injected through 24-gauge intravenous lines.

The use of automated detection of con-
contrast agent in the cavernous carotid artery and subsequent triggering of the 3D gradient-echo sequence with elliptic centric ordering of phase encoding to ensure that the center of k-space is collected at a time when a high concentration of contrast material is intravascular has been shown to yield a high vessel-to-background signal intensity (8). We did not use automated detection of the gadolinium chelate or triggering of the 3D MR venographic acquisition because these features are not available with our MR units. Predicting the time to maximum contrast agent accumulation in the dural venous sinuses after intravenous administration of a gadolinium chelate is problematic in young children with rapid heart rates. Hence, we used a 2D MR fluoroscopic sequence, during which the technologist began performing 3D MR venography once contrast material was observed in the superior sagittal sinus.

The transverse sinuses are the major conduits of cerebral venous drainage. There are anatomic variations in the appearance of the transverse sinuses at 2D MR venography that may simulate the appearance of thrombus (3,6). Ayanzen et al (3) observed flow gaps in the nondominant transverse sinuses in 30% of healthy patients, mostly adults, who underwent 2D MR venography performed in the coronal plane and commented that these flow gaps could potentially be indistinguishable from dural sinus thrombosis. In a study with adults that involved 2D and 3D MR venography, Farb et al (6) noted that flow gaps seen in nondominant transverse sinuses at 2D MR venography were sometimes not visible when 3D MR venographic sequences were performed after intravenous administration of a gadolinium chelate. In the pediatric population described herein, flow gaps seen within the nondominant transverse sinuses of four patients at 2D MR venography were not present after the gadolinium chelate was administered. As Farb et al (6) suggested, we therefore suggest that the presence of flow gaps at 2D MR venography should be considered an indication to proceed to 3D MR venography before the diagnosis of thrombus within a transverse sinus is made.

Previous reports have also indicated that 3D contrast-enhanced gradient-echo techniques enable better differentiation between atretic and hypoplastic sinuses than does 2D MR venography (4). The insensitivity of 2D MR venography to a low volume of flow or slow flow through a hypoplastic transverse sinus may result in overestimation of the prevalence of transverse sinus atresia. In one large study of pediatric patients involving normal routine MR imaging, 2D MR venography revealed atresia of a transverse sinus in about 13% of patients (9). In the patient population reported herein, 2D MR venography results suggested the absence of one transverse sinus in eight patients in whom 3D MR venography revealed hypoplastic transverse sinuses. Of the 37 patients we evaluated, none had unilateral transverse sinus atresia.

Congenital absence of both transverse sinuses is rare, having been reported in less than 1% of healthy children, and may be misdiagnosed as dural sinus thrombosis (9). In our study, the level of confidence with which the diagnosis of transverse sinus atresia rather than thrombosis was made in patient 20 was strengthened on the basis of 3D MR venography results (Figure 2). In one of the patients in the present study, absence of the straight sinus was seen but could not be diagnosed with any level of certainty at 2D MR venography; 3D MR venography revealed a persistent falci

b. Coronal MIP from 2D MR venography (23/5.1; flip angle, 50°; section thickness, 2 mm) data set shows an apparent filling defect (arrow) in the middle region of the left transverse sinus.

c. Coronal MIP from 3D MR venography (6.0/2.0; flip angle, 35°; section thickness, 0.8 mm) data set shows that the left transverse sinus is patent and dominant.
sis, as seen with 3D MR venography, occurs in the pediatric population and the clinical and hemodynamic importance of venous stenoses have not been studied in children. As part of an ongoing investigation at our institution, patients with pseudotumor and age-matched control patients are being evaluated with 3D MR venography.

Limitations of this study relate to the lack of blinded interpretation by the neuroradiologists, the small number of patients studied, and the low incidence of thrombotic venous disease. Further comparison of 2D and 3D MR venography in pediatric patients suspected of having or known to have dural sinus thrombosis is needed before definitive conclusions can be drawn about which technique is preferable in this clinical setting. However, in this prospective nonblinded study, gadolinium-enhanced 3D gradient-echo MR venography performed by using segmented central k-space ordering appeared to be superior to 2D time-of-flight MR venography in the depiction of major draining veins in infants and children. At our institution, 3D gadolinium-enhanced MR venography is performed when there is a high clinical suspicion of disease of the major cerebral venous structures or when 2D time-of-flight MR venography results suggest clinically important disease.

References
US Features of the Normal Appendix and Surrounding Area in Children

PURPOSE: To evaluate prospectively the frequency of depiction with ultrasonography (US) of the appendix in children without clinical suspicion of acute appendicitis and to evaluate the US appearance of the normal appendix.

MATERIALS AND METHODS: Between March 2003 and July 2003, 146 consecutive patients (62 boys and 84 girls; mean age, 7 years; age range, 2–15 years) without clinical suspicion of acute appendicitis were examined with US. Patients with cystic fibrosis and those with acute abdominal pain were excluded from the study. Outer diameters, mural thickness, and color Doppler flow were measured. Appendiceal lumen and surroundings of the appendix were determined. The overall diameter and mural thickness of the appendix were examined for relationship to age, weight, or height of the patient. For statistical analysis, the Mann-Whitney test, Student t test, and linear regression analysis were applied.

RESULTS: In 120 (82%) children, the appendix was depicted with US; in 26 (18%) children, this was not possible. In 114 (95%) of the depicted appendices, the position was classical; we observed six (5%) retrocecal appendices. All appendices were compressible. Mean diameter of the appendix was 0.39 cm (range, 0.21–0.64 cm), and the mean mural thickness was 0.18 cm (range, 0.11–0.27 cm). The appendiceal lumen was empty in 74 (62%) children. The others were filled with fecal material, gas, or both. In 75 (51%) of the 146 children, lymph nodes were present in the right lower quadrant of the abdomen. We found no relation between the age, weight, or height of the examined child and the overall diameter or wall of the appendix.

CONCLUSION: The results of this study show that a normal appendix can be depicted with US in 82% of asymptomatic children.

© RSNA, 2005

In 1986, graded-compression ultrasonography (US) was reported to have been used in the diagnosis of acute appendicitis in children and adults (1). Because of technical limitations of the US machines used in those days, the only criterion for acute appendicitis was mere depiction of the appendix with US. A normal appendix that was not inflamed, however, would not be depicted. US machines in general and resolution in particular have improved dramatically since then; now, both inflamed and normal appendices can be depicted with US.

To discriminate an inflamed appendix from a normal appendix, several additional criteria for appendicitis were established, such as the outer diameter of the appendix, compressibility of the appendix, shape of the appendix, and absence of gas in the appendiceal lumen (2–4). These criteria have been investigated in the adult population. Less is known, however, about the US appearance of the normal appendix in the pediatric population without clinical suspicion of acute appendicitis. Thus, the purpose of our study was to evaluate prospectively the frequency of depiction with US of the appendix in children without clinical suspicion of acute appendicitis and to evaluate the US appearance of the normal appendix.

MATERIALS AND METHODS

Patients

From March 2003 to July 2003, consecutive patients aged 2–15 years referred to the department of radiology of the Juliana Children’s Hospital to undergo US were included in
our study. The majority of the patients were referred for examination of the urinary tract or because of chronic abdominal pain. Patients who were referred because of acute abdominal pain were excluded from the study. Children with cystic fibrosis were also excluded from this study because in a recent study, an enlarged appendix was found in children with cystic fibrosis but without clinical signs of acute appendicitis (5). Patients younger than 2 years were excluded because of difficulty in performing the examination; these patients are very active, and sometimes they are reluctant to undergo an extensive examination. Furthermore, the incidence of acute appendicitis in such a young age group is low. Patients who had undergone previous appendectomy were also excluded. From a total of 482 abdominal US examinations performed at our department, 146 patients with a mean age of 7 years (age range, 2–15 years) were included in our investigation. The mean age for girls was 7.3 years (age range, 2–14 years), and the mean age for boys was 6.5 years (age range, 2–15 years). Additional US evaluation of the right lower quadrant of the abdomen was performed. The patient population consisted of 62 boys and 84 girls. Informed consent was obtained from each patient and/or his or her parents, as per the rules of our country. Institutional review board approval was obtained for our study. For each patient, the weight (measured in kilograms) and height (measured in meters) were recorded and used to calculate the body mass index (BMI).

Measurements

Transverse and longitudinal images of the appendix were obtained by using an HDI 5000 scanner (ATL HDI 5000; Beethel, Wash) with a 5–12-MHz linear-array transducer. The US examinations were performed by a pediatric radiologist (H.C.H.) or an experienced resident of radiology. The radiologist had 12 years of experience in pediatric abdominal US. The three residents were in the 3rd year of their education, and they each had about 6 months of specific experience.

The maximum time to depict the appendix was set at 15 minutes. If 15 minutes were not enough, the examination was concluded, and the appendix was considered unable to be depicted on the basis of our prior experience. We labeled a structure as a normal appendix when it appeared as a blind-ending lamellated structure without peristalsis (Fig 1). The depiction of the appendix was subdivided into partial or complete (including the tip of the appendix). The position of the appendix was determined (ie, classical or retrocecal). The classical position above the iliac vessels pointing in the direction of the bladder and the position of the appendix above the colon in the paracolic gutter were both set as classical (Fig 2). Subsequently, the overall diameter with and without compression was measured. For measurement of the diameter with compression, we used the graded compression technique described by Puylaert (1). To measure diameters, the electronic calipers were placed between the outer borders of the hypoechoic tunica muscularis. The outer diameters were measured in the transverse plane of the appendix. The mural thickness of the appendix was defined as the distance from the hyperechoic luminal interface to the outer hyperechoic line. Subsequently, the content of the appendiceal lumen was determined. The content was classified either as empty or as filled with gas, fecal material, or both (Fig 3). Furthermore, the blood flow in the appendiceal wall was determined with color Doppler US. The criteria for blood flow were classified as follows: no detect-
able flow, a single vessel within the appendiceal wall, and multiple vessels in the appendiceal wall. Finally, we examined the area surrounding the appendix for the presence of fluid, inflamed fat, and/or mesenteric lymph nodes. We labeled tissue as inflamed fat when it was recognized with US as hyperechoic, noncompressible intraabdominal fatty tissue around the appendix (6). All the measurements and findings were noted immediately after imaging.

Statistical Analysis

The distributions of the continuous variables were tested for normality by using the Kolmogorov-Smirnov test. The upper limit of the normal range of the continuous variables was set at the mean plus two times the standard deviation or the 95th percentile for variables that were not normally distributed. The results that concern discrete variables were expressed as frequencies. For the comparison of subgroups, the Student t test or the Mann-Whitney test for normally distributed subgroups was used. Linear regression analysis was performed by analyzing one variable at a time and was used to examine whether the overall diameter or the thickness of the appendiceal wall is related to age, weight, or height of the examined child. The regression coefficient of the continuous variable indicates the increase or decrease of the diameter or mural thickness of the appendix for every unit increase in the studied continuous variables (eg, age and weight). A P value of less than .05 was considered to indicate a statistically significant difference (F.W., A.S.).

RESULTS

General Characteristics

The general characteristics of the study population are shown in Table 1. Of the 482 patients referred to our department for abdominal US, 146 were included in our study. Reasons for exclusion were cystic fibrosis (n = 3), age younger than 2 years (n = 141), and acute abdominal pain or a medical history of appendectomy (n = 192). The mean age for boys and girls was 6.50 and 7.30 years, respectively. There was no significant difference in age between boys and girls (P = .4).

Appendiceal Appearance

In 118 (81%) patients, the appendix was depicted completely (Table 2), whereas in 26 (18%) patients, it was not depicted at all. Of the 26 patients in whom the appendix could be depicted, eight (31%) had urinary tract problems, 14 (54%) had chronic abdominal pain, and four (15%) had other indications. In two (1%) patients, the appendix was only partially visible. There was a difference in depiction between children with urinary tract problems and chronic abdominal pain (difference, 10% 95% confidence interval: −3%, 23%). In children with chronic abdominal pain, the appendix was depicted less frequently than in children with urinary tract problems. Results of the measurements of the anteroposterior, left-to-right diameter, and appendiceal mural thickness are summarized in Table 2. The position of the appendix was classical in 114 (95%) patients and retrocecal in six (5%). All of the depicted appendices were compressible (Table 3). The appendiceal lumen was empty in 74 (62%) of the patients. In 14 (12%) of the patients, the appendix was filled with gas, and in 29 (24%), the appendix was filled with fecal material. Two appendices were filled with gas and fecal material. Two appendices exhibited a sparse flow. None of the depicted appendices contained an appendicolith.

Surrounding Environment

In three patients, we detected inflamed fat in the right lower quadrant of the abdomen. One of these patients later proved to have Crohn disease and had a terminal ileitis. In the remaining two patients, the appearance of fat had returned to normal by the time follow-up examinations were performed. The cause of the inflamed fat cannot be clarified. This might possibly be due to a local self-limiting infection (Fig 4).

In 75 (51%) of the 146 patients, lymph nodes were present in the right lower quadrant of the abdomen (Fig 5). In 35 (24%) of the patients, we observed more than five nodes. There was no significant difference (P = .1) in presence of lymph nodes between children who were referred for urologic problems and children with chronic abdominal problems.

Regression analysis showed no evidence of a relationship between age, weight, and height of the examined child and the diameter or mural thickness of the appendix. Although there may be a relationship between these variables, the present study with 120 patients had 90% power to depict correlations of more

### Table 1: Patient Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)*</td>
<td>7 ± 3.5</td>
</tr>
<tr>
<td>Weight (kg)*</td>
<td>28.5 ± 13.2</td>
</tr>
<tr>
<td>BMI*</td>
<td>17.3 ± 3.9</td>
</tr>
<tr>
<td>Indication</td>
<td>Urinary tract problems 66, Chronic abdominal pain 63, Other reasons* 17</td>
</tr>
</tbody>
</table>

* Data are mean ± standard deviation. The other reasons consisted of gynecologic disorders, liver disorders, gastric carcinoma, inguinal hernia, umbilical hernia, coccygeal pain, rupture of the spleen, and chronic vomiting.

### Table 2: Appendiceal Diameter and Mural Thickness

<table>
<thead>
<tr>
<th>Diameter and Thickness</th>
<th>Range (cm)</th>
<th>Mean (cm)</th>
<th>Standard Deviation (cm)</th>
<th>Two Standard Deviations (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diameter</td>
<td>Anteroposterior 0.21–0.64</td>
<td>0.39</td>
<td>0.08</td>
<td>0.55</td>
</tr>
<tr>
<td>Left to right</td>
<td>0.28–0.85</td>
<td>0.56</td>
<td>0.11</td>
<td>0.78</td>
</tr>
<tr>
<td>Anteroposterior with compression</td>
<td>0.16–0.64</td>
<td>0.38</td>
<td>0.09</td>
<td>0.56</td>
</tr>
<tr>
<td>Left to right with compression</td>
<td>0.24–0.85</td>
<td>0.57</td>
<td>0.11</td>
<td>0.79</td>
</tr>
<tr>
<td>Mural thickness*</td>
<td>0.11–0.27</td>
<td>0.18</td>
<td>0.06</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Note.—Results are in the 120 children with depicted appendix out of 146 studied. * Nonnormal distribution. † 95% confidence interval: 0.17, 0.19.
TABLE 3
Appendiceal Appearance and Environment

<table>
<thead>
<tr>
<th>Appearance and Environment</th>
<th>No. of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compressibility</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>120 (100)</td>
</tr>
<tr>
<td>No</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Lumen content</td>
<td></td>
</tr>
<tr>
<td>Empty</td>
<td>74 (62)</td>
</tr>
<tr>
<td>Gas</td>
<td>14 (12)</td>
</tr>
<tr>
<td>Fecal material</td>
<td>29 (24)</td>
</tr>
<tr>
<td>Gas and fecal material</td>
<td>3 (2)</td>
</tr>
<tr>
<td>Color Doppler flow</td>
<td></td>
</tr>
<tr>
<td>None or normal</td>
<td>118 (98)</td>
</tr>
<tr>
<td>Increased</td>
<td>2 (2)</td>
</tr>
<tr>
<td>Surrounding area*</td>
<td></td>
</tr>
<tr>
<td>No fat, fluid, or lymph nodes</td>
<td>68 (47)</td>
</tr>
<tr>
<td>Inflamed fat</td>
<td>3 (2)</td>
</tr>
<tr>
<td>Lymph nodes</td>
<td>75 (51)</td>
</tr>
<tr>
<td>Lymph nodes (&lt;1 cm)*</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>71 (49)</td>
</tr>
<tr>
<td>&lt;5</td>
<td>40 (27)</td>
</tr>
<tr>
<td>&gt;5</td>
<td>35 (24)</td>
</tr>
</tbody>
</table>

Note.—Data in parentheses are percentages. Unless otherwise indicated, results are in the 120 children with depicted appendix out of 146 studied.

* Results are in the 146 children studied.

DISCUSSION

In this study, we analyzed how frequently the normal appendix could be depicted in children. We found that 82% of appendices were depicted. Šimonovský (7) found that 49% of appendices were depicted in a group of patients in whom clinical appendicitis was not suspected. In our study, we analyzed only pediatric patients (age range, 2–15 years), while Šimonovský included patients from all age groups (age range, 1–82 years). In a later study, Šimonovský (8) found that 58% of appendices were depicted in asymptomatic patients (age range, 1–82 years).

The difference in size between the diameter with and without compression was small. Our explanation for this finding was that an appendix is usually depicted only when compression is applied during the US examination. With maximal pressure, most appendices could be compressed more.

To our knowledge, only Šimonovský (8) had investigated one characteristic of the appendix in a pediatric population. He analyzed the difference in mean appendiceal mural thickness between children and adults. He found a significant difference in mural thickness between infants (1–6 years of age) and adolescents and adults (16–82 years of age); mural thickness was 0.19 cm in infants and 0.21 cm in adults. In our investigation, we found a mean mural thickness of 0.18 cm. Moreover, we found no relationship between the age of the examined child and the mural thickness of the appendix. An explanation for this finding might be the presence of some lymphoid hyperplasia of the appendix in children, independent of age.

Only 13.5% of the depicted appendices contained intraluminal gas. Rettenbacher et al (2) showed that presence of gas in the appendix is useful in the exclusion of appendicitis, while the absence of gas helps in proving appendicitis. Our results are supported by the results of Rao et al (9). They conclude that intraluminal gas is not a diagnostic finding in patients who undergo imaging for clinical suspicion of appendicitis. Air is seen in both inflamed appendices and normal appendices. Intraluminal air cannot be presumed to indicate that the appendix is normal (9).

We were able to depict the appendix with US in 120 (82%) of 146 patients. In 26 (18%) patients, the appendix was not depicted. An explanation could be the size of the patient. In several studies, obesity is given as a major reason for not depicting the appendix (4,10). We too assumed that the appendix was more difficult to depict in obese children than in patients with normal BMI; however, we found no significant difference in BMI between children in whom the appendix was not depicted and children in whom the appendix was depicted. Our subjects are thinner than average American children. Obesity is a bigger problem in the United States than it is in Europe (11,12).

Another explanation for not depicting the appendix with US is the position of the appendix. Puig et al (10) suggested that a retrocecal position of the appendix...
could be an explanation. In a report by Rettenbacher et al (4), the retrocecal position of the appendix is a reason for not depicting the appendix. Ceres et al (13) found that 28% of the examined children had a retrocecal appendix. In our study, only six (5%) children had a retrocecal appendix. A retrocecal appendix might be a reason for failure to depict the appendix in our investigation.

In eight (12%) of the 66 patients with urinary tract problems and 14 (22%) of the 63 patients with chronic abdominal pain (difference, 10%; 95% confidence interval: −3%, 23%), the appendix could not be observed. This difference in depiction might be due to the fact that in populations with chronic abdominal pain, constipation is a more frequent symptom. Depiction of the appendix is made more difficult by intestines filled with fecal material and air.

A prominent finding in our study was the high percentage of children in which mesenteric lymph nodes smaller than 1 cm were observed in the right lower quadrant of the abdomen. In almost 53% of the patients, we found lymph nodes in this region of the abdomen. In nearly 25% of these patients, we observed more than five lymph nodes. Because there was no difference between children with urinary tract problems and children with chronic abdominal complaints, we may assume that the presence of lymph nodes at US in the abdomen of a child is a nonspecific finding with no clinical importance. In the current literature, little is known about this.

There are some limitations in this study. First, there was a lack of comparison with pediatric patients in whom appendicitis was clinically suspected. We indeed had no pathologic proof that the examined patients actually had a normal appendix. Another limitation was the potential operator experience because of the three residents.

The results of this study show that in 82% of the children without clinical suspicion of appendicitis, the normal appendix can be depicted with US. A transverse diameter of 0.55 cm and 0.54 cm with compression and a mural thickness of 0.25 cm might be considered normal in this population. The lumen of the appendices in most of the children was empty, and the lumen in more than half of the children’s lymph nodes was present at US. We did not find any relationship between the overall diameter or mural thickness and the age, weight, or length of the examined child. Furthermore, our results indicate that the presence of mesenterial lymph nodes smaller than 1 cm in the abdomen of a child is a nonspecific finding.

Acknowledgment: We thank Julien B. C. M. Puylaert, MD, PhD, radiologist at Medical Centre Haaglanden, the Hague, the Netherlands, for his advice.

References
8. Šimonovsky V. Normal appendix: is there any significant difference in the maximal mural thickness at US between pediatric and adult populations? Radiology 2002; 224:333–337.
The Spur Sign

APPEARANCE

The spur sign is seen on transverse computed tomographic (CT) images or conventional obturator oblique (45° internal rotation and 15° cephalic tilt) radiographs of the pelvis. The appearance is the result of an inferiorly directed apex of a triangular fragment of bone (the spur) (Figs 1, 2).

EXPLANATION

The spur sign is produced by a triangular fragment of iliac bone that remains attached to the sacroiliac joint but is separated from the fractured acetabulum. This spur is exposed when the fractured acetabular columns are medially displaced. The spur sign is indicative of fracture in both the anterior and posterior acetabular columns (both-column fracture).

DISCUSSION

The acetabulum is composed of two columns and two walls (Fig 3). The columns represent condensation of trabecular bone along lines of stress and transfer weight-bearing force from the hip joints to the axial skeleton. The walls stabilize the hip joint. The columns are unequal in size and, together, form an inverted Y (Fig 3). The anterior column is composed of the anterior portion of the ilium, including the iliac crest; the anterior superior portion of the acetabular roof; and the pubic symphysis. Though shorter, the posterior column provides the major support for the hip joint and includes the weight-bearing dome of the acetabulum and posterior portion of the ischium.

The columns and walls may be fractured in combination or isolation. The anatomic classification developed by Letournel and Judet illustrates the pathologic and anatomic characteristics of variations and combinations of such fractures (1–3). This classification divides the fractures into five elementary and five associated fracture patterns. Elementary fractures involve the four components of the acetabulum—that is, the two walls and the two columns—and include simple disruption as a transverse fracture. The associated fractures represent variations of the elementary fractures.

The mechanism and force of injury, in part, determine the type of fracture (4). Fractures of the anterior column are often the result of force applied in external rotation, whereas fractures of the posterior column are the result of force applied in internal rotation. Injuries to the posterior wall, referred to as “dashboard-type” injuries, are caused by applying force to a flexed knee. Transverse fractures may occur as the result of force applied to the adducted (high transverse) or abducted (low transverse) hip. The both-column and transverse fracture are among the more common acetabular fractures and account for 18.8%–29.0% and 10.4%–11.3% of fractures, respectively (5,6).

Surgical approach and prognosis vary with fracture pattern. For example, management of both-column fractures commonly employs the iliinguinal approach for surgical reduction and has the lowest success rate (7).

Radiographs, when combined with transverse CT images and three-dimensional CT models, may help simplify the task of fracture classification. Brandser et al (6) and Brandser and Marsh (8) divide the Letournel and Judet classification of acetabular fractures into three basic types: wall fractures, column fractures, and transverse fractures. These basic fracture types have distinguishing imaging features. Wall fractures, unlike transverse and column fractures, do not interrupt the illiotibial and iliopsoas lines and the obturator ring. Rather, anteroposterior and obturator oblique conventional radiographs depict interruption of the margins of the acetabular walls.

Distinguishing transverse fractures from column fractures is more complicated because both fracture types interrupt the two (anterior and posterior) acetabular columns, although along different planes. Column fractures separate the acetabulum into front and back halves. Transverse fractures separate the acetabulum into top and bottom halves. On the anteroposterior radiograph, column and transverse fractures both interrupt the iliopubic and ilioischial lines (Fig 3a). Additional observations, such as those suggested by Brandser et al (6) and Brandser and Marsh (8), enable one to distinguish between these two fracture types.

One observation is a fractured iliac wing. Never seen in transverse fractures, iliac wing disruption occurs in only those fractures involving the anterior column of the acetabulum (6,8). This includes the both-column fracture. Iliac wing frac-
tures, however, may not be detected on radiographs because of overlying bowel gas (6,8).

Another observation is the interruption of the obturator ring. The obturator ring is interrupted in column and T-shaped fractures. The T-shaped fracture is a variant of the transverse fracture, with an added vertical fracture line splitting the ischiopubic segment. Obturator ring fractures, however, can be subtle on radiographs. These hairline fractures through the obturator ring are better seen on transverse CT images. CT may be used to further differentiate column fractures from transverse fractures by means of the orientation of the primary fracture line. In column fractures, the primary fracture line has a horizontal, or medial to lateral, orientation; in transverse fractures, the primary fracture line has a sagittal, or anterior to posterior, orientation (8).

The spur sign, seen on conventional obturator oblique radiographs and transverse CT images, is encountered only in both-column fractures. When the two acetabular columns are simultaneously fractured, the connection between the axial skeleton and the acetabular articular surface provided by the sacroiliac joint is severed, and the acetabulum is set free. Thus, in a both-column fracture, the columns are separated not only from each other but also from the axial skeleton. A piece of iliac bone or the part of the acetabulum that remains attached to the sacroiliac joint may produce the spur sign when the detached acetabular fragments are displaced medially into the pelvis.

CT may also aid in the identification of intraarticular fragments and articular involvement. Three-dimensional models, created from transverse images, may help simplify the task of fracture classification. The goal of fracture management is to establish a congruent joint so as to prevent premature osteoarthrisis (7).

In conclusion, the spur sign is indicative of fracture to both the anterior and posterior acetabular columns.

References
Feasibility of Integrating High-Spatial-Resolution 3D Breath-hold Coronary MR Angiography with Myocardial Perfusion and Viability Examinations

The study was institutional review board approved and Health Insurance Portability and Accountability Act compliant. All subjects provided informed consent. Three-dimensional breath-hold coronary magnetic resonance (MR) angiography with use of steady-state free precession was performed in 12 patients up to 20 minutes after 0.2 mmol gadolinium-based contrast material per kilogram of body weight was administered. Within 24 heartbeats, a spatial resolution of up to 1.0 × 1.2 × 2.0 mm was achieved. Sixty-five (82%) of the 79 visualized coronary artery segments had a grade of 3 or 4 on a four-point scale of depiction. Twenty-seven percent (n = 21) of the 79 segments were assigned a grade of 4; 56% (n = 44), a grade of 3; 16% (n = 13), a grade of 2; and 1% (n = 1), a grade of 1. Coronary MR angiography performed as part of a first-pass myocardial perfusion and viability assessment MR imaging examination is feasible and does not involve additional imaging time.

Three-dimensional (3D) breath-hold coronary magnetic resonance (MR) angiography (1–5) has substantial advantages in terms of time efficiency compared with free-breathing navigator MR imaging of coronary arteries (5–8). With current 3D volume coronary MR angiography techniques, short imaging times for breath-hold acquisitions can be achieved but at the expense of spatial resolution. To reduce imaging time requirements, these techniques typically involve segmentation of the linear k-space trajectory by dividing the number of phase-encoding steps per segment such that three heartbeats (or cardiac R-R intervals) per section are required to yield temporal acquisition windows of 120–150 msec per section. Consequently, the section thickness is fairly thick (3.0 mm) for a typical coronary MR angiographic acquisition because adequate spatial coverage (20–30-mm thickness) during a single breath hold limits the number of sections to six to eight (3,9).

An increasingly important cardiac MR imaging examination is that performed to assess myocardial viability (10–12). To maximize the information derived from this examination, first-pass myocardial perfusion MR imaging is typically performed during the gadolinium chelate contrast material delivery. Subsequently, there is an interval of 10–20 minutes between the administration of the contrast material and the delayed-enhancement (ie, viability) examination. It would be useful to conduct a coronary artery examination during this interval to maximize the use of the imaging time.

A coronary MR angiography examination could yield knowledge about the diseased vessel to complement the information derived from the myocardial perfusion and delayed-enhancement (viability) MR imaging studies. For example, on coronary MR angiograms, one could delineate the approximate location of the culprit lesion in an ischemic myocardial territory or construct a suitable preproce-
dural planning “road map.” Coronary MR angiography findings could also provide advance warning of the presence of an aberrant coronary anatomy before further interventional procedures are performed. Thus, the objective of this study was to determine the feasibility of integrating a coronary MR angiography examination with first-pass myocardial perfusion and viability MR imaging examinations.

1 Materials and Methods

For this pilot evaluation, 12 consecutive patients (11 men, one woman; mean age, 60 years ± 19 [standard deviation]; age range, 22–80 years; mean weight, 72 kg ± 15) who were already scheduled to undergo myocardial viability MR imaging were enrolled. Participation in this study was voluntary, and all subjects provided informed consent. The study was performed in accordance with the human-use guidelines of the participating institutions: It was approved by the institutional review boards of Uniformed Services University of the Health Sciences and Johns Hopkins University or performed according to the local ethics guidelines of Chinese PLA General Hospital and Mie University. The study was also Health Insurance Portability and Accountability Act compliant. D.A.B. is a paid consultant for GE Healthcare Biosciences.

Acquisition Strategy

We used an acquisition strategy that allows imaging of a moderately thick slab—between 24 and 28 mm—with section thicknesses of between 2.0 and 2.4 mm. This strategy involves the use of a 3D, electrocardiographically gated, breath-hold acquisition with steady-state free precession (by using steady-state free precession or fast imaging employing steady-state acquisition techniques) with fat suppression. No additional magnetization preparation (eg, T2 preparation or IR preparation) was used. Unlike conventional gradient-recalled-echo MR angiography, which requires image acquisition during the first pass of contrast material for maximal contrast enhancement, this acquisition technique enables us to exploit the persistent gadolinium chelate-induced contrast enhancement achieved with fast imaging employing steady-state acquisition pulse sequences for an improved vessel image signal-to-noise ratio (SNR) several minutes after the first pass of the contrast material bolus (13).

To complete the acquisition within a short breath hold, a variable temporal sampling scheme (14) was used to divide the acquisition into two R-R intervals per section-encoding partition. This scheme has already been successfully used for the assessment of myocardial viability with a 3D sequence (14,15). Hence, for a 12-section acquisition volume, 24 heartbeats were needed to complete the data acquisitions. With such a small acquisition slab thickness, targeted volume acquisitions (1,2) were made for each coronary artery vessel.

The pulse sequence that we used is illustrated in Figure 1, which shows the relative position of the spectrally selective inversion radiofrequency pulse (16) with respect to the MR imaging sequence for fat suppression. A half α–half repetition time prepulse followed by a 20-dummy radiofrequency excitations was also used to accelerate the approach to the steady state (17). Because the acquisitions were contrast material enhanced, no additional magnetization preparation schemes were used (18). To further reduce the total imaging time, a partial-Fourier (with 0.5 signal acquired) acquisition with 20 overscans were used. The overscan region was defined as the number of additional k-space views in the conjugate k-space region that was acquired to provide an estimate of the low-spatial-frequency phase for partial-Fourier (ie, homodyne) image reconstruction (19).

Imaging Protocol

All MR imaging experiments were conducted by using 1.5-T CV/i or TwinSpeed imaging units (GE Medical Systems, Waukesha, Wis) equipped with high-performance gradient systems (40–50 mT/m and 150 T/m/sec). For the patient examinations, an initial first-pass myocardial perfusion MR imaging examination, with six to nine sections acquired every two R-R intervals (20), was performed by using 0.1 mmol gadolinium chelate contrast material per kilogram of body weight followed by a second 0.1 mmol/kg dose of the agent. This protocol yielded a cumulative 0.2 mmol/kg dose of gadolinium chelate contrast material that was administered as part of the myocardial viability assessment. No imaging was conducted during the second contrast material injection. The time between the two injections was the time required for myocardial perfusion MR imaging, which lasted approximately 1–2 minutes.

After myocardial perfusion MR imaging, targeted-volume 3D coronary artery examinations of as many coronary artery vessels (right coronary artery [RCA] and left main [LM], left anterior descending [LAD], and left circumflex [LCX] arteries) as possible were performed before the start of the delayed-enhancement (viability) study. Imaging planes were optimized for visualization of the longest proximal length of each vessel. Because all subjects had initially been referred for the clinical assessment of myocardial viability, the initiation of viability MR imaging limited the time that was available for coronary artery imaging to 15–20 minutes, and in some patients, this amount of time did not allow the imaging of all coronary arteries. In nine of the 12 patients examined, a precontrast 3D breath-hold acquisition of the RCA also was performed. This enabled us to mea-
sure the vessel SNR and the vessel contrast-to-noise ratio (CNR) before and after contrast material administration so that we could assess the advantages of imaging the coronary arteries after the first-pass perfusion examination. The RCA was chosen for these measurements because it was the most consistently visualized vessel.

Imaging parameters were as follows: 3.9–4.7/1.5–1.9 (repetition time msec/echo time msec), a 24–28-cm field of view, a 2.0–2.4-mm section thickness interpolated to 1.0–1.2 mm, 10–12 acquired sections, a 256 × 192 to 256 × 224 acquisition matrix, and a 65° flip angle. This protocol yielded acquired voxel spatial resolutions of 0.9–1.0 × 1.1–1.3 × 2.0–2.4 mm. The imaging time was 20 or 24 heartbeats (two heartbeats per acquired section), respectively, depending on whether 10 or 12 sections were selected. The mean breath-hold duration for the acquisition of 12 sections was 22 seconds ± 5 (standard deviation), with the imaging time varying according to patient heart rate.

Image Evaluation and Statistical Analyses

The acquired images were independently assessed by two radiologists experienced in cardiac MR imaging (D.A.B. and V.B.H., with 9 and 11 years experience, respectively). The radiologists graded each depicted coronary artery segment (proximal, middle, and distal) by using a four-point depiction scale: Grade 1 indicated poor depiction—that is, the coronary vessel was barely seen or was obscured by noise; grade 2, marginal depiction—that is, the vessel was visible, but confidence in the diagnosis was low; grade 3, good depiction—that is, the vessel was adequately visualized, with confidence in the diagnosis; and grade 4, excellent depiction—that is, the vessel was well visualized. The readers were blinded to the patients' medical histories, other MR image data (ie, perfusion and delayed-enhancement examination results), and conventional coronary angiography results. Each coronary artery segment was defined as a 3-cm vessel segment. Qualitative vessel grades of 3 and 4 were deemed to be diagnostic.

Although each set of vessel segments was from the same coronary artery, the individual segments were in different spatial locations and had different orientations and flow-gradient interactions. Thus, we considered the segments to be separate vessels and did not deem that there was any clustering of data that necessitated different statistical analysis. Vessel SNRs and CNRs were also measured in the proximal RCAs before and after contrast material administration. Paired t test analysis (two-tailed, with two samples for means) of the vessel SNR and CNR data was performed using computer software (Excel 2000; Microsoft, Redmond, Wash). Vessel SNR was defined as follows: $SNR = SIm/SDa$, as measured in a small region of interest, where $SIm$ is the mean vessel signal intensity and $SDa$ is the standard deviation in a region of interest measured in air. Vessel CNR was defined as follows: $CNR = 100 \times (SIm - Slb)/(SIm \times SDa)$, where $Slb$ is the signal intensity of the adjacent background tissue. Assessment of inter-reader correlations was performed by using the Cohen $\kappa$ test (MEDCALC 7.4; MedCalc Software, Mariakerke, Belgium). Coronary artery vessel lengths were measured in each subject, and the mean lengths were calculated (by T.K.F.F.).

Results

Images of 11 RCAs and 10 LAD, 10 LM, and 10 LCX arteries were acquired. With each major coronary artery divided into three segments (proximal, middle, and distal), with the exception that the LM coronary artery was evaluated as a single vessel segment, a total of 103 segments possibly could have been visualized.

Of the 103 vessel segments that could be visualized, 79 were visualized, and the mean grade was 3 or 4 for 29 of 33 "possible" RCA segments, for 15 of 30 possible LAD artery segments (with 10 of 30 segments not seen), for nine of 10 possible LM artery segments, and for 12 of 30 possible LCX artery segments (with 14 of 30 segments not seen). Thus, the majority (65 of 79 [82%]) of all vessel segments observed were noted to have a qualitative grade of 3 or 4: 29 (88%) of 33 visible RCA segments, nine (90%) of 10 visible LM artery segments, 15 (75%) of 20 visible LAD artery segments, and 12 (75%) of 16 visible LCX artery segments. The overall results, with the grades averaged for the two observers, are summarized in the Table. Of all 79 visualized segments, 21 (27%) were assigned a grade of 4; 44 (56%), a grade of 3; 13 (16%), a grade of 2; and 1 (1%), a grade of 1. When the data were evaluated by using $\kappa$ statistics, the $\kappa$ value for the correlation of findings between the two readers was 0.60, indicating moderate agreement.

As noted in the Table, all RCA segments and the LM coronary artery were assigned a mean grade of 3 or higher. In addition, the proximal segments of the LAD and LCX arteries were also deemed to be evaluable and of good image quality. The middle and distal segments of the LAD and LCX arteries were not as well visualized and had mean depiction grades of between 2 and 3. Of all 103 possible segments, 24 (mostly distal LAD and LCX segments) were not visualized.

Image SNRs were compared between the pre- and postcontrast MR images obtained in nine subjects. The postcontrast coronary artery images exhibited a mean

<table>
<thead>
<tr>
<th>Vessel Segment</th>
<th>No. of Visualized Segments*</th>
<th>Mean Grade†</th>
<th>Visualized Segments with Diagnostic and Good Image Quality (%)‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal</td>
<td>11/11</td>
<td>3.7 ± 0.5</td>
<td>91 (10/11)</td>
</tr>
<tr>
<td>Middle</td>
<td>11/11</td>
<td>3.5 ± 0.4</td>
<td>100 (11/11)</td>
</tr>
<tr>
<td>Distal</td>
<td>11/11</td>
<td>3.0 ± 0.9</td>
<td>73 (8/11)</td>
</tr>
<tr>
<td>LM artery</td>
<td>10/10</td>
<td>3.4 ± 0.5</td>
<td>90 (9/10)</td>
</tr>
<tr>
<td>LAD artery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal</td>
<td>10/10</td>
<td>3.2 ± 0.4</td>
<td>90 (9/10)</td>
</tr>
<tr>
<td>Middle</td>
<td>8/10</td>
<td>2.9 ± 0.7</td>
<td>63 (5/8)</td>
</tr>
<tr>
<td>Distal</td>
<td>2/10</td>
<td>2.5 ± 0.7</td>
<td>50 (1/2)</td>
</tr>
<tr>
<td>LCX artery</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal</td>
<td>10/10</td>
<td>3.1 ± 0.6</td>
<td>80 (8/10)</td>
</tr>
<tr>
<td>Middle</td>
<td>5/10</td>
<td>2.7 ± 0.7</td>
<td>60 (3/5)</td>
</tr>
<tr>
<td>Distal</td>
<td>1/10</td>
<td>3.0</td>
<td>(100) (1/1)</td>
</tr>
</tbody>
</table>

* Number of vessel segments visualized/total number of segments.
† Mean qualitative vessel depiction grades (± standard deviation) assigned by two independent observers and tabulated as a function of specific coronary artery vessel segments. Grades were assigned by using a scale of 1–4.
‡ Vessels assigned a score of 3 or higher were considered to be diagnostic and of good image quality. The RCAs and the LM arteries were well depicted in all cases, whereas only the proximal LAD and proximal LCX arteries were consistently well visualized with an acceptable qualitative grade. Numbers in parentheses are the numbers of segments used to calculate the percentages.

**Results**

Images of 11 RCAs and 10 LAD, 10 LM, and 10 LCX arteries were acquired. With each major coronary artery divided into three segments (proximal, middle, and distal), with the exception that the LM coronary artery was evaluated as a single vessel segment, a total of 103 segments possibly could have been visualized.

Of the 103 vessel segments that could be visualized, 79 were visualized, and the mean grade was 3 or 4 for 29 of 33 "possible" RCA segments, for 15 of 30 possible LAD artery segments (with 10 of 30 segments not seen), for nine of 10 possible LM artery segments, and for 12 of 30 possible LCX artery segments (with 14 of 30 segments not seen). Thus, the majority (65 of 79 [82%]) of all vessel segments observed were noted to have a qualitative grade of 3 or 4: 29 (88%) of 33 visible RCA segments, nine (90%) of 10 visible LM artery segments, 15 (75%) of 20 visible LAD artery segments, and 12 (75%) of 16 visible LCX artery segments. The overall results, with the grades averaged for the two observers, are summarized in the Table. Of all 79 visualized segments, 21 (27%) were assigned a grade of 4; 44 (56%), a grade of 3; 13 (16%), a grade of 2; and 1 (1%), a grade of 1. When the data were evaluated by using $\kappa$ statistics, the $\kappa$ value for the correlation of findings between the two readers was 0.60, indicating moderate agreement.

As noted in the Table, all RCA segments and the LM coronary artery were assigned a mean grade of 3 or higher. In addition, the proximal segments of the LAD and LCX arteries were also deemed to be evaluable and of good image quality. The middle and distal segments of the LAD and LCX arteries were not as well visualized and had mean depiction grades of between 2 and 3. Of all 103 possible segments, 24 (mostly distal LAD and LCX segments) were not visualized.

Image SNRs were compared between the pre- and postcontrast MR images obtained in nine subjects. The postcontrast coronary artery images exhibited a mean
improvement in SNR of 150% ± 29 (P < .002), despite being obtained at a substantial delay after the first pass of the contrast material bolus. The mean time after the second contrast material injection (yielding a cumulative dose of 0.2 mmol/kg) for the SNR measurements was 9 minutes ± 4. Pre- and postcontrast MR images (single 2.0-mm sections) obtained in the same patient are shown in Figure 2. Note that the suppression of signal in the myocardium, which was comparable between the pre- and postcontrast images, was due to the high flip angle used in the experiments. The postcontrast image exhibited a mean improvement in CNR of 115% ± 27 (P < .18).

When we measured vessel lengths by using the reformatted 3D volume acquisitions, a mean of 102 mm ± 16 of the RCA was visualized. The mean lengths of the other vessels visualized were 18 mm ± 3 of the LM coronary artery, 56 mm ± 15 of the LAD coronary artery, and 45 mm ± 17 of the LCX coronary artery. These measured lengths were consistent with those reported in other studies (2,3,6,7).

I Discussion

Our study results demonstrate the possibility of using 3D coronary MR angiography in combination with myocardial perfusion and delayed-enhancement visibility MR imaging examinations. As shown in Figure 3, a first-pass perfusion study, a delayed-enhancement study, and visualization of the coronary arteries (LCX artery and RCA in Fig 3) can be completed during a single examination owing to the short imaging times that are possible with 3D breath-hold coronary MR angiography performed by using a variable temporal sampling scheme. With use of the proposed acquisition technique, 82% of the visible coronary artery segments were determined to be of sufficiently high image quality for diagnosis. However, 63% of all possible segments (including visualized and nonvisualized segments) were evaluable (grade ≥ 3). Note that in prior studies (6,7), only the visible segments were considered in the assessment of the diagnostic effectiveness of coronary MR angiography (ie, in terms of percentage of diagnostic segments).

The nonvisualized segments were primarily distal LCX and distal LAD coronary artery segments. This limited visualization probably was due to the fact that the positioning of the imaging planes was targeted for imaging the proximal coronary artery segments and thus not optimal for visualization of the distal coronary artery segments. The use of thicker sections would have improved anatomic coverage and potentially facilitated an increase in the number of evaluable segments. Because high-spatial-resolution MR imaging was desired, however, thicker section thicknesses were not used. Furthermore, the research protocol required the acquisition of delayed-enhancement MR images (for the viability study) 20 minutes after the perfusion examination, and this limited the time that was available for coronary artery imaging. Additional imaging of the coronary arteries could have been performed after the delayed-enhancement examination and thus potentially facilitated an increased number of evaluable segments; however, this approach was not used in the current feasibility study.

Our study results are consistent with those of earlier reported clinical studies (6,7). On the basis of our qualitative evaluations, the percentage of evaluable segments was comparable to the percentages reported with use of free-breathing acquisitions (6,7) but at a substantial reduction in imaging time. The ability to image each coronary artery during a single breath hold allows all three major coronary vessels (RCA and LM, LAD, and LCX arteries) to be imaged within three to four breath holds, and this capability is an advantage for incorporating coronary artery visualization into a comprehensive cardiac examination. The administration of contrast material improved the vessel SNR to 150% of the precontrast image SNR within 20–30 minutes after the initial contrast agent bolus injection.

In the patients enrolled in this study, only two R-R cardiac intervals per section were used. In patients with higher heart rates (>80 beats per minute), segmentation into three R-R intervals per section can be used. This will reduce the temporal acquisition window in patients with high heart rates while permitting the imaging time to remain at less than 27 seconds. With this approach, higher heart rates will result in shorter imaging times, and, thus, increasing the segmentation will not adversely affect the breath-hold period.

There were several limitations to the described MR image acquisition approach. Although 24 heartbeats may constitute a short breath-hold period, it may be challenging for—and even beyond the capability of—some patients, especially those with compromised pulmonary function. For patients who are unable to maintain a 20–24-second breath hold, parallel imaging techniques (21–23) can be used. Parallel imaging allows reduction of the breath hold period by a factor of two or greater. Alternatively, if the breath-hold period can be maintained, higher spatial resolution can be achieved. This permits the acquisition of submillimeter in-plane–spatial-resolution images during a 20–24-second breath hold. With use of short imaging times, the performance of free-
breathing navigator techniques can be improved. Applying parallel imaging further reduces the image acquisition time. Higher spatial resolution navigator-gated MR images can then be acquired by using the proposed 3D steady-state free precession acquisition technique in less than 3 minutes (using a conservative assumption of 10%–20% efficiency), as compared with the 10–15-minute acquisition time required in other reported free-breathing navigator studies (5–7).

Because our study involved the use of a small sample size, it was not a definitive assessment of the sensitivity and specificity of the described approach for assessment of coronary artery disease. Thus, a much larger study with conventional angiographic correlation in each patient is needed.

In conclusion, 3D coronary MR angiography performed in combination with first-pass myocardial perfusion and viability assessments is feasible. With use of a variable temporal sampling scheme, acquisition times (ie, breath-hold times) are short (24 seconds) and SNRs are significantly improved, compared with the acquisition times required and SNRs achieved with precontrast MR imaging. The described technique is most suitable for assessment of proximal and middle coronary artery segments. Thus, it seems feasible to complete a cardiac MR imaging examination that facilitates myocardial perfusion, myocardial viability, and proximal and middle coronary artery vessel assessments within 30–45 minutes.

References
Quantitative Assessment of Left Ventricular Function: Steady-State Free Precession MR Imaging with or without Sensitivity Encoding

Quantitative left ventricular (LV) function was assessed with magnetic resonance imaging in 20 patients by using standard multisection multiphase steady-state free precession (SSFP) imaging and sensitivity encoding (SENSE)-accelerated cine SSFP imaging with identical spatial, contrast, and temporal resolution. The local institutional review board approved the protocol, and all patients gave signed informed consent prior to imaging. The study complied with the Health Insurance Portability and Accountability Act. Results of Bland-Altman analysis showed that both techniques produced similar estimates of LV ejection fraction, LV mass, and blood-to-muscle contrast and demonstrated minimal interobserver variability. The authors showed that it is possible, by combining SENSE with cine SSFP imaging, to reduce acquisition time by 50% without compromising spatial resolution, temporal resolution, or blood-to-muscle contrast-to-noise ratio compared with those achieved by using SSFP imaging without SENSE for quantitative LV function assessment.

An important and routine component of cardiac magnetic resonance (MR) imaging is the determination of global left ventricular (LV) function. Standard multisection multiphase T1-weighted gradient-echo sequences (variously known as fast field echo [FFE], fast spoiled gradient echo, and fast low-angle shot) have been used routinely to acquire a series of short-axis sections throughout the LV volume to assess global LV function (1–3). Steady-state free precession (SSFP)-based sequences (also referred to as balanced FFE, true fast imaging with steady-state precession [4], or fast imaging employing steady-state acquisition [5]) are gaining wider acceptance for the evaluation of LV function (6). SSFP-based sequences are preferred because, unlike conventional T1-weighted FFE sequences in which the signal is proportional to $\sqrt{\text{TR}}$ (where TR is the repetition time), the intrinsically high signal intensity available with SSFP makes it possible to use the high-frequency bandwidth to shorten repetition time and thereby increase temporal resolution. Second, unlike T1-weighted FFE sequences that rely on the modest T1 difference between blood and myocardium (7), SSFP sequences advantageously exploit the greater difference in the T2-to-T1 ratio between the two tissues to provide much higher intrinsic contrast. Third, SSFP sequences do not suffer from the progressive spin saturation over the cardiac cycle that characterizes T1-weighted FFE imaging, because the sequence structure ensures flow compensation, eliminates flow-induced dephasing, and preserves blood-to-mycardium contrast throughout the cardiac cycle. Thus, the advantages of SSFP sequences compared with conventional T1-weighted FFE sequences—higher temporal resolution, better blood-to-mycardium contrast, and higher signal intensity—have made them the sequences of choice for routine evaluation of LV function (8).

Despite these advances, MR imaging assessment of LV function remains a cumbersome process that requires patients to hold their breath repeatedly. It would be desirable to make this process more patient friendly by decreasing the...
number of breath holds required or the duration of individual breath holds. In this regard, recently developed parallel imaging techniques such as sensitivity encoding (SENSE) involve the use of multiple receiver coils to acquire MR images more rapidly without compromising spatial or contrast resolution, though at the cost of a decreased signal-to-noise ratio (SNR) (9,10). The higher intrinsic SNR of the SSFP sequences may offer a means for trading this SNR for faster acquisition by using SENSE to reduce the breath-hold duration and increase patient comfort or to acquire more sections within the same breath-hold duration and thereby decrease the examination time. Therefore, the purpose of this study was to test the hypothesis that the use of SENSE in multislice multiphase cine SSFP imaging could help to reduce acquisition time while producing LV functional assessment data comparable to those produced with standard cine SSFP imaging.

1 Materials and Methods

Study Population

Twenty consecutive patients (10 women with a mean age of 55 years ± 9.3 and age range of 41–67 years; 10 men with a mean age of 59 years ± 13.9 and age range of 41–79 years) were referred for MR imaging assessment of LV function for various clinical reasons (ischemic heart disease [n = 9], cardiomyopathy [n = 4], aortic disease [n = 2], aortic valve disease [n = 1], LV function [n = 1], constrictive pericarditis [n = 1], cardiac mass [n = 1], and congestive heart failure [n = 1]). The results of a two-tailed Student t test revealed no statistically significant difference in the age of the women versus the men (P = .53). Each patient examination consisted of both standard SSFP and SENSE-accelerated SSFP acquisitions. The local institutional review board approved the protocol, and all patients gave written informed consent prior to imaging. Our study complied with the Health Insurance Portability and Accountability Act.

MR Image Data Acquisition

All imaging was performed by using a 1.5-T MR imager (Gyroscan NT-Intera; Philips Medical Systems, Best, the Netherlands) with a five-element synergy cardiac coil and a vector cardiographic gating. The same operator (M.P.) performed all examinations. After initial scout imaging and a reference acquisition, both a conventional cine SSFP sequence and a SENSE-assisted cine SSFP sequence (applied in random order) were used to obtain a series of short-axis sections throughout the LV volume (10–14 sections, each with a thickness of 8 mm, with an intersection gap of 2 mm). In the standard SSFP sequence, the following acquisition parameters were used: repetition time msec/echo time msec, 3.4/1.7; flip angle, 55°; temporal resolution or cardiac phase interval, 36–40 msec; in-plane spatial resolution, 1.50–1.75 mm² (depending on patient size); and breath-hold duration, 10–12 heartbeats per section acquired. In the SENSE-assisted cine acquisition, all parameters were identical to those used in the standard SSFP acquisition except that the number of in-plane phase encoding steps was halved, which resulted in an acquisition time of five to six heartbeats per section. This reduction in acquisition time per section permitted the acquisition of two sections per breath hold, with maintenance of the same breath-hold duration as that with the standard SSFP sequence. A total of 40 complete LV volumes were acquired: 20 with conventional SSFP, and 20 with SENSE-assisted SSFP. From the breath-hold duration and the number of breath holds required to image the entire left ventricle, the acquisition time was calculated for both the SSFP and the SENSE-assisted SSFP acquisitions.

Reference acquisition.—The SENSE reconstruction requires information about the coil-sensitivity profiles, as described by Pruessmann et al (9). It is known that coil sensitivity functions are relatively smoothly varying functions, and, therefore, a coarse sampling of the coil sensitivities is sufficient (11). Complex radiofrequency receiver coil sensitivity within the entire imaging volume was calculated from a low-resolution FFE acquisition with a voxel size of 9 × 9 × 9 mm. Coil sensitivities were estimated by interleaving a body coil acquisition with the synergy coil acquisition for each repetition time. Also, to minimize the potential for misregistration due to differences in breathing position between the reference examination and the acquisition for evaluation of LV function, the low-resolution reference examination was averaged with eight acquisitions (11). The total acquisition time of the reference examination was 52 seconds.

Postprocessing.—Data were transferred to a postprocessing workstation (EasyVision, release 5.0; Philips Medical Systems) for analysis of LV function, which was performed after all 40 image data sets were acquired. The 40 data sets were presented in random order to two independent observers (S.D.F., 14 years of experience in cardiac MR imaging; R.D.K., 1 year of experience in cardiac MR imaging) who were blinded to whether the images were acquired with or without SENSE. For each data set, the observers drew the endocardial and epicardial contours of each section of the LV at end diastole and end systole (6). From these contours, end-diastolic volume, end-systolic volume, and LV mass were computed by using voxel summation. These values were then used to calculate stroke volume with the equation SV = EDV – ESV and ejection fraction with the equation EF = SV/EDV, where SV is stroke volume, EDV is end-diastolic volume, ESV is end-systolic volume, and EF is ejection fraction.

Three irregular regions of interest were drawn on the central short-axis section at end diastole that included areas of myocardium in the septum (100–150 mm²), blood in the middle part of the LV cavity (240–300 mm²), and air outside the body (240–300 mm²). The blood-to-muscle contrast-to-noise ratio was computed as the ratio of the difference between the mean signal intensity of blood and that of myocardium to the standard deviation of noise measured in air outside the body.

Data Analysis

All data are presented as mean ± standard deviation. The agreement between the standard SSFP and SENSE-assisted SSFP measurements of LV ejection fraction and LV mass was assessed by using the method of Bland and Altman (12), as was interobserver reliability for the standard SSFP and SENSE-accelerated SSFP sequences. The statistical significance of the results was assessed by using a paired Student t test. A P value less than .05 was considered to reflect statistical significance. Power calculations were not performed a priori, as the standard deviation of the difference between standard SSFP measurements and SENSE-assisted SSFP measurements was unknown. After these differences were determined, power calculations were performed to calculate the minimum detectable differences.

1 Results

Representative diastolic and systolic LV short-axis images from SENSE-assisted SSFP and standard SSFP acquisitions in the same patient are shown in Figure 1. Quantitatively, there was good agreement between the conventional SSFP and SENSE-assisted SSFP data sets for both ejection fraction (mean bias, −0.2% ± 0.6%).
The limits of agreement (±2 standard deviations) between the two acquisition techniques were narrow: 2.6% to −2.8% for ejection fraction and 10.3 to −6.7 g for LV mass (Fig 2). Further, interobserver variability was low for ejection fraction (mean bias, −0.4% ± 3.2 for conventional SSFP measurements and 0.2% ± 3.0 for SENSE-assisted measurements) and for LV mass (mean bias, 8.0 g ± 6.7 for conventional SSFP measurements and 4.6 g ± 8.3 for SENSE-assisted measurements) (Fig 3). The study had 80% power with an α level of .05 to detect differences of 7% in ejection fraction and of 20 g in LV mass between conventional SSFP and SENSE-assisted SSFP measurements. Finally, no statistically significant difference was detected in the blood-to-muscle contrast-to-noise ratio between the two techniques: 50.1 ± 39.8 with conventional SSFP and 50.4 ± 48.4 with SENSE-assisted SSFP (P = .95, two-tailed paired t test).

Acquisition time for SENSE-assisted examinations (140 seconds ± 21) was half of that for conventional SSFP examinations (278 seconds ± 37) (P < .001, two-tailed paired t test). The total examination time reduction with use of SENSE, including time required for completion of the reference acquisition, was 31% (192 seconds ± 21) (P < .001, two-tailed paired t test).

**Discussion**

Evidence from previous studies suggests that the use of cardiac MR imaging for the assessment of LV function, volume, and mass produces consistently more reliable results than does that of echocardiography (13,14), which in turn means that cardiac MR imaging is also more valuable for the repeated assessment of LV function in patients. In addition, it has been shown that, although LV function parameters such as ejection fraction and cardiac mass estimated by using the SSFP techniques are very similar to those estimated with T1-weighted FFE cine imaging sequences, there are some systematic differences in the measurements obtained with the different techniques (8,15). It has been suggested that these differences are due primarily to the increased conspicuity of the endocardial boundary depicted with the SSFP sequences. Recently, Alfakih et al (6), while documenting the systematic differences in the quantitative parameters estimated by using the T1-weighted FFE and SSFP

---

**Figure 1.** Short-axis cine images in the middle part of the left ventricle. A, C, Images acquired at end diastole. B, D, Images acquired at end systole. A, B, Images obtained with the standard SSFP technique, without SENSE (3.4/1.7; flip angle, 55°; acquisition time, 12 heartbeats). C, D, Images obtained with SENSE-assisted SSFP (3.4/1.7; flip angle, 55°; acquisition time, six heartbeats). The endocardial and epicardial boundaries (white circular contour lines) marked on A and C and the endocardial boundaries marked on B and D were used to compute LV ejection fraction and LV mass. Note the high blood-to-myocardium contrast and the clear delineation of anatomic structures, such as papillary muscles, within the LV cavity.

**Figure 2.** Bland-Altman plots show the degree of agreement between conventional SSFP and SENSE-assisted SSFP in evaluation of LV ejection fraction (EF; left) and LV mass (right). Central lines indicate bias, and outer lines indicate limits of agreement (±2 standard deviations). The bias in the estimation of ejection fraction (−0.2%) and LV mass (1.7 g) between the two methods is small, reflecting a high degree of agreement.
sequences, also provided sex- and age-specific normal ranges for LV and right ventricular dimensions measured with both SSFP and standard FFE cine imaging sequences. This information should facilitate the routine adoption of cine MR imaging for evaluation of LV function. New parallel imaging techniques such as SENSE, which allow faster image acquisition without a reduction in spatial or temporal resolution, could further increase the routine clinical use of MR imaging for cardiac evaluation.

The results of this study suggest that SENSE-assisted cine acquisitions with SSFP can help to reduce examination time while producing LV functional assessment data comparable to those achieved with conventional cine SSFP acquisitions. Both techniques provide similar estimates of ejection fraction and LV mass. Blood-to-muscle contrast-to-noise ratio is also nearly identical with both methods. However, the SENSE-assisted SSFP acquisition can be performed in half the time needed for a conventional SSFP acquisition. It should be noted, however, that while the reference acquisition does not require any special planning and can be performed by any operator, it does introduce an overhead of 52 seconds for the SENSE-assisted SSFP acquisition.

Between the two observers, only the measurements of LV mass showed a large bias (mean, 8 g). This may be because the epicardial boundary of the dark myocardium (in the absence of pericardial fat) is more difficult to delineate against the background of the lung than is the endocardial boundary, which markedly contrasts with both the bright blood and the dark myocardium.

With regard to LV functional assessment, the addition of SENSE to the conventional SSFP cine acquisition has a drawback: an inherent loss of SNR. When the number of phase-encoding steps is reduced by 50%, as in the present study, SNR is decreased by a factor of \( \sqrt{2} \) (9). Fortunately, unlike conventional T1-weighted FFE cine sequences, SSFP sequences have high intrinsic SNR and are less affected by the decrease in SNR imposed by parallel acquisition techniques such as SENSE. The high level of agreement between the results produced with conventional SSFP and SENSE-assisted SSFP acquisition methods in the present study suggests that the loss of SNR in SENSE acquisitions does not substantially alter the results of functional analysis. Interestingly, while our contrast-to-noise results did not show a difference between the SENSE-assisted SSFP and conventional SSFP sequences, in theory the contrast-to-noise ratio for the SENSE-assisted SSFP acquisition should also be decreased by a factor of \( \sqrt{2} \). This may be due to the fact that, unlike conventional SSFP acquisitions, in which the noise is uniformly distributed across the image, the noise in a SENSE acquisition is spatially variant (9). This makes the estimation of noise more difficult in SENSE imaging, and we speculate that the conventional procedure of estimating noise in a region outside the body may lead to the underestimation of noise (9).

A limitation of the study was the lack of a reference standard to validate the quantitative results for measurement of ejection fraction and LV mass with both conventional SSFP and SENSE-assisted SSFP sequences.

In summary, it is feasible to combine SENSE with an SSFP cine sequence to reduce total acquisition time, without compromising spatial or temporal resolution or blood-to-myocardium contrast. Despite the significant reduction in acquisition time, quantitative measurements of LV function obtained with the use of SENSE-accelerated cine SSFP sequences are similar to those obtained with SSFP sequences without SENSE. This reduction in acquisition time can be traded for either faster patient throughput or shorter breath-hold duration to accommodate sicker patients who are not capable of long breath holds.

Acknowledgment: The authors thank Stephen N. Palmer, PhD, for his editorial assistance in the preparation of the manuscript.

References
1. van der Meulen P, Groen JP, Cuppen JJ. Very fast MR imaging by field echoes and

Figure 3. Bland-Altman plots show low intraobserver variability for both conventional SSFP (left) and SSFP with SENSE (right) for evaluation of LV ejection fraction (EF; top) and LV mass (bottom). Central lines indicate bias, and outer lines indicate limits of agreement (±2 standard deviations).


Assessment of Multiple Sclerosis Lesions with Spherical Harmonics: Comparison of MR Imaging and Pathologic Findings

Spherical harmonics (SH) were used to approximate the volume and three-dimensional geometry of multiple sclerosis (MS) lesions in deceased patients. The institutional ethical committee does not require its approval for studies involving pathologic specimens. Pathologic findings were used as the reference standard. In addition, lesion volume was measured with cylindrical approximation (CA). Volumetric comparisons of biases were based on summary statistics, Spearman correlation, Wilcoxon test, and two-way analysis of variance. Shape comparison metrics included mean distance and Dice similarity coefficient (DSC). Eight of 11 lesions had smaller biases with SH method ($P < .001$). Median biases with SH and CA did not differ significantly, as compared with pathologic findings ($r = 1.00$ vs $0.99$, respectively). Variances of the biases were significantly smaller for SH ($P = .04$). Ranges of normalized distance and DSC were 0.1%–2.5% and 75%–96%, respectively. Mean DSC was significantly higher than 70% ($P < .001$). SH method provided unbiased lesion volume and added geometric information that may enable a better understanding of the pathogenesis and lesion evolution over time.

Magnetic resonance (MR) imaging of the brain is the preferable paraclinical test for the diagnosis and assessment of disease progression in patients with multiple sclerosis (MS). MR imaging provides a direct measure of the extent of pathologic changes compared with clinical symptoms, enables detection of subclinical activity, and is sensitive to the long-term accumulation of disease burden within the brain. However, it remains difficult to correlate the exact relationship between the evolution of MS lesions on MR images and the details of pathologic findings. Kirshner et al (1) found that MR imaging can help visualize lesions as small as 3 mm in diameter and that the extent of lesions detected with MR imaging was similar to the extent seen in pathologic examinations. Another comparative study reported that even when extensive areas of abnormal MR imaging signal intensity were observed, only small periventricular plaques were found at dissection (2). De Groot et al (3) found that 44% of the MR imaging–detectable abnormalities were neither visible nor palpable macroscopically. They also found that the size and shape of the lesions in different brain samples derived from the same patients were homogeneous, thus suggesting a patient-specific lesion formation.

A modeling approach has been developed to assess brain lesions in patients with MS by using spherical harmonics (SH) (4). SH can quantitatively define and calculate complicated three-dimensional (3D) geometric features to assess the shape of MS lesions. Moreover, with SH, it is possible to obtain the morphometric measurements of the lesions, as well as an analytic description of the shape, yielding accurate estimations of its features. Consequently, it is possible to quantitatively compare the change in 3D lesions and monitor differences over time. The geometry of the studied lesions...
can be characterized by indices derived from SH, which are invariant to space rotation (5), enabling the reduction of errors resulting from inaccurate patient repositioning during MR imaging examinations.

As a limitation, however, in the previous investigation (4), only artificial MS lesions were simulated to perform statistical validations of the volume approximation with SH. Furthermore, the shape approximation was visually inspected only artificially and in vivo (4). Therefore, in the present study, our purpose was to evaluate SH to measure volume and approximate the 3D geometry of MS lesions in comparison to pathologic findings, which were used as the reference standard.

1 Materials and Methods
Specimens and MR Images

Imaging data were obtained from two pathologic specimens. A half of the brain from each of two deceased patients with MS was obtained. The brains were fixed and preserved in formaldehyde and were donated externally by the Multiple Sclerosis Human Neurospecimen Bank, Los Angeles, Calif. The postmortem specimens we received had been obtained with authorization from the next of kin or an appropriate legal representative for use for research proposes. The ethical (Helsinki) committee at the Multiple Sclerosis Center of Sheba Medical Center does not require its approval for studies involving pathologic specimens.

MR images were acquired with a 2-T imager (Prestige; Elscint, Haifa, Israel) by using a clinically existing T1-weighted protocol (repetition time msec/echo time msec of 550/20, field of view of 24 cm, and matrix size of 256 × 256 pixels). Images were acquired in the coronal plane, with a 2-mm section thickness and image resolution to avoid distortions in the estimated 3D surface. Therefore, n was set to be equal to the number of coronal contours when two or more transverse contours were used for the approximation. A value of n = 2 was selected when only one coronal contour was used for the approximation to prevent the final 3D shape from becoming spherical.

As a result of the uneven distribution of the data points and the lack of information about the lesion surface in the intersection space, the choice of a higher n value can cause distortions to the final surface. By using the sets of points expressed in spherical coordinates, the corresponding SH functions Ylm (Eq [A1] in Appendix A) were calculated to estimate the coefficients rlm (Eq [A2] in Appendix A), which define the estimated regression radii R (Eq [A2] in Appendix A). A new set of points that covers the entire 3D space (ie, θ in the range 0–π and φ in the range 0–2π) was obtained, and the corresponding SH functions were again estimated. Finally, R was recalculated by using the previously estimated coeffi-
Radiology

Pathologic Findings

After MR imaging, the brain halves were cut into 5-mm-thick coronal slices (10). Each pathologic slice was independently examined macroscopically for the presence of MS lesions by two pathologists (B.S. and D.N., with 5 and 15 years of experience in brain pathology, respectively). The dimensions of each lesion were measured, particularly focusing on the thickness, by obtaining 2-mm consecutive slices at the lesion area. From each selected region, a tissue block was fixed in 4% formalin and embedded in paraffin. Five microscopic millimeter-thick sections of each block were stained with hematoxylin-eosin, luxol fast blue, and Bielschowsky stains to ascertain the presence of MS lesions (Fig 2).

Photographs of the brain slices (original magnification, ×40) were scanned into a computer. The area of each lesion was determined by contouring the edge of the lesions by using mouse tracking (D.G.Z.) and by applying Green’s theorem in the plane (6) (see mathematic details and reasons for using this theorem in Appendix B). Lesion volume was calculated by multiplying its area by the 2-mm thickness according to macroscopic examination. Pathologic findings were considered the reference standard conventionally used in image validation and comparison, as in other prospective multicenter clinical trials (11–13). The macroscopically detected lesions were then compared with the corresponding MR imaging sections to match as many lesions as possible.

MR Imaging Features versus Pathologic Findings

To match each of the identified lesions at MR imaging with its corresponding pathologic lesion, the following five-step procedure was implemented: (a) For the sets of MR images, the most anterior and the most posterior sections were considered to be the first and the last sections, respectively; (b) the approximated location of each lesion at MR imaging was determined by taking the number of the MR sections in which each lesion appeared and multiplying it by the section thickness; (c) the brains were cut in an anterior-posterior direction (B.S.); (d) the pathologic lesions were sought one at a time in the four pathologic slices closer to their corresponding location determined in the second step; and (e) correspondence between the pathologically identified lesions and the lesions identified on MR images was determined by three readers (D.G.Z., A.A., and B.S. [who had 5 years of experience in brain pathology]) by means of consensus. The readers compared each pathologic slice in which lesions were identified with the corresponding MR imaging sections.

Hypotheses Tested

Two main hypotheses were tested (D.G.Z., K.H.Z.): first, that the SH method would show improvement over the CA method in 3D volumetric measurement, as correlated with the pathologic reference standard; and second, that the SH method would follow closely the shape of the image-segmented contour.

Detailed statistical methods are elaborated on later for each of these two hypotheses. Analytic and statistical software programs used included Matlab 6.0 (www.mathworks.com), S-Plus 6 (www.insightful.com), and Excel 2000 (www.microsoft.com).

Statistical Volumetric Comparisons: SH and CA versus Pathologic Findings

1. For each of the volumetric variables, summary statistics, including range, median, mean, and standard deviations, were computed. Furthermore, a z test of normality of the distributions of the data was conducted for each continuous variable (14,15). Thus, we chose the appropriate subsequent statistical hypothesis tests, either by means of ln (a natural log with a base of e) transformation of the data for symmetrization and normalization or on the basis of nonparametric statistical tools. The ln transformation is necessary because the distribution of volume tends to be skewed and have a restricted range. It is often employed in statistical analysis (17). In addition, lesion-specific percentage bias (B) of each approximation method was determined by standardizing the approximated volume (V_{SH} and V_{CA}) against the pathologic volume (V_{Path}): R(V_{SH}) = (V_{SH} − V_{Path})/V_{Path} × 100%, and similarly to compute R(V_{CA}) for the CA method. Furthermore, absolute biases were also calculated.

2. Because the volumetric distribution (in cubic millimeters) is nonnormal, nonparametric methods and Spearman rank correlation coefficient were used to evaluate the correlation between the volumetric data of each approximation method and the pathologic reference standard (16,17). The Wilcoxon signed-rank test for paired data was conducted to compare the underlying median biases (16).
3. To compare the underlying means of the percentage biases between the two approximation methods, two-way analysis of variance was performed, where the two factors were the lesions versus the methods. The equivalence of variances of the percentage biases by these methods was also tested by using an F test. Similar analyses were conducted for comparing the means and variances of the absolute biases after log transformation. Because absolute biases were often not normally distributed, these hypotheses tests were conducted on the absolute biases under a log transformation, with normality verified by using a z test of normality (14,15).

4. To visually compare the magnitude of the standardized absolute biases of SH versus CA, a scatterplot was created for all lesions, along with a 45° line through the origin with a slope of 1. The Spearman rank correlation coefficient of these standardized absolute biases between these two methods was computed.

5. The coefficient of variation, calculated by dividing the standard deviation by the mean and then multiplying the result by 100%, was compared for the absolute bias of each method.

### Statistical Shape Representation of SH against Segmented Contour

1. The reconstructed surface lying close to the segmented contour in the same coordinate system was recut along the coronal plane into discrete slices. A shape comparison between the segmented contour and its corresponding contours from the recut surface was performed.

2. For each lesion, the shape representation of the SH method was assessed against the segmented contour contour-wise. Euclidean distance was computed between these two corresponding points of the contours. Furthermore, because of the size variability, we only reported percentage normalized distance according to area of the corresponding segmented contour. Lesion-specific minimum, mean, and maximum values of the mean normalized distance were graphed for all lesions. The hypothesis that the median distance is 0 was tested by using a one-sided Wilcoxon rank-sum test, and the correlation between the distance and the pathologic volume of the lesions was assessed by using a Spearman rank correlation coefficient.

3. Because the traditional distance-based metrics might be influenced by variable underlying lesion volume, Dice similarity coefficient (DSC) was also used as an additional validation metric of spatial shape representation according to SH against segmented contour (SC), where DSC is defined as $DSC = \frac{2(SH \cap SC)}{SH + SC}$, where $\cap$ represents intersection. The value of this coefficient ranges from 0 to 1, indicating no overlap to complete overlap (18,19). A DSC of 70% was interpreted as satisfactory spatial overlap according to the literature (19). Lesion-specific minimum, mean, and maximum DSC values were graphed for all lesions. The underlying mean of the logit-transformed mean DSCs was tested at 70% by using a Student t test after applying the z test of normality (14,15), where $logit(DSC) = \ln[DSC/(1 − DSC)]$. The correlation between DSC and the pathologic volume of the lesions was also computed by using a Spearman rank correlation coefficient.

### Results

#### MR Imaging Features versus Pathologic Findings

Nineteen lesions were identified macroscopically by the pathologists and were confirmed to be MS lesions. Most of them were chronic according to histologic examination. Of these 19 identified MS lesions, 11 (58%) were matched with the T1-weighted MR imaging sections. On in vivo T1-weighted MR images, MS lesions and cerebrospinal fluid appear as areas of hypointensity. In our study, because of the relaxation time shortening inherent with formalin fixation, the lesions and the cerebrospinal fluid appeared as areas of hyperintensity. These lesions were then represented by SH and CA methods, along with the segmented contour method. MR imaging–detectable abnormalities were neither visible nor palpable macroscopically in the remaining eight (42%) lesions.

### Statistical Volumetric Comparisons: SH, CA, and Pathologic Findings

The actual volumetric measurements and biases (in cubic millimeters) of the 11 lesions varied from lesion to lesion (Table 1, with the lesions listed in increasing order of the pathologic volumes). The summary statistics were the following, according to pathologic volume: range of 10.7–2409.2, with a median of 56.1; according to SH, range of 12.1–2132.5, with a median of 53.6; and according to CA, range of 15.7–2117.3 with a median of 51.8. The volume data were significantly nonnormal on the basis of the z test of normality ($P = .001$, $P = .002$, and $P < .001$, according to pathologic, SH, and CA volumes, respectively).

By using the ln transformation, we normalized the distribution of the data ($P = .19$, $P = .21$, and $P = .12$ for ln-transformed pathologic, SH, and CA volumes, respectively). After applying the ln transformation of the volume measurements, the two-way analysis of variance

---

**TABLE 1**

Volumetric Measurements and Biases of Matched Lesions according to SH and CA Methods

<table>
<thead>
<tr>
<th>Lesion No.</th>
<th>Pathologic Volume (mm$^3$)</th>
<th>SH Volume (mm$^3$)</th>
<th>Bias of SH (%)</th>
<th>CA Volume (mm$^3$)</th>
<th>Bias of CA (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10.7</td>
<td>12.1</td>
<td>13.1</td>
<td>15.7</td>
<td>46.7</td>
</tr>
<tr>
<td>2</td>
<td>33.2</td>
<td>34.5</td>
<td>3.9</td>
<td>39.1</td>
<td>17.8</td>
</tr>
<tr>
<td>3</td>
<td>34.4</td>
<td>36.6</td>
<td>6.4</td>
<td>36.0</td>
<td>4.7</td>
</tr>
<tr>
<td>4</td>
<td>37.6</td>
<td>38.7</td>
<td>2.9</td>
<td>40.4</td>
<td>7.4</td>
</tr>
<tr>
<td>5</td>
<td>40.0</td>
<td>45.1</td>
<td>12.8</td>
<td>46.6</td>
<td>16.5</td>
</tr>
<tr>
<td>6</td>
<td>56.1</td>
<td>53.3</td>
<td>−5.0</td>
<td>51.8</td>
<td>−7.7</td>
</tr>
<tr>
<td>7</td>
<td>116.6</td>
<td>123.8</td>
<td>6.2</td>
<td>125.9</td>
<td>8.0</td>
</tr>
<tr>
<td>8</td>
<td>214.4</td>
<td>225.3</td>
<td>5.1</td>
<td>230.0</td>
<td>7.3</td>
</tr>
<tr>
<td>9</td>
<td>335.9</td>
<td>385.0</td>
<td>14.6</td>
<td>388.2</td>
<td>−8.2</td>
</tr>
<tr>
<td>10</td>
<td>1512.0</td>
<td>1358.3</td>
<td>−10.2</td>
<td>1519.2</td>
<td>0.5</td>
</tr>
<tr>
<td>11</td>
<td>2409.2</td>
<td>2132.5</td>
<td>−12.2</td>
<td>2117.3</td>
<td>−12.1</td>
</tr>
</tbody>
</table>

Note.—Data are correlated with the pathologic standard of reference and are presented in increasing order of the pathologic volumes. Spearman rank correlation coefficient was $r = 1.00$ between SH and pathologic volumes and $r = 0.99$ between CA and pathologic volumes.
of 8.7% ± 4.1; for CA, range of 0.5%–46.7%, median of 8.0%, and mean of 12.4% ± 12.4.

Both SH and CA methods were shown to be unbiased: By testing the mean percentage biases at 0 for SH and CA methods separately, we found that neither method was biased statistically. In addition, the log-transformed absolute biases satisfied the normality assumption, as verified by the z test (P = .68 and .12 for SH and CA, respectively), and the mean absolute percentage biases were not statistically different from 0 for each method. However, results of the F test of the variances of the percentage biases were significantly smaller for SH than for CA (P = .04), as well as for the variances of the absolute percentage biases (P = .001), suggesting that SH was more robust for different lesions than was CA. The coefficient of variation of the logarithm transformed absolute bias was only 0.25 for SH, in comparison with 0.53 for CA.

The absolute percentage biases showed that eight (ie, lesions 1, 2, 4, 5, 7, 8, 10, and 11) of the 11 lesions (73%) had biases smaller for SH than for CA (P = .04), as well as for the variances of the absolute percentage biases (P = .001), suggesting that SH was more robust for different lesions than was CA. The coefficient of variation of the log-transformed absolute bias was only 0.25 for SH, in comparison with 0.53 for CA.

The absolute percentage biases showed that eight (ie, lesions 1, 2, 4, 5, 7, 8, 10, and 11) of the 11 lesions (73%) had biases smaller for SH than for CA, as these points lie below the 45° diagonal line (Fig 3). However, Spearman rank correlation was r = 0.33 between the paired standardized absolute biases of these two methods, suggesting a slightly positive correlation. In other words, while the bias in CA increases, this scatterplot suggests that the bias in SH increases only slightly, but overall, SH performed better than CA.

### Statistical Shape Representation of SH against Segmented Contour

The shape of each MR imaging–identified lesion was approximated by using SH. Figures 4 and 5 demonstrate two different examples of MS lesions, showing the brain slice (Figs 4a, 5a) on the corresponding MR imaging section in Figures 4b and 5b and the approximated 3D shape of the lesion according to SH in Figures 4c and 5c. Figures 4b and 5b are examples of periventricular white matter lesions that were extended into two and three MR imaging sections, respectively. Visual comparisons of each of the contours of these lesions outlined during the segmentation process with their corresponding contours extracted from the 3D SH-reconstructed shape are shown in Figures 4d, 4e, 5d, 5e, and 5f. The lesion shape in Figure 4c is similar to the classical elliptically shaped MS lesions, while the shape obtained in Figure 5c is similar to a pyramidal shape.

The results of the statistical comparisons are as follows:

1. The number of contours per lesion is given in Table 3. In Figure 6, the normalized distance measures (percentages), standardized according to the sum of pathologic and SH areas (analogous to the calculation of DSC described later), are displayed with regard to individual lesions in increasing order of pathologic volume. The overall range of the normalized distance was 0.1%–2.5%. The mean distance was slightly negatively correlated with lesion size (r = −0.38), suggesting that the normalized distance was smaller for larger lesions with a larger normalizing lesion area.

2. Figure 7 shows the DSC versus the individual lesions, also displayed in increasing order of pathologic volume. The overall range of the DSC was 75%–96%. The logit (a function of the ln) transformation significantly improved normality, as the z test of normality showed P = .004 and P = .61 before and after logit transformation, respectively.

The underlying mean of the logit-transformed DSC values was significantly above logit (0.70) (P < .001) but was not significantly correlated with lesion size (r = 0.08). Thus, on the basis of DSC as a validation metric, SH was robust with respect to lesion size.

### Discussion

In previous pathologic studies of MS lesions (20–24), investigators found that the number, pattern of distribution, and size of lesions varied considerably between and within patients; furthermore, lesions can occur in all areas of the central nervous system. In an extensive study of postmortem brains (21), it was established that lesions had different shapes—for example, spherical, egg-shaped, ring-shaped, or conical, with the long axis oriented in the
same direction as the long communicating venules. Polyhedral or bizarre ramifying shapes were also present (20). The spherical and elliptical shapes appear most frequently in the white matter and are attributed to a perivascular extension of the inflammatory and demyelinating process along straight venules (24).

Because MR imaging is an important tool for detecting and monitoring MS lesions, further assessment of the geometry of the lesions could contribute to better understanding of their evolution over time. Moreover, such information can assist examination of several clinical variables, such as rate of progression to disability and response to drug treatments. In a couple of studies (25,26), investigators have also underlined the importance of MR imaging in the diagnosis of suspected MS and in predicting the future conversion of the diagnosis to clinically definite MS.

It is apparent that although great improvement has been achieved so far, specific aspects related to MR imaging technology, section thickness, or variations in lesion load measurements are still not standardized, and thus, it is extremely difficult to compare findings between patients or between studies. The development of analytic methods, such as the SH technique for assessment of MS lesion volume and geometry, is crucial for future studies and will enable early diagnosis and sensitive monitoring of disease activity.

It was reported that lesion size and shape of different lesions within the same patient were often similar (3). However, this homogeneity was not observed in our study; instead, significantly different MS lesion patterns, with high variations in size and shape in the same patient, were demonstrated on the basis of two-way analysis of variance.

SH is an analytic modeling approach that provides accurate measurements of MS lesion volumes, yields 3D shape information (4), and enables monitoring of temporal changes in lesion geometry (5). In the present study, as well as in a previous study (4), on the basis of the fact that most MS lesions manifest a spherical or elliptical shape, the center of gravity of the data points was used as that of the spherical coordinate system. Thus, one and only one pair of values for the elevation and azimuth angles (θ and ϕ, respectively) can represent each point on the lesion’s surface. For more irregular lesions with a half-moon (known as Dawson fingers in MS), star-like, or other shape, this parameterization (ie, assigning a pair of θ and ϕ values to each point on the lesion’s surface) may be limited. This is due to the possibility that more than one pair of θ and ϕ values could represent some of the points on the lesion’s surface. For those cases, a more sophisticated procedure to parameterize the data is needed. Those complex problems are usually solved by flattening and conformal mapping to map the data to a unitary sphere (27–29).

In the present study, we have applied and validated this method on the basis of brain MS lesions with known pathologic findings used as the reference standard. Our investigation was mainly devoted to volume assessment, since the evaluation of shape parameters was investigated previously (5). Since the goal of the current investigation was not comparison of section thickness and partial volume averaging, we tried to get the same 2-mm thickness of the MR imaging sections for the pathologic examination, the findings of which were used as the reference standard. Because of the manual pathologic slicing process, however, the achieved slice thickness was inaccurate. We found that the pathologic lesions appeared in only one slice, and then we estimated their volumes according to the CA method after measuring the individual thickness for each particular pathologic slice. The decision to use CA for measuring the pathologic volumes of the lesions was based on a previous, more extended analysis (4), which shows that on average, CA is slightly better than SH when only one contour per lesion is available for estimating the volume. The fact that only one pathologic slice was available for estimating the volumetric reference standard could be a cause of measurement uncertainties. As stated earlier, however, for this kind of situation, where only one contour per lesion is available, CA yields a good approximation of the lesion volume (4).

Our volumetric analysis confirmed that SH closely correlated with the patho-
logic lesion volume with no significant bias. In addition, SH yielded significantly reduced variability in the bias with regard to CA ($P < 0.04$), even for significantly different lesion types and sizes ($P = .001$). The lesion volume assessed according to SH highly correlated with pathologic volume ($r = 1.00$).

The main difference in the proposed method and the previously used CA method was in the variances of the biases for the volume measurement in comparison to the pathologic volume. Thus, SH was found to be closer to histologic measurements. Even though the deviation from the pathologic volume seems normally distributed with all measurements, for small lesions (<50 mm$^3$), the volume measured according to SH or CA is consistently larger than that at pathologic examination. This may suggest that there was systematic measurement error in one or all lesions, which was beyond the scope of the correlation analysis when considering pathologic findings as the “practical” standard of reference by convention. However, the fact that the SH measurement was closer to the pathologic measurement might not have guaranteed that the SH was a better measurement than the CA method. One of the possible causes for small lesion volume at pathologic measurement may be shrinking of samples during fixation.

Next, in our shape analysis, both the distance and DSC as spatial validation metrics showed negligible shape differences between the image-segmented contours and the corresponding contours derived from SH. The range of the mean distance was only 0.1%–2.5%, and the range of the DSC was 75%–96%. The mean DSC was also statistically significantly above 0.7 ($P < .001$) (19).

Table 3. Ranges of Normalized Mean Distance and DSC Measured between Contours according to SH and Segmented Contours Methods

<table>
<thead>
<tr>
<th>Lesion No.</th>
<th>No. of Contours</th>
<th>Range of Normalized Mean Distance (%)</th>
<th>Range of DSC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.8–0.8</td>
<td>88.0–88.0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.4–0.4</td>
<td>93.9–93.9</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1.0–1.6</td>
<td>80.0–82.8</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>0.6–0.7</td>
<td>86.7–90.9</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>0.6–0.9</td>
<td>85.7–91.4</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>0.7–1.4</td>
<td>81.1–87.8</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>0.5–1.0</td>
<td>75.4–92.9</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>0.2–0.4</td>
<td>91.3–94.1</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>0.4–1.2</td>
<td>75.6–86.3</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>0.1–1.2</td>
<td>82.5–95.5</td>
</tr>
<tr>
<td>11</td>
<td>14</td>
<td>0.1–2.5</td>
<td>74.6–95.1</td>
</tr>
</tbody>
</table>

Note.—The mean DSC is significantly above 0.7 ($P < .001$).
were robust with respect to lesion size and shape. This study had several limitations. First, only 11 lesions were available in two brain halves with confirmed pathologic findings obtained externally for statistical validation purposes. This limited sample size was due to two possible factors—first, the relative inaccuracy of the slicing process, and second, the uncertainties introduced by the fact that only one pathologic slice per lesion was available. Next, there were the difficulties in corresponding all brain slices with a precisely corresponding MR image. These difficulties result from different imaging plane and thickness (30) and the lack of topographic landmarks (3). Better MR imaging techniques may yield improved visualization of the MS lesions; as a result of higher spatial resolution of images (eg, 1-mm section thickness or less) and more automated MR imaging segmentation techniques, MS lesion volume and shape will be better estimated by using SH. Another limitation was the existing MR imaging acquisition protocol used in our clinical practice, as alluded to earlier. T1-weighted multisection MR imaging was used in the study. In the future, a 3D acquisition protocol may be considered as an alternative, since thinner sections can be obtained, even with isotropic voxels, which allows a better quantification of volume and irregular shape of lesions. Finally, the voxel size was approximately 2 mm³, which may introduce additional uncertainty when dealing with small lesions. An attempt to identify such uncertainty was done by using computerized simulations of MS lesions (4), and additional attempts will be made in the future.

In summary, SH provided unbiased lesion volume and smaller variable biases, along with 3D shape information. The demonstration of its utility in the sense of providing diagnostic information may be considered to be a wider scope of evaluation, which is beyond the scope of this work. However, we suggest including SH as an additional tool for assessing MS lesions, which is beyond the scope of this work. Therefore, the researchers suggested including SH as an additional tool for assessing MS lesions, which is beyond the scope of this work.

### Appendix A

SH are orthonormal functions over the unit sphere used to describe complicated surfaces in 3D space. These functions are defined by the following equation (31):

\[
Y_{lm}(\theta, \varphi) = \frac{2l+1}{4\pi} P_l^m(\cos \theta) \exp(im\varphi),
\]

where \( Y_{lm}(\theta, \varphi) \) is the corresponding SH function defined in a spherical coordinate system \((R, \theta, \varphi)\), \( P_l^m(\cos \theta) \) is the associated Legendre polynomial, the parameter \( l \) (function’s degree) is zero or a positive integer, and the integer \( m \) (function’s order) can have only the values of \(-l, \ldots, -(l-1), \ldots, 0, \ldots, (l-1), l\).

For a spherical coordinate system \((r, \theta, \phi)\), the surface radii \( R \) can be presented discretely with the following equation:

\[
R(\theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} r_{lm} Y_{lm}(\theta, \varphi),
\]

where \( r_{lm} \) are the amplitudes of the corresponding SH functions (represented by complex numbers, with \( l \) running from 0 to the selected number of harmonics “\( n \)”).

Equation (A2) can be represented as a multiple linear regression model. Thus, by using a least squares approximation, an MS lesion can be described analytically by calculating the corresponding \( r_{lm} \) coefficients. The term \( r_{00} \) is the general average radius of the approximated surface. Hence, \( r_{00} \) represents the size or volume of the shape. Unlike \( r_{00} \), which represents a global measure (ie, size or volume), the terms \( r_{lm} Y_{lm}(\theta, \varphi) \) contain more local information of the 3D SH surface shape.

### Appendix B

A cross-sectional area bounded by a closed contour can be estimated by Green’s theorem in the plane:

\[
\text{Area} = \frac{1}{2} \int \text{xdy} - y\text{dx}.
\]

The integral can be calculated by integrating over each contour segment and summing the results. By using a summation notation and substituting \( \Delta x = x_i - x_{i-1} \) for \( dx \) and \( \Delta y = y_i - y_{i-1} \) for \( dy \), the formula is given thus:

\[
\text{Area} = \frac{1}{2} \sum_{i=1}^{N} [x_i(y_i - y_{i-1}) - y_i(x_i - x_{i-1})],
\]
where $N$ is the number of points in the contour, $x_i$ and $y_i$ are the coordinates of each contour point, and $x_0$ and $y_0$ are also the coordinates of the last contour points $x_N$ and $y_N$, respectively.

Acknowledgments: The authors express their deep gratitude to the people of the Multiple Sclerosis Human Neurospecimen Bank, Los Angeles, Calif, for providing us with the in vitro brains, as well as to Rachel Moscovitz, from the MRI Unit, Sheba Medical Center, Tel Hashomer, Israel, for her outstanding technical assistance.

References
New Suction Guide Needle Designed to Reduce the Incidence of Biopsy-related Pneumothorax: Experimental Evaluation in Canine Model

In an attempt to remove air that enters the pleural space during computed tomography (CT)-guided coaxial transthoracic needle biopsy, the authors fashioned an 18-gauge experimental suction guide needle and evaluated the incidence of pneumothorax with this needle in comparison to the incidence of pneumothorax with a standard 18-gauge guide needle in a canine model. This experiment had animal care and use committee approval. Ten dogs underwent a biopsy of each lung, for a total of 20 lung biopsies. Half of the biopsies were performed by using the experimental needle (five right lungs, five left lungs), and half were performed by using a standard guide needle. CT revealed pneumothorax during the procedure and was performed to reveal pneumothorax 1 and 3 hours after the procedure. A significant reduction ($P < .016$) in intraprocedural lung biopsy–associated pneumothorax was found when the experimental guide needle was used.

Transthoracic needle biopsy (TNB) became popular after Nordenstrom introduced the fine-needle aspiration technique in 1965 (1). Wallace et al (2) have recently reviewed the literature and found that this percutaneous technique has reported diagnostic accuracy rates in excess of 93% and sensitivity rates in excess of 95%. Because of its high diagnostic accuracy and less-invasive nature compared with surgery, TNB has become the procedure of choice for the diagnosis of pulmonary lesions. In addition, some thoracic surgeons advocate the use of TNB prior to mediastinoscopy for tissue sampling of mediastinal lesions (3).

Pneumothorax remains the most common complication of TNB and has been reported to occur in 0%–61% of procedures, with 1.6%–17.0% of such pneumothoraces requiring chest tube treatment (4). Pneumothoraces are rarely life threatening, but their treatment, with either close observation or chest tube drainage, exposes patients to increased morbidity and expense. Several factors thought to increase the risk of a biopsy-related pneumothorax include large needle size, increased number of pleural punctures, increased needle dwell time, obstructive lung disease, small lesion size, and increased skin-to-lesion depth (2,5–7). Although some of these factors are controllable, there has been conflicting evidence regarding the importance of many of these factors in clinical trials. However, several reports have consistently shown that obstructive lung disease significantly increases the risk that a symptomatic pneumothorax that requires chest tube treatment will occur (8–10). Unfortunately, this factor is independent of operator or patient control during TNB.

Because many complicated mechanisms affect the risk that a biopsy-related pneumothorax will occur, the incidence of this complication has not dramatically changed over the years. Two strategies for decreasing the chances of pneumothorax and chest tube placement are dependent
biopsy site positioning and injection of a blood patch or fibrin product in the needle path to help seal the pleural puncture site (11,12). Although some authors have reported success with these techniques, they are not routinely used owing to the inconsistency of their effectiveness. The exact mechanisms and numerous factors that affect the incidence of pneumothorax may never be completely understood. It is known that the space between the parietal and visceral pleura maintains a negative pressure that helps the lung stay fully inflated. We hypothesized that the incidence of biopsy-related pneumothorax should be lower if positive-pressure air does not enter the pleural space and disrupt this natural vacuum. Thus, the purpose of our study was to evaluate, in a canine model, the incidence of pneumothorax when an experimental suction guide needle designed to remove air that enters the pleural space is used during TNB.

1 Materials and Methods

Animals and Sedation Protocol

Our institution’s Animal Care and Use Committee approved this canine experiment after ensuring it would be conducted according to the guidelines and requirements set forth by the National Institutes of Health Public Health Service Policy on Humane Care and Use of Laboratory Animals (13) and the U.S. Department of Health and Human Services Guide for the Care and Use of Laboratory Animals (14). Ten large dogs (The University of Texas-Houston Medical School, Houston, Tex) with a mean body weight of 30 kg were used in this protocol because results of our preliminary experiments with computed tomography (CT)-guided lung biopsies (conducted by using a similar protocol) indicated that the lungs of large dogs better tolerated an indwelling 18-gauge needle than did the lungs of pigs or smaller animals (eg, rabbits). These preliminary feasibility experiments in rabbit and pig lungs consistently resulted in severe pneumothoraces that could not be relieved with suction.

Veterinary personnel performed all animal care and sedation procedures under the direct supervision of a veterinarian (T.M.L.). After a peripheral leg intravenous catheter was placed, sedation was induced with a 0.75 mg/m² intravenous bolus of medetomidine (Domitor; Pfizer Animal Health, Eaton, Pa) and an intramuscular injection of 0.04 mg of atropine (Atropine Sulfate; American Pharcaceutical Partners, Schaumburg, Ill) per kilogram of body weight. To duplicate the clinical scenario of a human lung biopsy and to avoid pulmonary pressure changes that might affect the natural occurrence of a biopsy-related pneumothorax, we used neither endotracheal intubation nor mechanical ventilation.

Sedation was maintained during the procedure with mask delivery of 1.5% isoflurane (IsoFlo; Abbott Laboratories, Chicago, Ill) in oxygen that was delivered at a rate of 5 mL per pound per minute. The effect of the medetomidine was reversed at the end of the procedure with a 3.0-mg intravenous bolus of atipamezole (Antisedan; Pfizer Animal Health). In accordance with the instructions on the atipamezole label, the volume of atipamezole that was introduced by intramuscular injection was identical to the previously administered volume of intravenous medetomidine. After the biopsy, all dogs were monitored for sedation recovery and potential biopsy complications.

Experimental Suction Guide Needle and Biopsy Technique

Each of the 10 dogs underwent one biopsy procedure in each lung. In the first half of the experiment, one lung was sampled for biopsy in each dog, and, after a 2-week recovery, the contralateral lung was sampled for biopsy in each dog in the second half of the experiment. Throughout the experiment, no more than two dogs underwent biopsy in 1 day. For half of the total number of procedures (n = 10 [five right lungs and five left lungs]), the new suction guide needle was used. This needle was created by first cutting several side holes that communicated with the inner needle lumen along the shaft of a standard 5-cm-long 18-gauge guide needle (Fig 1). The side holes were cut in a spiral fashion along the needle shaft so as not to weaken the needle and cause inadvertent fracture. A Tuohy-Borst type adaptor with a side port was attached to the guide needle hub by means of the Luer lock end of the adaptor (Boston Scientific, Natick, Mass) (Fig 2), and its side port was attached to standard wall suction set at 100 mm Hg of suction. A needle stylet for a 10-cm-long 19-gauge Chiba needle (Cook) was then cut to fit the length of the 5-cm-long 18-gauge needle plus the length of the attached valve. A 19-gauge stylet was used because it is smaller in diameter than an 18-gauge stylet, and air entering the needle could travel through the needle lumen around the stylet and out the valve side port to the wall suction container. This system maintained a vacuum once the guide needle side holes were in the thorax. The valve was opened for placement of the coaxial biopsy needles and then closed around them to maintain the suction during tissue sampling.

Biopsy Procedure and Imaging

All CT images were reviewed by the same radiologist (F.A.M.), who had 9 years of experience reviewing chest CT images. For this protocol, a pneumothorax was defined as the presence of air between the lung and chest wall as seen at transverse CT. The degree of pneumothorax was not evaluated for this introductory experiment, as our purpose was to observe the presence or absence rather than the severity of a pneumothorax. However, CT images were evaluated to determine if any pneumothorax was
greater than 50% of the estimated total hemithoracic volume.
All fabrication and assembly of experimental needle systems, as well as the use of these systems during TNB, was performed by the same operator (F.A.M.), who had 8 years of experience in performing TNB. During each lung biopsy, two pleural punctures were performed and two fine-needle aspirate samples and two core samples were obtained. For the first set of biopsies (n = 10 [one lung in each dog]), the use of the experimental suction guide needle or the standard guide needle was alternated, as was the choice of right or left lung (five right lungs, five left lungs).

After sedation was achieved, each dog was placed in a decubitus position with the biopsy lung side up on the table of a HiSpeed Advantage CT scanner (GE Medical Systems, Milwaukee, Wis). A preliminary nonenhanced scan through the thorax was obtained by using 80 kVp, 100 mA, and 5-mm section increments. A lower lobe lung access site was marked, and the skin was shaved and prepared in a sterile fashion. After the guide needle was placed in the lower lung lobe, an aspirate sample was obtained with a coaxial 22-gauge needle, and a core tissue sample was obtained with a coaxial 20-gauge core needle; interval CT scans were obtained to document pneumothorax. The guide needle was then removed and replaced in the lower lobe, creating a second pleural puncture. The fine-needle aspiration and core sampling were repeated, and all needles were removed. CT was performed immediately after the procedure, and each dog was sedated again and rescanned 1 and 3 hours after each biopsy for detection of delayed pneumothorax.

Two weeks later, the contralateral lungs were sampled for biopsy with this same protocol. For the contralateral lung biopsies, the guide needle selected was the one that had not been previously used in that particular dog. Thus, a total of 20 lung biopsy experiments were performed. Both the experimental and the standard guide needles were used in each dog, and each needle was used in an equal number of right and left lungs.

Statistical Analysis

The McNemar test was used to compare the proportion of intraprocedural pneumothoraces between the set of lungs sampled for biopsy with the suction guide needle and the set of lungs sampled for biopsy with the standard guide needle. A similar comparison was made for the proportion of overall pneumothoraces between the suction and standard guide needle groups. All reported P values are two-sided at a significance level of 5%. Analyses were performed by using SAS software version 8.2 (SAS Institute, Cary, NC).

Assuming a two-sided significance level of 5%, a sample of 10 dogs (10 lung pairs) yields 98% power to detect a 70% difference in proportions of intraprocedural pneumothorax when the proportion of discordant pairs is approximately the same and the method of analysis is a McNemar test of equality of paired proportions.

To predict how many more biopsies would be necessary to indicate statistical significance for overall pneumothorax reduction, we performed the McNemar test of equality of paired proportions with a two-sided significance level of 5%. Results of this test indicated that a sample size of 19 pairs would have 80% power to reveal a difference in proportions of 0.500 when the proportion of discordant pairs was expected to be 0.700. The small cohort in this protocol precluded analysis of significant differences between right and left lungs or the significance of the timing of an intraprocedural pneumothorax as related to number of pleural punctures performed or number of aspirate or core samples obtained.

Results

Complications

No mechanical complications were encountered. There were no animal complications other than pneumothorax. None of the pneumothoraces were more than 50% of the estimated total hemithoracic volume as seen at transverse CT. No dog with a pneumothorax showed obvious respiratory signs that necessitated additional therapy. Two weeks elapsed between right and left lung biopsies in each dog, and no prior pneumothoraces were detected at the time of the contralateral lung biopsy.

Pneumothorax Incidence

For the 10 biopsies performed by using the experimental suction guide needle (Fig 3), four (40%) pneumothoraces occurred. Two (20%) of these occurred during the procedure, and two (20%) were detected at postprocedure CT imaging. Nine (90%) pneumothoraces occurred in the 10 biopsies performed by using the standard guide needle (Fig 4); all of these occurred during the biopsy procedure. The Table shows the distribution of biopsies performed by using the two guide needles and the incidence of pneumothoraces.

Results of McNemar testing indicated that there was no significant difference in the overall rate of pneumothorax between the experimental suction needle and the standard guide needle (P = .125). However, when only intraprocedural pneumothoraces were considered, a significant difference was found between the two types of needles (P < .016), with the suction guide needle producing significantly fewer pneumothoraces during the biopsy procedure.

Discussion

Biopsy-related pneumothorax remains a complication that generally cannot be
predicted or prevented despite efforts to alleviate it through procedural adjustments and postbiopsy maneuvers. Treatment is usually observation, but sometimes insertion of a chest tube is required to relieve the pulmonary-related symptoms. Because many factors affect the incidence of pneumothorax, including ones that are uncontrollable, it seems logical to concentrate more on removing the positive-pressure air and maintaining the pleural vacuum. It has previously been shown that manual aspiration of a pneumothorax that occurs during a lung biopsy might prevent progression and the subsequent need for chest tube placement (15). Because the suction guide needle produces constant aspiration during a biopsy procedure, the significant reduction in the incidence of pneumothorax that we observed with this technique supports the results of that earlier report. This canine experiment was the first step in a process intended to reproduce similar results in human percutaneous lung biopsies.

Careful attention was devoted to reproducing a clinical situation similar to that of a human lung biopsy performed with conscious sedation. With our canine sedation protocol, we eliminated pulmonary pressure changes associated with endotracheal intubation. Despite this, our pneumothorax rate of 90% with the standard guide needle is believed to be too high compared with the previously cited rate of human lung biopsy-associated pneumothorax (0%–61% [4]). We believe that this difference underscores inherent differences between canines and humans. One such difference is that the hemithoraces in the dog communicate, whereas those in humans do not. This communication increases the risk of a bilateral pneumothorax during a unilateral lung biopsy. In fact, Wittich et al (16) reported that such a communication occurred in a patient who underwent cardiothoracic surgery that created a common pleural cavity.

Although we did not observe bilateral pneumothoraces in a single lung biopsy in our present study, during preliminary studies to determine the most appropriate species to use as an experimental model, a severe unilateral pneumothorax in a dog produced a contralateral pneumothorax, and neither pneumothorax responded to suction. This communication creates an area of uncertainty regarding canine pulmonary physiology and affected the adjustment of several factors in our protocol. For example, normal negative pleural pressure in the canine thorax is difficult to predict. For this reason, we chose to set the wall suction at a maximum pressure of 100 mm Hg. We observed no clinical or imaging evidence that this degree of suction was excessive or damaging to the lung parenchyma. If a lesser degree of suction had been chosen, this choice would have been entirely arbitrary, and we could not have determined whether the occurrence of a pneumothorax during a lung biopsy was due to insufficient suction. However, it is uncertain whether the use of a greater degree of suction than 100 mm Hg would reduce the pneumothorax incidence even more. This will need further investigation, and our better understanding of human pulmonary physiology may lead to improved adjustments and better results in human clinical trials.

Several types of needles have previously been studied in lung biopsies (17). Although many needle types are used to obtain various tissues for diagnosis, no specific needle type has been statistically proved to significantly reduce the incidence of pneumothorax. The results of the present experiment are promising because we significantly reduced the incidence of intraprocedural pneumothoraces by using the suction guide needle. In addition to this, there are potential benefits of the suction guide needle that have not yet been investigated.

In a clinical setting, reducing intraprocedural pneumothoraces could lead to better sampling accuracy for difficult lesions. If a pneumothorax occurs during a biopsy procedure and the lesion falls away from the needle as a result, diagnostic accuracy may be compromised. Keeping the lung fully inflated during the procedure by using a suction needle may remedy this sampling difficulty.

Pneumothoraces were not graded in this protocol, nor was an attempt made to evaluate the development of a pneumothorax related to the number of pleural punctures, aspiration versus core sampling, or the timing of a delayed occurrence at 1 versus 3 hours. The significance of both the number of pleural punctures and the acquisition of an aspirate versus a core tissue sample has been studied in human lung biopsies (2,5–7). Such study results have not been consistent with the importance of these factors, but this also does not mean that uncontrolled differences in these parameters do not affect the results. However, we found that pneumothoraces occurred even with the use of the suction guide needle. This underscores the multifactorial and complicated etiology of a pneumothorax. We speculate that such factors do have an important role in the development of a lung biopsy–associated pneumothorax. With the limiting factors of a small cohort, a fixed number of pleural punctures, and a fixed number of biopsy passes in each experiment, our analysis could not reveal the statistical significance of these limiting factors. Because our protocol did not allow for the placement of chest tubes, grading the severity of a pneumothorax was not necessary. Because in humans the number of pleural punctures varies and the severity of a pneumothorax

<table>
<thead>
<tr>
<th>Lung Biopsy Details and Incidence of Pneumothorax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dog No.</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>10</td>
</tr>
</tbody>
</table>

Note.—ND = no data (no further data were collected).
affects the placement of a chest tube, these issues would be better evaluated in human clinical trials.

A number of limitations affected our data analysis. The operator was not blinded to the type of needle used. Although we do not believe this attributed to a major difference in technique, it does introduce bias. A further selection bias involved the nonrandom selection of right and left lungs, as well as the needle choice for each procedure. Finally, the degree of pneumothorax was not quantified in this experiment. For the purpose of this initial study, we chose to concentrate on the presence or absence of pneumothorax. Of course, in human lung biopsies, the size of the pneumothorax and the presence of tension, not simply the presence of a small pneumothorax, are main causes of morbidity. On the other hand, operators who perform lung biopsies with real-time fluoroscopic guidance may observe small pneumothoraces that prove to be inconsequential. Certainly, the degree of pneumothorax is important to evaluate. This requires further study in subsequent experiments.

Our initial positive results with this experimental guide needle have prompted an application for a provisional patent and permission to undertake the steps required to create a prototype for use in human clinical trials.

Acknowledgment: We thank Marcella M. Johnson, MS, for her contribution to the statistical analysis of the experiment results.

References
ATTENUATION OF ACUTE AND CHRONIC PULMONARY EMBOLI

PULVOSE: To compare retrospectively the attenuation coefficients of acute and chronic pulmonary embolism (PE).

MATERIALS AND METHODS: Institutional review board approval was obtained, and informed consent was waived. The study was compliant with requirements of the Health Insurance Portability and Accountability Act. All patients with chronic PE, from July 2001 to January 2004, were identified via a radiology report search system; of the 39 identified, 25 were excluded because the thrombi were too small to measure or were obscured by streak artifacts or because there was no corroborative evidence of chronic PE. Of 27 consecutive patients with acute PE who were also identified, two were excluded because of streak artifacts. The final study group included six women and eight men with chronic PE (mean age, 50 years; range, 26–76 years) and 11 women and 14 men with acute PE (mean age, 61 years; range, 33–83 years) (*P* < .01 for age). Images were acquired with a four-detector row computed tomographic scanner and 1.25-mm collimation. Two readers made independent attenuation measurements of the largest thrombus in each patient at a workstation. Statistical analysis included calculation of means and standard deviations, the *t* test, and the Bland-Altman test.

RESULTS: Reader 1 found mean attenuation of 90 HU ± 30 (range, 54–155 HU) for chronic PE and 33 HU ± 15 (range, 6–63 HU) for acute PE (*P* < .001). Reader 2 found mean attenuation measurements of 83 HU ± 32 (range, 32–135 HU) for chronic PE and 33 HU ± 14 (range, 13–65 HU) for acute PE (*P* < .001). The mean attenuation for both readers was 33 HU for acute PE (95% confidence interval: 26, 41 HU) and 87 HU for chronic PE (95% confidence interval: 66, 107 HU). The Bland-Altman test demonstrated agreement between readers.

CONCLUSION: The mean attenuation measurement in chronic PE is significantly higher than in acute PE.

© RSNA, 2005

Pulmonary embolism (PE) is the third most common acute cardiovascular disease, after myocardial infarction and stroke, and it leads to thousands of deaths each year because it often goes undetected (1). In a study performed to evaluate trends of inpatient thoracic radiology use over a decade at an academic medical center, Wittram et al (2) showed that use of computed tomography (CT) in the investigation of patients suspected of having thromboembolic disease has increased considerably (2). Although the use of ventilation-perfusion scintigraphy and pulmonary angiography has decreased, these imaging modalities still have roles in the evaluation of patients in whom PE is suspected.

Acute PE is often a complication of deep leg or pelvic vein thrombosis. Fragmentation and fibrinolysis cause dissolution of the vast majority of acute PE. Chronic PE persists in a number of patients, and it has been suggested that chronic PE results from individual variations in the fibrinolytic activity of pulmonary vessels (3). To our knowledge, the attenuation of chronic PE has not been measured. This information may help in differentiating acute from chronic thrombus at first encounter if old CT studies are not available. Thus, the purpose of our study was to compare retrospectively the attenuation coefficients of acute and chronic PE.
MATERIALS AND METHODS

Patients

Our institutional review board approved this retrospective study and waived informed consent. Our study was compliant with requirements of the Health Insurance Portability and Accountability Act. Between July 2001 and January 2004, all patients with a diagnosis of chronic PE or acute PE at CT pulmonary angiography were identified from our radiology information database search program (Folio, Woburn, Mass). The patients’ charts were collected consecutively on the basis of the report text, and they were reviewed by the first author (C.W.) for corroborative evidence (imaging or clinical data) to confirm the diagnosis of PE, which was assigned at CT. The corroborative criteria for chronic PE used in this study included acute PE at the same site for more than 3 months or chronic PE documented with CT or angiographic findings, a previous medical history of chronic PE, or a clinical history consistent with pulmonary arterial hypertension. Corroborative evidence of acute PE included no history of prior PE, with previous imaging findings—when available—negative for thromboembolism.

Thirty-nine patients with chronic PE were identified; 20 were excluded because the lesions were too small to measure (ie, not present on three consecutive images or too small to allow the area of the region of interest [ROI] to be half the diameter of the thrombus being measured), two were excluded due to streak artifacts, and three were excluded because there were no imaging findings or history to confirm a diagnosis of chronic PE. Thus, 14 patients with chronic PE were available for our study (six women, eight men; age range, 26–76 years; mean age, 50 years). Twenty-seven consecutive patients with acute PE were identified, but two were excluded because of streak artifacts, leaving 25 patients with acute PE in our study (11 women, 14 men; age range, 33–83 years; mean age, 61 years).

The two-tailed t test was performed with SAS software (version 8; SAS Institute, Cary, NC) and demonstrated a significant difference in age (P = .01) but not sex between the two groups.

CT Technique

A multi–detector row CT scanner (four rows of detectors) (LightSpeed; GE Medical Systems, Milwaukee, Wis) was used to acquire the images. Images were acquired from 2 cm below the lowest hemidiaphragm to the top of the aortic arch, in a caudocranial direction. The imaging field of view used was the widest rib-to-rib distance acquired in apnea after inspiration. For intravenous access, an antecubital vein was used, with 135 mL of ioxilan (300 mg iodine per milliliter) injected through an 18-gauge catheter at a rate of 4 mL/sec. Images were acquired with the following settings: collimation and reconstruction width, 1.25 mm; table speed, 7.5 mm per rotation; pitch, 6:1; peak voltage, 120 kVp; amperage, 300 mA; and tube rotation time, 0.8 second. Image acquisition was started 20 seconds after commencement of the intravenous contrast medium injection, and a standard algorithm was used.

Diagnostic Criteria of PE

Both acute and chronic PE were identified as intraluminal filling defects that demonstrated a sharp interface with intravascular contrast material (C.W.). The diagnostic criteria for acute PE included (a) complete arterial occlusion with failure to opacify the entire lumen—the artery may be enlarged compared with pulmonary arteries of the same order of branching (4–6), (b) a central arterial filling defect surrounded by intravascular contrast material (4), and (c) a peripheral intraluminal filling defect that forms acute angles with the arterial wall (5,6).

The diagnostic criteria for chronic PE included (a) complete occlusion of a vessel smaller than pulmonary arteries of the same order of branching (5,6), (b) a peripheral eccentric filling defect that forms obtuse angles with the vessel wall (5,6), (c) contrast material flowing through apparently thick-walled arteries that are smaller due to recanalization (5,6), (d) a web or band within a contrast-material-filled artery (5,6), and (e) an intraluminal filling defect with morphologic features of acute PE, which has been present for more than 3 months.

Measurements

The images were viewed at a picture archiving and communication system, with an IMPAX monitor (version 4.1; Agfa, Teterboro, NJ). Images were displayed with three gray scales for interpretation of the lung window (window width, 1500 HU; window level, −600 HU), mediastinal window (window width, 350 HU; window level, 40 HU), and PE-specific settings (window width, 700 HU; window level, 100 HU).

Two radiologists (C.W. and M.M.M., with 8 and 2 years of experience, respectively, in thoracic radiology interpretation) independently measured attenuation in Hounsfield units. They identified the largest diameter thrombus for an individual patient and measured the ROI from the middle image of three that demonstrated the thrombus. The ROI diameter was selected to be half the diameter of the thrombus being measured and positioned so as to avoid streak artifacts.

Each reader independently selected the largest image of the main pulmonary artery, used a circular ROI approximately half the diameter of the main pulmonary artery—away from streak artifacts—and recorded the area of the ROI and the mean attenuation for the pulmonary artery. Other CT findings of PE were read by consensus. The additional findings of chronic PE included enlarged main pulmonary artery and collateral bronchial artery, mosaic lung attenuation, and pericardial effusion (7). The additional findings of acute PE included a right ventricle short axis larger than left ventricle short axis, consolidation or ground-glass opacification peripheral to PE, atelectasis, mosaic attenuation, and pleural effusion.

Statistical Analysis

Statistical analysis included the calculation of means and standard deviations, and χ2 and two-tailed t tests were performed with SAS software (version 8; SAS Institute). P values less than .05 were considered to indicate statistical significance. The Bland-Altman test was used to analyze the degree of measurement agreement between readers.

RESULTS

Chronic PE

Review of the charts of patients in the chronic PE group revealed previous imaging with CT pulmonary angiography in five of 14 patients and contrast material–enhanced CT with 5-mm collimation in two of 14 who demonstrated acute PE in the same region as those with chronic PE. The interval from previous CT to CT pulmonary angiography ranged from 3 months to 4 years 11 months (mean, 1 year 1 week). In one additional patient, chronic PE was diagnosed at CT pulmonary angiography 2½ months before CT pulmonary angiography was performed in this study. In two more patients, pulmonary angiograms that confirmed chronic PE were obtained 5 and 7 weeks after the study CT pulmonary angiograms were obtained. The remaining four of 14 patients underwent no prior or subsequent imaging
Chronic PE is characterized by complete occlusion of a vessel that is smaller than pulmonary arterial hypertension. The circular ROI measurement was 84 HU.

Acute PE

Review of the charts of patients in the acute PE group revealed no history of prior or recurrent PE. Twenty-one of 25 patients had undergone no previous imaging for thromboembolic disease. In two patients, however, negative CT pulmonary angiograms had been obtained 21 and 30 days before the study CT pulmonary angiograms were obtained. In one more patient, conventional contrast-enhanced CT findings were negative 15 days before positive CT pulmonary angiography findings were obtained in this study. The final patient underwent bilateral leg ultrasonography (US) 10 days before the positive CT pulmonary angiogram was obtained, and the US findings were negative for deep vein thrombosis.

The emboli in patients with chronic PE were identified on the right side within the apical segmental artery in the upper lobe in one patient, the upper lobe artery in one, the interlobar artery in two, the lower lobe pulmonary artery in three (Fig 1), the posterior basal segmental artery in the lower lobe in three (Fig 2), and the medial basal segmental artery in the lower lobe in one. One embolus was located in the left pulmonary artery, one in the interlobar artery, and another within the posterior basal segment arteries in the left lower lobe. Other findings of chronic thromboembolic disease included an enlarged main pulmonary artery in 10 patients, collateral bronchial artery enlargement in nine, mosaic lung attenuation in seven, and pericardial effusion in five.

DISCUSSION

CT pulmonary angiography has become the standard of care for the investigation of thromboembolic disease at some institutions (2). There are many reasons that PE is misdiagnosed at CT pulmonary angiography, and one is flow phenomena. For reviewing cases of suspected PE, modified PE-specific window settings are recommended to help identify small emboli that may be obscured by very bright vessel contrast (8). In our experience, we have found that PE-specific window settings improve the detection of small PE and increase the conspicuity of flow-related artifacts. To the untrained eye, some flow-related artifacts can be mistaken for acute PE. Flow artifacts are not sharply defined and might be identified more confidently if one knows the normal range of attenuation in patients with acute PE. Further imaging with repeated CT pulmonary angiography or conventional pulmonary angiography may be necessary to exclude thrombus hidden in poorly opacified vessels. This is because the detection of a low-contrast abnor-
Findings in Patients with Acute and Chronic Pulmonary Emboli

<table>
<thead>
<tr>
<th>Observer</th>
<th>Acute PE</th>
<th>Chronic PE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Thrombus</td>
<td>Main Pulmonary Artery</td>
</tr>
<tr>
<td>Mean Attenuation (HU)</td>
<td>Mean Area (cm²)</td>
<td>Mean Attenuation (HU)</td>
</tr>
<tr>
<td>Reader 1</td>
<td>33 ± 15</td>
<td>0.22 ± 0.27</td>
</tr>
<tr>
<td>Reader 2</td>
<td>33 ± 14</td>
<td>0.20 ± 0.26</td>
</tr>
<tr>
<td>Mean Attenuation (HU)</td>
<td>Mean Area (cm²)</td>
<td>Mean Attenuation (HU)</td>
</tr>
<tr>
<td>Reader 3</td>
<td>90 ± 30</td>
<td>0.08 ± 0.08</td>
</tr>
<tr>
<td>Reader 4</td>
<td>83 ± 32</td>
<td>0.10 ± 0.10</td>
</tr>
</tbody>
</table>

Note.—Data are mean ± standard deviation.

mality is not accurate when the standard deviation of the mean of the abnormality exceeds the difference in the means of the lesion and the surrounding region (9).

The morphologic outcome of acute PE has been previously evaluated (10). Remy-Jardin et al (10) reviewed findings in 62 patients who were referred to a cardiac intensive care unit with massive acute PE, received anticoagulation therapy, and underwent follow-up CT pulmonary angiography after a mean interval of 11 months. The CT pulmonary angiograms demonstrated that 30 patients (48%) had complete resolution of acute PE and 32 (52%) had incomplete resolution, with a significant difference in the severity of PE between the two groups.

The attenuation in patients with acute PE has been measured previously (11). By using 3-mm collimation CT pulmonary angiography, Cham et al (11) found a mean attenuation of 49 HU in patients with acute PE (95% CI: 44 HU, 34 HU). Unlike in that study, all our patients underwent imaging with 1.25-mm collimation. In addition, we used strict criteria to ensure that partial volume artifact did not affect measurements by using the middle image of three that demonstrated thrombus. In our study, we found a mean attenuation of 33 HU ± 15 in patients with acute PE.

To our knowledge, attenuation in patients with chronic PE has not been previously reported. We calculated the mean attenuation in patients with chronic PE to be 87 HU ± 30. This value was significantly higher than the mean for patients with acute PE ($P < .001$). Although there was some overlap between chronic and acute PE for individual results, this overlap was not reflected in the lower and upper 95% CIs for the two groups. The reasons for the higher attenuation in patients with chronic PE compared with those with acute PE are likely related to enhancement of organizing thrombus, retraction of thrombus with concentration of hemoglobin and its iron moiety, and possibly, calcium deposition. Paydarfar et al (12) used preoperative contrast-enhanced magnetic resonance (MR) imaging to evaluate 15 consecutive patients with intraluminal cardiac masses. Contrast enhancement on MR images was seen in three of the patients with subsequently proved organized thrombus (12). These findings support our observations and our hypothesis that enhancement is seen in patients with chronic PE.

Limitations of this study include the fact that the diagnoses of acute and chronic PE were not confirmed with the reference standard, pulmonary angiography, in the vast majority of cases. The reason is that in practice, within our institution, conventional angiography is used to confirm or refute PE in patients with indeterminate findings at CT pulmonary angiography, and both the acute and chronic PE groups in our study had imaging or clinical evidence to corroborate the CT findings. Also, in a study conducted to evaluate chronic PE detection, Bergin et al (13) demonstrated that CT pulmonary angiography was more accurate than angiography in depicting central vessel disease.

A second limitation of our study is that only one PE was measured per patient.
The study was designed this way, to optimize the ability to measure the thrombus by selecting the largest thrombus. A third limitation is the relatively small number of patients with chronic PE examined, an unavoidable consequence of applying our strict measurement criteria to these patients. A fourth limitation is that unenhanced CT was not performed just prior to contrast-enhanced CT. We assume that the higher attenuation values seen in patients with chronic PE are most likely due to enhancement of organizing thrombus; however, without unenhanced CT, it is impossible to be certain if they are due to hemoglobin-iron concentration from clot retraction or calcium deposition.

The finding of enhancement in patients with chronic PE raises the problem of differentiating between PE and an uncommon cause of an intraluminal arterial filling defect, a primary pulmonary artery sarcoma (14,15). Both chronic PE and pulmonary artery sarcoma can demonstrate contrast enhancement, although other findings may permit differentiation between the two entities. Signs that suggest a pulmonary artery sarcoma at CT pulmonary angiography include a lobulated mass that forms acute angles with its vessel wall, vascular distention, and local extravascular spread (14–16). Chronic PE often forms obtuse angles with the adjacent arterial wall (5,6). To differentiate between acute PE and pulmonary artery sarcoma, as both can form acute angles with the vessel wall, pulmonary artery sarcomas demonstrate enhancement (14,15), whereas acute PE does not. If uncertainty remains in the differentiation between PE and vascular tumor, angiographically guided biopsy may be necessary.

In summary, knowing the attenuation range for acute PE may be useful in the incremental evaluation of a pulmonary artery filling defect and may aid in differentiating between acute PE and flow artifact. This study demonstrates that chronic PE has a higher mean attenuation than acute PE, and we believe this difference is likely due to enhancement of organizing thrombus.

References
Airway Wall Thickness in Cigarette Smokers: Quantitative Thin-Section CT Assessment

**PURPOSE:** To design and validate a dedicated software tool to measure airway dimensions on thin-section computed tomographic (CT) images and to use the tool to prospectively compare airway wall thickness in nonsmokers with normal lung function with that in smokers with and without chronic obstructive pulmonary disease (COPD).

**MATERIALS AND METHODS:** All subjects gave written informed consent. The study was approved by local ethics committee. With Laplacian of Gaussian algorithm, software was tested in phantom and excised sheep lung fixed in inflation and validated with Bland-Altman analysis. Study prospectively included nine nonsmokers (six women, three men; mean age, 53 years ± 5.6 [standard error of the mean]) with normal lung function (group 1), seven smokers (three women, four men; mean age, 56 years ± 5.6) with normal lung function (group 2), and eight smokers (zero women, eight men; mean age, 65 years ± 4.0) with COPD. Calculations were determined with spirometrically gated CT: For each selected bronchus, the wall area (WA), internal area (IA), airway caliber (sum of IA and WA), and WA/IA ratio were calculated. For each patient, summation of WA to summation of IA (WA/IA) ratio, which reflected normalized airway wall thickness, was calculated. Groups were compared by using analysis of variance with generalized linear model and unpaired t test. Pearson correlation coefficient was used to assess correlation between software measurements and pulmonary function test results.

**RESULTS:** Comparison of measurements in phantom and excised sheep lung with algorithm measurements revealed that the latter were reliable and repeatable. In clinical study, WA/IA ratio was significantly different among three groups (P < .001). Normalized airway wall thickness and IA were significantly related to lung function test data, including forced expiratory volume in 1 second (FEV1) and forced expiratory flow between 25% and 75% of vital capacity (FEF25%–75%). Normalized airway wall thickness was larger in smokers with COPD than it was in smokers or nonsmokers without COPD.

**CONCLUSION:** This software provides accurate and reproducible measurements of IA and WA of bronchi on thin-section CT images and demonstrates that in vivo normalized airway wall thickness was larger in smokers with COPD than it was in smokers or nonsmokers without COPD.

© RSNA, 2005

Airway wall remodeling observed in chronic obstructive pulmonary disease (COPD) contributes to alteration in the function of the airways. By using a histologic approach, airflow obstruction has been shown to be caused by airway wall thickening and to be related to airway inflammation in patients with COPD (1). On the other hand, it has been shown that cigarette smoking is associated with an inflammatory process that involves both proximal and distal airways and lung parenchyma (2), even in patients who exhibit normal lung function (3). Although several authors observed differences in T lymphocyte bronchial infiltration when they compared smokers who developed COPD with smokers who did not, the characteristics of the inflammatory process remain largely unclear (1).
Thin-section computed tomography (CT) is increasingly used for the study of airway morphologic characteristics in vivo. Theoretically, thin-section CT can depict the dimensions of airways as narrow as approximately 1–2 mm in diameter. For clinical purposes, the use of thin-section CT for assessment of bronchial wall thickness in patients with airway diseases has been, so far, mainly subjective (4). Although subjective grading has led to new insights into the pathophysiologic features of various airway diseases, objective methods are suitable to perform longitudinal studies and to compare airway dimensions before and after bronchial provocation testing or therapeutic intervention in chronic diseases (5). Therefore, it has been suggested that CT be used to evaluate airway dimensions in patients with COPD (6), and quantitative algorithms for analysis of thin-section CT images have been designed and validated in studies with phantoms and animals (7). Although a large number of studies have been conducted to detect and to assess emphysema, for example, by using CT, very few have been devoted to measurement of airway dimensions in patients with COPD. Nakano and colleagues (8) evaluated the airway wall dimension at the origin of the apical bronchus of the upper lobe of the right lung in smokers. Since a single bronchus was measured in each patient, however, the heterogeneity of inflammation within the bronchial tree was not assessed (8). Little et al (9) measured airway caliber (sum of internal area [IA] and wall area [WA], hereafter referred to as IA + WA) in a variety of different bronchi in asthmatic patients, but their method was manual and, thus, observer dependent. IA is defined by the area limited by the internal border of a bronchus wall.

The aim of the present study, therefore, was to design and validate a dedicated software tool to measure airway dimensions on thin-section CT images and to use the tool to prospectively compare airway wall thickness in nonsmokers with normal lung function with that in smokers with and without COPD.

**MATERIALS AND METHODS**

**Study Design**

A dedicated software tool that was based on a Laplacian of Gaussian algorithm (Detection of Airway Contours with Laplacian of Gaussian Algorithm [DACLOG]) was developed to measure IA and airway WA of bronchi on two-dimensional thin-section CT images. DACLOG was tested in two ways: by using a phantom made with silicone tubes and by using an excited and fixed sheep lung. DACLOG was used to measure and to compare airway dimensions in a large group of airways in three groups of subjects: a group of smokers with COPD, a group of smokers without COPD (with normal lung function), and a control group (nonsmokers with normal lung function).

**Image Analysis Software: DACLOG**

A Laplacian of Gaussian algorithm was used to develop a software tool for extraction of bronchial geometric parameters defined as IA and WA from thin-section CT images. Laplacian filters are second-derivative filters used to find areas of rapid change (edges) on images. Since they are very sensitive to noise, it is common to smooth the image (eg, by using a Gaussian filter) before application of the Laplacian filter, a one-step process called the Laplacian of Gaussian operation (10) (Fig 1). The software was implemented on a personal computer (Maxdata, Würselen, Germany). The first step for the image analysis was to import thin-section CT images by means of a local area network in Digital Imaging and Communications in Medicine format. With the software, we were able to assess the IA and WA of any bronchus visible on the thin-section CT image. Although there was no limitation in regard to the number of bronchi studied, the software was developed to enable evaluation of five bronchi on the same section. After the observer (V.P. or F.L.) independently traced a square region of interest that encompassed the selected bronchus and, therefore, that depended on the size of the selected bronchus, the algorithm was used to calculate a discrete $3 \times 3$ Laplacian of Gaussian mask. A spatial convolution was applied to the whole image to get an accurate depiction of airway contours. The selected region of interest was then displayed and magnified with a factor of five to make the region more visible to the operator. The resultant image was binary: white pixels represented the bronchial wall and black pixels represented the bronchial IA. When this first step did not allow complete separation of the bronchial wall from adjacent structures such as blood vessels or when the bronchus contour was not closed, manual editing of pixels was performed by two authors (V.P. or F.L.) independently (Fig 1, E–G). IA and WA were then calculated by using a growing algorithm. A white pixel was selected within the bronchial wall, and all adjacent pixels connected at four points were aggregated automatically as far as the edge of the wall; the same operation was repeated for the IA by using a black pixel. Last, the number of pixels was converted to an area expressed in square millimeters.

**Validation of the Software**

**Phantom study.**—A phantom was made of five silicone tubes (mean attenuation, 84 HU; range, 70–94 HU) embedded in a foam block (mean attenuation, −800 HU). Real IA and WA of the hollow silicone tubes were calculated from the external diameters and wall thickness measurements obtained by one author (V.P.) by using a micrometer (Outilsam, Paris, France); the measurements were determined with an accuracy to the closest 0.05 mm and ranged from 5.60 to 52.04 mm$^2$ for IA and from 6.53 to 30.16 mm$^2$ for WA. Thin-section CT scans (Somatom Plus 4S; Siemens, Erlangen, Germany) were acquired in the helical mode by using a single-section scanner with the following parameters: 1-mm collimation, 120-kV voltage, 165-mA current, 750-msec rotation time, and pitch of 2. Thin-section CT acquisitions were performed by one author (F.L.) from 0° (strictly perpendicular to the long axis of the tubes) to 60°, corresponding to a ratio of largest to smallest diameter from 1.00 to 2.00, to measure tubes at different oblique angles. CT data were reconstructed with a high-spatial-frequency algorithm, 13-cm field of view, and 512 × 512 matrix (voxel size, 0.25 × 0.25 × 1.00 mm) and displayed on the monitor with parenchymal window width and level of 1600 HU and −450 HU, respectively. Images were transferred to the workstation and were segmented by using DACLOG. Comparisons were performed between micrometer measurements and DACLOG measurements.

**Excised sheep lung study.**—A 1.350-kg lung was removed from a freshly sacrificed sheep in the local slaughterhouse and was fixed in inflation with formalin fumes by using a method adapted from the technique described by Welbel and Vidone (11). Once the lung was fixed, it was cut into 1-cm-thick transverse slices perpendicular to its largest axis, in a plane similar to that which is used with CT.

Airways on the cut surface of each thick slice were visualized and photo-
graphed by using a method adapted from King and colleagues (7). This technique was thought to be the most suitable to match the thin-section CT images with the digitized photographs of the slices. The images were digitized with a high-resolution camera (Nikon, Tokyo, Japan). Two rulers with a graduation of 0.5 mm were included on the image for calibration. The rulers were positioned at 90° to minimize the measurement errors caused by image distortions. The images were transferred to a personal computer. The measurement of airways was performed manually with graphic analysis software (Scion Image, version beta 4.0.2; Scion, Frederick, Md). Care was taken to perform a calibration before each measurement to prevent image distortion. The software measured the area within the external border of the bronchus (or EA) and IA; WA was calculated (WA = EA – IA). After the selection of bronchi to be analyzed by one author (V.P.), two independent observers (V.P., F.L.) performed the measurements of each selected bronchi. The first observer (V.P.) performed two measurements.

The lung slices were then scanned by using thin-section CT. To ensure the best fit between thin-section CT images and digitized images, each thick slice of the lung was placed between two cardboard sheets, which were oriented in the scanning plane to ensure that as much of the slice surface was in the acquisition plane of the scanner as possible. The volume of acquisition encompassed the whole lung volume between cardboard sheets for each thick slice. Therefore, because the lung surface of the thick slice was not perfectly flat, the next contiguous 1-mm slice was used for analysis if parts of the cut surface were incomplete on the first 1-mm slice. The same scanner, settings, and reconstruction algorithm that were used for the images of the phantom were used for those of the excised sheep lung. Selected thin-section CT images of the cut surface were segmented by using DACLOG, and the same two independent observers performed the measurements of each selected bronchus.

Clinical Study

Subjects.—The study prospectively included 24 subjects during 2 months, without any occupational dust exposure, as follows: nine nonsmokers (mean age, 53 years ± 5.6 [standard error of the mean]; six women and three men) who never smoked and had normal lung function (group 1); seven smokers (mean age, 56 years ± 5.6; three women and four men) with normal lung function (group 2); and eight smokers (mean age, 65 years ± 4.0; zero women and eight men) with COPD (group 3). COPD was assessed by one author (P.B.) and was defined by a marked decrease in forced expiratory volume in 1 second (FEV1), specific airway conductance, and forced expiratory flow between 25% and 75% of vital capacity (FEF25%–75%). All subjects gave their written informed consent to participate in the study, which was approved by the local ethics committee. The time between performance of thin-section CT and that of pulmonary function tests was less than 7 days.

Thin-section CT and analysis.—Single-section thin-section CT scans were obtained with the same CT unit as was used for the phantom study in sequential and spiral modes for assessment of emphysema and airway dimensions, respectively. The spiral mode was chosen for assessment of airway dimensions to scan a large volume of data during a breath hold, despite the use of single-section CT. Sequential 1-mm-thick thin-section CT sections were acquired every 10 mm (120-kV voltage, 70-mA current, 1-mm
collimation, and 750-msec rotation time) and encompassed both lungs, were reconstructed on a 35-cm field of view with a $512 \times 512$ matrix, and were visualized on lung windows (window width $= 1800$ HU, window level $= -700$ HU). The visual score described by Goddard et al (12), was used to obtain a qualitative CT score for emphysema. Destruction of a certain percentage of the lung caused by emphysema was identified by one author (F.L., with 15 years of experience), with the following scale: score of 1, destruction of 1%–25%; score of 2, destruction of 26%–50%; score of 3, destruction of 51%–75%; and score of 4, destruction of more than 75% of the lung. The sum of the scores for each of the three lobes was used as the global severity score. The maximum possible score was 12 per lung.

In addition, for measurement of airway dimensions, single-volume thin-section CT scans were acquired in the spiral mode with parameters similar to those used for the excised sheep lung study (1-mm collimation, 750-msec rotation time, pitch of 2), through an anatomic region of interest of 40-mm length starting 4 cm below the level of the carina in the lower lobe of the right lung. We used spirometric gating at a level of breath hold corresponding to 90% of vital capacity (13). This method was used to trigger signals for scans to be obtained at a selected level of respiration and to interrupt airflow during scanning. The level of inspiration was defined by means of a small hand-held transducer (Micro Medical Instruments, Rochester, England) through which the patient was asked to breathe. A microcomputer that was connected to the transducer was used to determine vital capacity and to generate trigger signals for scans to be obtained at a user-selected level of inspiration. These levels could be chosen in percentage of vital capacity. As the trigger signal was generated and sent to the CT scanner, airflow was interrupted mechanically by the closing of a valve attached to the transducer. Therefore, the momentary status was kept constant for the duration of CT scanning.

All the patients were alert and cooperative and prepared for the examination by practicing the breathing maneuver before the CT study. When the patient was positioned in the scanner, a spirometric measurement of vital capacity was obtained. Then, acquisitions were performed at 90% of vital capacity. The patients completed the study within 10 minutes. Reconstructions were obtained by one author (F.L., with 15 years of experience) with a 13-cm field of view and a $512 \times 512$ matrix. These sections were then used for image evaluation and were transferred to a workstation for image analysis with DACLOG.

The area of the lung sampled for the study was situated in the lower lobe of the right lung to avoid as many cardiac artifacts as possible, and all sections that were free from artifacts were used. According to the results of our phantom study, measurements were performed on all bronchi with a ratio of largest bronchial diameter to smallest bronchial diameter of less than 1.50. A mean number of 25 sections were available per patient. For the purpose of a comparison assessed by one author (P.B.), in addition to WA and IA, the IA/VA ratio and the WA/IA ratio for each selected bronchus were calculated. For each patient, the summation of WA to the summation of IA (ie, $\Sigma$WA/$\Sigma$IA) ratio and the summation of WA to the summation of IA + WA (ie, $\Sigma$WA/$\Sigma$(IA + WA)) ratio were also determined. To evaluate the relative contribution of airway size, images of the bronchi were subcategorized in four sets according to the IA + WA value (set A, $<10$ mm$^2$; set B, between 10 and 15 mm$^2$; set C, between $>15$ and 20 mm$^2$; and set D, $>20$ mm$^2$). The lung attenuation in the parenchyma surrounding the analyzed bronchi was also measured on the sections selected for airway dimension measurements by the observer (F.L.) with use of round regions of interest of 10–15 mm in diameter in areas away from blood vessels and equally distributed between anterior (nondependent), lateral, and posterior (dependent) areas (14). Mean values of lung attenuation were compared.

**Statistical Analysis**

Since WA and IA were not normally distributed, values were log transformed. In the phantom study, the accuracy of measurements with DACLOG was compared with that of measurements with the micrometer and with data obtained at various angles on the segmented image by using Pearson correlation coefficients and one-tailed paired t tests. According to Bland-Altman analysis (15–17), the reliability of DACLOG was evaluated by one author (P.B.) after log transformation of data to obtain more normal data by using (a) the Pearson correlation coefficient, (b) the lack of agreement (ie, bias estimated with the mean difference and the standard deviation of the difference), (c) the intraclass correlation coefficient, and (d) the means of DACLOG values and those of the observer plotted against their differences. The repeatability of DACLOG was also analyzed by the same author as follows: (a) The measurement error was evaluated graphically, with the plotting of the individual subject standard deviations against their means, and analytically, with the Spearman rank correlation coefficient, (b) The within-subject standard deviation also was evaluated. For the clinical study, comparison between groups of patients and/or sets of airways was determined by the same au-

**TABLE 1**

Phantom Study: Comparison of DACLOG and Micrometer Measurements at Various Angles on Thin-Section CT Images

<table>
<thead>
<tr>
<th>Parameter, Tube No., and P Value</th>
<th>Micrometer Measurement</th>
<th>DACLOG Measurement and Angle*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0° (1.00)</td>
<td>15° (1.04)</td>
</tr>
<tr>
<td></td>
<td>30° (1.15)</td>
<td>45° (1.41)</td>
</tr>
<tr>
<td></td>
<td>60° (2.00)</td>
<td></td>
</tr>
<tr>
<td>IA (mm$^2$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5.60</td>
<td>4.96</td>
</tr>
<tr>
<td>2</td>
<td>6.98</td>
<td>5.93</td>
</tr>
<tr>
<td>3</td>
<td>13.27</td>
<td>13.10</td>
</tr>
<tr>
<td>4</td>
<td>20.11</td>
<td>22.11</td>
</tr>
<tr>
<td>5</td>
<td>52.04</td>
<td>52.41</td>
</tr>
<tr>
<td>P value†</td>
<td>.45</td>
<td>.94</td>
</tr>
<tr>
<td>WA (mm$^2$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>6.53</td>
<td>8.51</td>
</tr>
<tr>
<td>2</td>
<td>5.59</td>
<td>9.02</td>
</tr>
<tr>
<td>3</td>
<td>13.80</td>
<td>14.31</td>
</tr>
<tr>
<td>4</td>
<td>30.16</td>
<td>27.01</td>
</tr>
<tr>
<td>5</td>
<td>26.50</td>
<td>26.40</td>
</tr>
<tr>
<td>P value†</td>
<td>.28</td>
<td>.19</td>
</tr>
</tbody>
</table>

*Values in parentheses are the ratios of large bronchial diameter to small bronchial diameter. The paired t test was performed on log-transformed values. There was a significant difference between micrometer and DACLOG data at 60° (ratio, >1.50).
†Statistically significant.
tor by using the following: (a) one- and two-way analysis of variance (ANOVA); (b) ANOVAs with generalized linear model procedures, with consideration of both between- and within-subject variation; and (c) unpaired t test of DACLOG measurements and the Pearson correlation coefficient for determination of the correlation between the DACLOG measurement and the pulmonary function test results. Results were considered significant when there was a difference with P < .05. All analyses were performed with software (NCSS 2001; NCSS Statistical Software, Kaysville, Utah).

RESULTS

Phantom Study

The comparison between measurements performed with either the micrometer or DACLOG on thin-section CT images did not show a significant difference for both IA and WA values when acquisition was performed at an angle that was less than 60° (Table 1). The measurements were perfectly reproducible over multiple sessions, since no addition or retrieval of pixels was necessary. Thus, the subsequent analysis was limited to bronchi with a ratio of large bronchial diameter to small bronchial diameter that was smaller than 1.50.

Excised Sheep Lung Study

Forty-eight bronchi with IA ranging from 1.6 to 137.4 mm² (median, 13.4 mm²) and WA ranging from 5.3 to 54.1 mm² (median, 16.1 mm²) were evaluated by using both DACLOG on thin-section CT images and a manual method on specimens (Fig 2). To validate the DACLOG measurements, we compared these measurements with those performed by two different trained observers, the first of whom performed measurements at two different times by using the manual method (Table 2). For both IA and WA, there was a strong correlation between the DACLOG data and the data obtained by both observers (Fig 3, Table 2). The lack of agreement was greater for WA values than it was for IA values (Fig 4, Table 2). In addition, for IA values, the difference between the DACLOG method and the manual method decreased when the bronchial size increased (Fig 4a). The measurement error assessed by using the within-subject standard deviation (0.01 for IA and 0.02 mm² for WA), however, was minimal. Standard deviations of IA and WA assessed by using DACLOG were not correlated with the mean values of IA and WA, respectively (Fig 5). The repeatability of DACLOG measurement of IA and WA was as accurate as that of the manual method (Table 2).

Clinical Study

Patient characteristics are reported in Table 3. There was no difference between groups in terms of age or total lung capacity. Functional parameters related to airway obstruction or complications of the disease (ie, FEV₁, specific airway conductance, FEF₂₅%-₇₅% and residual volume) in patients with COPD (group 3),
however, were significantly different from those of nonsmokers and smokers without COPD (groups 1 and 2, one-way ANOVA). Nevertheless, group 1 did not differ from group 2 in terms of functional parameters (unpaired t test). In regard to the score for emphysema, patients with COPD had a higher score than did the other subjects. Lung attenuation in the surrounding parenchyma was, however, consistent within the three groups (Table 3).

A total of 970 bronchi from 24 patients were selected according to the largest diameter to smallest diameter ratio, which was assessed graphically as smaller than 1.50. IA and WA were then measured with DACLOG on thin-section CT images (mean of analyzed bronchi per patient, 44.4; range, 20–67). There was no significant difference among the three groups in terms of number of selected bronchi (P > .05, ANOVA and unpaired t test). Since IA and WA values were not normally distributed, statistical analysis was performed after log transformation of data. No significant difference was found among the three groups concerning the caliber of 970 bronchi (Fig 6a) with generalized linear model ANOVA (F = 1.46, df = 2, P = .25), which confirmed that bronchi samples of similar caliber were selected within the three groups.

With a similar approach, IA values in 970 bronchi were found to be significantly different among groups (Fig 6b) with generalized linear model ANOVA (F = 9.72, df = 2, P = .001), whereas WA values did not differ (Fig 6c) with generalized linear model ANOVA (F = 3.14, df = 2, P = .06). The WA/IA ratio in 970 bronchi, which reflects normalized airway thickness, was significantly different among groups (Fig 6d) with generalized linear model ANOVA (F = 13.12, df = 2, P < .001). By using multiple comparison procedures, whereas WA/IA ratio was higher in COPD patients (group 3) than it was in healthy control subjects or in smokers without COPD (groups 1 and 2), no difference was found between groups 1 and 2. Moreover, with consideration of heterogeneity of airway wall thickness within the bronchial tree, we found a significant difference in terms of WA/IA ratio when we compared airway sets A–D in 970 bronchi (one-way ANOVA, df = 6, P < .001).

We took into account both the various sizes of the bronchi and the heterogeneity of airway thickness by using the \( \Sigma \text{WA}/\Sigma \text{IA} \) ratio (Fig 6e) or the \( \Sigma \text{WA}/(\Sigma \text{IA} + \Sigma \text{WA}) \) ratio (Fig 6f) for each patient. Since the number of bronchi per patient

---

**TABLE 2**

Measurement of Agreement and Repeatability of DACLOG in Evaluation of 48 Bronchi

<table>
<thead>
<tr>
<th>Parameter, Method, and Time</th>
<th>IA</th>
<th>WA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurements of agreement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DACLOG vs manual method 1, time 1</td>
<td>0.974</td>
<td>0.878</td>
</tr>
<tr>
<td>DACLOG vs manual method 1, time 2</td>
<td>0.953</td>
<td>0.881</td>
</tr>
<tr>
<td>DACLOG vs manual method 2, time 1</td>
<td>0.954</td>
<td>0.859</td>
</tr>
<tr>
<td>DACLOG vs manual method and time</td>
<td>0.985</td>
<td>0.883</td>
</tr>
<tr>
<td>Repeatability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DACLOG over manual method and time</td>
<td>0.999</td>
<td>0.998</td>
</tr>
<tr>
<td>Manual method 1 and 2 over time</td>
<td>0.954</td>
<td>0.954</td>
</tr>
</tbody>
</table>

Note.—Values were obtained after log transformation of IA and WA. ICC = intraclass coefficient, SD = standard deviation.
TABLE 3
Patient Clinical, Functional, and Radiologic Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>P Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients</td>
<td>9</td>
<td>7</td>
<td>8</td>
<td>. . .</td>
</tr>
<tr>
<td>Sex</td>
<td>M</td>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3</td>
<td>4</td>
<td>8</td>
<td>. . .</td>
</tr>
<tr>
<td>F</td>
<td>6</td>
<td>3</td>
<td>0</td>
<td>. . .</td>
</tr>
<tr>
<td>Age (y)</td>
<td>53 ± 5.6 (27–70)</td>
<td>56 ± 5.6 (29–71)</td>
<td>65 ± 4.0 (42–77)</td>
<td>.24</td>
</tr>
<tr>
<td>Smoking (pack-years)</td>
<td>0 ± 0</td>
<td>21 ± 6.4</td>
<td>42 ± 3.5</td>
<td>.001</td>
</tr>
<tr>
<td>FEV&lt;sub&gt;1&lt;/sub&gt;</td>
<td>103 ± 5.6</td>
<td>108 ± 5.7</td>
<td>61 ± 7.6</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Specific airway conductance</td>
<td>(1/kPa · sec&lt;sup&gt;−1&lt;/sup&gt;)</td>
<td>1.94 ± 0.3</td>
<td>2.8 ± 0.7</td>
<td>0.54 ± 0.1</td>
</tr>
<tr>
<td>FEF&lt;sub&gt;25–75%&lt;/sub&gt; Value</td>
<td>3.3 ± 0.3</td>
<td>3.2 ± 0.4</td>
<td>1.2 ± 0.3</td>
<td>.001</td>
</tr>
<tr>
<td>Percentage of predicted</td>
<td>98 ± 7.3</td>
<td>98 ± 7.8</td>
<td>37 ± 8.7</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Total lung capacity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value (L)</td>
<td>5.7 ± 0.3</td>
<td>5.4 ± 0.5</td>
<td>6.4 ± 0.4</td>
<td>.16</td>
</tr>
<tr>
<td>Percentage of predicted</td>
<td>103 ± 4.3</td>
<td>95 ± 6.7</td>
<td>95 ± 5.2</td>
<td>.41</td>
</tr>
<tr>
<td>Residual volume</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value (L)</td>
<td>2.1 ± 0.2</td>
<td>1.6 ± 0.2</td>
<td>3.2 ± 0.4</td>
<td>.002</td>
</tr>
<tr>
<td>Percentage of predicted</td>
<td>108 ± 10.1</td>
<td>80 ± 5.4</td>
<td>132 ± 15.2</td>
<td>.02</td>
</tr>
<tr>
<td>Emphysema CT score (0–24)</td>
<td>0.0 ± 0.0</td>
<td>0.3 ± 0.2</td>
<td>1.8 ± 0.5</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Lung attenuation (HU)</td>
<td>−851 ± 6.8</td>
<td>−847 ± 8.8</td>
<td>−860 ± 6.3</td>
<td>.46</td>
</tr>
</tbody>
</table>

Note.—Data are mean ± standard error of the mean. Data in parentheses are ranges.
* P value indicates the probability of the null hypothesis when we compared patients with normal lung function (group 1), smokers with normal lung function (group 2), and smokers with COPD defined by a significant decrease in FEV<sub>1</sub> (group 3) by using one-way ANOVA.

In this study, we validated an algorithm that was used to measure airway lumen and WA on thin-section CT images. We then used this algorithm to demonstrate that airway wall thickness assessed with quantitative measurements of IAs and WAs of bronchi was related to COPD in smokers. Finally, we identified a parameter that reflects airway wall thickness (ΣWA/ΣIA) that can be used to discriminate between smokers and nonsmokers even in the absence of COPD.

To circumvent interobserver variability and parallax errors ascribed to manual methods (6), we used dedicated software for tracing internal and external contours of bronchi. The resulting contour was also independent from the thin-section CT window setting, a fundamental condition for assessment of airway WA (18,19). The method is semiautomatic and allows correction of the airway external wall contour despite its connection to a blood vessel, without assumption of perfect roundness and symmetry of the WA. Previous investigators developed algorithms that allow measurement of airway lumen and WAs with CT.

Okazawa et al (20) and McNamara and associates (21) described a technique in which the internal and external perimeters of the airways are traced on images obtained at a window width of 1500 HU and at a window level of −450 HU. Amirav et al (19) developed a computerized algorithm for measurement of airway lumen area that was based on an edge detection contour algorithm by using the full-width-at-half-maximum principle, which has less subjectivity and greater speed. Webb et al (22) and Brown and colleagues (23) used algorithms for which the operator draws radial lines from the airway lumen through the whole thickness of the airway wall.

Wood et al (24) developed a threshold method to define airway lumen area; the lumen centroid was used to define the central axis of the airway so that the angle of orientation could be measured. Subsequently, these authors used this algorithm to perform a three-dimensional reconstruction of the airway tree after they converted the data into isotropic voxels and performed accurate measurements of cross-sectional diameters of phantom tubes at various angles (24). King et al (7) developed and validated an automatic image analysis algorithm with which they were able to take into account the airway angulations by using only three contiguous sections and to correct the effect of airway orientations on the values of luminal areas and WAs.

Prêteux et al (25) developed an automatic method of segmentation that was based on mathematical morphology (image processing tool for measurement of morphologic features) and that was able to provide an accurate cross section of airways larger than 4 mm in diameter. Owing to the anatomy of the lung, most of the airways likely run oblique to the plane of the section rather than perpendicular to it. Therefore, on a two-dimensional thin-section CT image, there is an error in the calculation of IA and WA, depending on how acutely the airway is angled with relation to the acquisition plane. In our phantom study, this error was acceptable when the analysis was restricted to the airways with a largest bronchial diameter–to–smallest bronchial diameter ratio that was smaller than 1.50.

Discrepancies between thin-section CT measurements and those performed on specimens can be related to volume averaging of the folds along the mucosal surface and to irregularities on the adventitial surface of the airway at sites of parenchymal attachment in obliquely orientated airways. In addition, CT scanners also have a point-spread function in which the attenuation value of any pixel is affected by the attenuation value of the adjacent pixels. As a consequence, this has the effect of enlarging small tubular structures that run parallel to the CT axis and causes artifactual thickening of thin planar structures such as airway walls.

Therefore, validation of such dedicated software before its use for clinical purposes is required. For validation, we used fixed inflated excised sheep lungs. With excised sheep lungs, the effect of volume
averaging can be taken into account, whereas this cannot be done with an airway phantom, in which irregularities of the mucosal and adventitial surfaces and presence of adjacent blood vessels cannot be duplicated. The techniques that have been published have been validated by using data from phantom studies (7,19, 21,24) and excised animal lungs (7) or by developing a realistic modeling of airways included in CT scans of animal lungs obtained in vivo (25).

Since a history of smoking is associated with bronchial inflammation (1,26,27), even in the absence of COPD (3), we evaluated airway dimensions with thin-section CT in three groups of patients; we compared data in smokers with COPD and smokers without COPD with the data in healthy nonsmoking control subjects. We found that the WA was increased in smokers with COPD compared with that in healthy control subjects. This parameter, however, was not different between

**Figure 6.** Box plots show distribution of parameters in 970 bronchi by using DACLOG. Group 1 corresponds to nonsmokers with normal lung function; group 2, smokers with normal lung function; and group 3, smokers with COPD. (a–d) Medians with 25% and 75% interquartiles; error bars represent 5th and 95th percentiles. (e, f) Bars represent the mean, and error bars represent the standard error of the mean for each group. The parameters that reflect airway thickening (ΣWA/ΣIA and ΣWA/ΣIA + WA) ratios could be used to discriminate smokers from nonsmokers even in the absence of COPD. (a) IA + WA. (b) IA. (c) WA. (d) WA/IA ratio. (e) Individual ΣWA/ΣIA ratio. (f) Individual ΣWA/ΣIA (WA + IA) ratio.

**Figure 7.** Representative transverse CT images. Squares were traced around the bronchi chosen for analysis by using DACLOG. By using a qualitative approach, it was difficult to differentiate smokers from nonsmokers even in the absence of COPD. (a) Male 51-year-old nonsmoker with normal lung function from group 1. (b) Female 53-year-old smoker with normal lung function from group 2. (c) Male 61-year-old smoker with COPD from group 3.
healthy control subjects and smokers without COPD or between smokers with COPD and smokers without COPD.

In vivo, WA seems to be more suitable, in theory, than IA for assessment of pathologic alterations in bronchi, because IA could be affected by both bronchial thickness and loss of elastic recoil caused by emphysema (14). In our study, we found that IA was related to functional parameters (ie, FEV₁, FEF₂₅%–₇₅%), whereas WA was not. In addition, IA was significantly different in smokers with COPD and in those without COPD. Since bronchial pathologic findings are not uniformly distributed throughout the whole bronchial tree, we searched for parameters that reflect both heterogeneity and various sizes of bronchi. The WA/IA ratio was mostly affected by IA and could not be used to discriminate healthy control subjects from smokers without COPD. In contrast, the WA/ΣIA ratio, which is calculated by using various bronchi in each patient, was significantly different among the three groups.

To the best of our knowledge, in only one study have the researchers evaluated airway dimensions with CT in smokers who had COPD (8). The authors showed that not only IA but also WA, assessed by using manual CT analysis of a single bronchus (ie, apical bronchus of the upper lobe of the right lung), was correlated, although weakly, with a decrease in FEV₁. The absence of correlation between WA and functional parameters in our study could be related to our small sample size (n = 24). The discrepancy between our results and those obtained by Nakano et al (8) could be explained by the absence of a control group in the latter study. Although our sample size was limited, we analyzed a large range of bronchi per patient and compared airway parameters with those in a control group. We also took into account the potential heterogeneity of pathologic findings with evaluation of various bronchus sizes.

As did Nakano et al (8), we evaluated global lung emphysema. Moreover, we also analyzed emphysema in the lung surrounding the bronchi that were examined. We found that, whereas patients with COPD have a greater amount of emphysema determined by means of the thin-section CT score for emphysema, especially in the upper lobes, the mean lung attenuation was similar within the surrounding parenchyma of the examined bronchi among the three groups. In group 3 of this study, the score for emphysema was lower than was usually found in patients with COPD disease. Nevertheless, our goal in this preliminary study was only to verify the absence of the role of emphysema in determination of the bronchial diameter, and this role clearly needs further investigation to evaluate how emphysema and bronchial abnormalities in such disease affect the bronchial diameter. CT sensitivity for detection of emphysema, however, is lower than sensitivity for detection at pathologic analysis (28). Therefore, we cannot exclude a loss of elastic recoil to explain the lower airway IA in patients with COPD. Further improvements, which include multisection CT acquisition, should enable us to analyze bronchial thickness of the whole bronchial tree in a procedure with a single acquisition and, thus, to optimize the measurement of the WA/ΣIA ratio parameter.

Our study had several limitations. The first one was the small number of subjects included, but it was a validation study, and further studies need to be performed to determine the clinical importance of these results. Therefore, this study was also affected by imperfect matching in terms of sex, age, and history of smoking among groups, and this may have affected the bronchial thickness, since a correlation of the bronchoarterial ratio with age and smoking has been found in asymptomatic subjects (29). Another limitation was that we chose the bronchi of the lower lobe of the right lung for assessment of the bronchial wall thickness in humans. This site was chosen because it allowed a more convenient method to obtain a tangential view of the bronchus and artery; therefore, the number of pixels that needed to be edited from the automated segmentation was limited. In addition, transmitted cardiac motion artifacts obscure the lower lobe of the right lung less than they do that of the left lung. Further studies are needed to evaluate the technique for other parts of the lung, such as the upper lobes, but a multisection acquisition will allow the whole lung to be scanned.

In conclusion, we validated an algorithm that can be used clinically to measure airway lumen and WA on thin-section CT images and determined a new parameter that reflects airway wall thickness and that allows discrimination of healthy subjects from smokers with COPD or smokers without COPD. By using these measurements, we found larger

### Table 4

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>IA (mm²)</th>
<th>WA (mm²)</th>
<th>WA/IA Ratio</th>
<th>IA + WA (mm²)</th>
<th>ΣWA/ΣIA Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>-0.58 (.004)</td>
<td>0.23 (.28)</td>
<td>0.54 (.007)</td>
<td>-0.28 (.20)</td>
<td>0.50 (.01)</td>
</tr>
<tr>
<td>Smoking (pack-years)</td>
<td>-0.62 (.001)</td>
<td>0.24 (.26)</td>
<td>0.62 (.001)</td>
<td>-0.28 (.19)</td>
<td>0.56 (.004)</td>
</tr>
<tr>
<td>FEV₁ Value (L)</td>
<td>0.71 (&lt;.001)</td>
<td>-0.13 (.54)</td>
<td>-0.66 (&lt;.001)</td>
<td>0.44 (.03)</td>
<td>-0.54 (.006)</td>
</tr>
<tr>
<td>Percentage of predicted</td>
<td>0.60 (.002)</td>
<td>-0.21 (.32)</td>
<td>-0.63 (.001)</td>
<td>0.28 (.19)</td>
<td>-0.51 (.01)</td>
</tr>
<tr>
<td>Specific airway conductance (1/kPa · sec⁻¹)</td>
<td>0.44 (.03)</td>
<td>-0.34 (.11)</td>
<td>-0.53 (.01)</td>
<td>0.07 (.76)</td>
<td>-0.45 (.03)</td>
</tr>
<tr>
<td>FEF₂₅%–₇₅% Value (L/sec)</td>
<td>0.77 (&lt;.001)</td>
<td>-0.22 (.30)</td>
<td>-0.74 (&lt;.001)</td>
<td>0.42 (.04)</td>
<td>-0.65 (&lt;.001)</td>
</tr>
<tr>
<td>Total lung capacity</td>
<td>0.66 (&lt;.001)</td>
<td>-0.23 (.27)</td>
<td>-0.69 (&lt;.001)</td>
<td>0.32 (.12)</td>
<td>-0.39 (.002)</td>
</tr>
<tr>
<td>Value (L)</td>
<td>-0.10 (.65)</td>
<td>0.28 (.19)</td>
<td>0.21 (.31)</td>
<td>0.15 (.49)</td>
<td>0.24 (.26)</td>
</tr>
<tr>
<td>Percentage of predicted</td>
<td>0.19 (.37)</td>
<td>0.08 (.71)</td>
<td>-0.12 (.59)</td>
<td>0.20 (.34)</td>
<td>-0.04 (.86)</td>
</tr>
<tr>
<td>Residual volume</td>
<td>-0.51 (.01)</td>
<td>0.29 (.17)</td>
<td>0.57 (.004)</td>
<td>-0.17 (.43)</td>
<td>0.54 (.007)</td>
</tr>
<tr>
<td>Value (L)</td>
<td>-0.28 (.18)</td>
<td>0.18 (.41)</td>
<td>0.34 (.11)</td>
<td>-0.08 (.72)</td>
<td>0.32 (.13)</td>
</tr>
</tbody>
</table>

Note.—Values are Pearson correlation coefficients that show correlation between DACLOG airway parameters and functional parameters. Values in parentheses are P values that indicate the probability of the null hypothesis that r = 0.
References


MR Imaging Follow-up after Percutaneous Radiofrequency Ablation of Renal Cell Carcinoma: Findings in 18 Patients during First 6 Months

PURPOSE: To prospectively evaluate the magnetic resonance (MR) imaging findings seen within the first 6 months after radiofrequency (RF) thermal ablation of renal cell carcinoma (RCC).

MATERIALS AND METHODS: After providing written informed consent, 18 patients (17 men, one woman; mean age, 71.2 years) with RCC underwent MR imaging–guided percutaneous RF thermal ablation, which was performed by using protocols approved by a comprehensive cancer center protocol committee and the institutional review board for human investigation. The study was Health Insurance Portability and Accountability Act compliant. Follow-up unenhanced T2-weighted MR images and unenhanced and gadolinium-enhanced T1-weighted MR images were acquired immediately, 2 weeks, 3 months, and 6 months after ablation. Thermal ablation zone size was analyzed, and contrast-to-noise ratios (CNRs) were calculated from the signal amplitudes of the thermal ablation zone, perirenal fat, and normal renal cortex on the MR images. Statistical analyses were performed by using the paired Student t test. P < .05 was considered to indicate statistical significance.

RESULTS: The mean follow-up time was 16.1 months (range, 6.0–41.2 months). The mean sizes of the thermal ablation zones were 6.8, 7.0, 6.1, and 4.7 cm², respectively, at immediate, 2-week, 3-month, and 6-month follow-up MR imaging examinations. Thermal ablation zones were uniformly hypointense and had a surrounding bright rim on T2-weighted images and were predominantly hyperintense on T1-weighted images. Thin rim enhancement with central hypointensity was noted on the gadolinium-enhanced images. Gadolinium-enhanced T1-weighted and unenhanced T2-weighted MR images showed significantly higher CNRs than unenhanced T1-weighted MR images. Residual tumor was detected after RF thermal ablation in two cases and was best seen on unenhanced T2-weighted and gadolinium-enhanced T1-weighted MR images.

CONCLUSION: After initially increasing in size within the first 2 weeks, renal RF thermal ablation zones involuted during the remainder of the MR imaging follow-up period.

© RSNA, 2005
Although radical nephrectomy has long been considered the standard treatment for localized RCC, nephron-sparing surgery has been used increasingly (5). Segmental resection is particularly valuable in patients who have previously undergone nephrectomy or had a contralateral nonfunctioning renal unit. Other minimally invasive treatment modalities, such as laser ablation and radiofrequency (RF) thermal ablation, also have become of interest. These procedures are being performed increasingly in patients who either are not surgical candidates because of comorbidities or refuse to undergo surgery.

During the past 5 years, limited experience with RF thermal ablation in patients with RCC has been gained by numerous research groups (6–15). The surveillance protocol after performing this procedure usually consists of dedicated contrast material–enhanced computed tomography (CT) of the kidney. However, a considerable number of eligible patients cannot receive contrast agents that contain iodine because of preexisting impaired renal function or severe contrast material allergies. These patients are usually referred for contrast-enhanced magnetic resonance (MR) imaging of the kidney. Thus, the purpose of our study was to prospectively evaluate the short- and midterm MR imaging findings seen after RF thermal ablation of RCC.

### MATERIALS AND METHODS

#### Patients

Between August 1999 and September 2003, after providing written informed consent, 18 patients (17 men, one woman; mean age, 71.2 years; range, 25–86 years) with solid RCC masses who met phase II trial eligibility criteria (described below) were treated with RF thermal ablation, which was performed by using protocols approved by a comprehensive cancer center protocol committee and the University Hospitals of Cleveland/Case Western Reserve University institutional review board for human investigation, between August 1999 and September 2003. Our study was Health Insurance Portability and Accountability Act compliant. The main inclusion criteria for this trial were as follows: (a) the patient had to have a neoplasm that was not amenable to surgical therapy with curative or substantial palliative intent, have undergone unsuccessful chemotherapy or therapy with biologic response modifiers, or have a tumor that was unlikely to respond to conventional chemotherapy; and (b) the maximal renal tumor diameter had to be 4 cm or smaller.

The 18 patients underwent 21 RF thermal ablation procedures with MR imaging guidance, which was performed by using an open 0.2-T system (Magnetom Open; Siemens Medical Solutions, Erlangen, Germany). One patient had two RCC foci, another patient had an additional contralateral RCC, and a third patient with RCC was treated a second time after incomplete ablation. All tumors were either located peripherally away from the renal sinus or exophytic. Detailed descriptions of the interventional MR imaging suite and the MR imaging–guided ablation procedure that we used have been reported previously (15).

#### Follow-up MR Imaging

Follow-up MR imaging was performed in all patients immediately after the completion of the RF ablation, while the patient was still in the open MR imaging system. The low-field-strength (0.2-T) follow-up MR imaging protocol involved the acquisition of transverse T1-weighted spin-echo (SE) images (440/15, five signals acquired, 169 or 215 × 256 matrix, 30 × 40-cm or 36 × 36-cm field of view) before and after intravenous administration of 0.2 mL of gadopentetate dimeglumine (Magnevist; Berlex Laboratories, Wayne, NJ) per kilogram of body weight (0.1 mmol/kg) or 0.2 mL/kg (0.1 mmol/kg) gadodiamide (Omni scan; Nycomed-Amersham, Princeton, NJ) and unenhanced coronal T2-weighted fast SE images (4000/102, four signals acquired, 126 or 182 × 256 matrix, 25 × 40-cm or 36 × 36-cm field of view, echo train length of seven).

High-field-strength (1.5-T) MR imaging was performed (with Magneton Vision, Magnetom Symphony, or Magnetom Sonata system; Siemens Medical Solutions) 2 weeks after RF ablation, quarterly within the first year, and every 6 months thereafter. The high-field-strength follow-up MR imaging protocol involved unenhanced coronal T2-weighted half-Fourier rapid acquisition with relaxation enhancement (RARE) imaging (single-shot turbo SE sequence, 800/42, one signal acquired, 169 or 205 × 256 matrix, 30 × 40-cm or 36 × 36-cm field of view) and transverse two-dimensional T1-weighted gradient-echo (GRE) imaging (fast low-angle shot sequence, 222/2.6, 70° flip angle, 169 or 205 × 256 matrix, 30 × 40-cm or 36 × 36-cm field of view) performed before and 3 minutes after intravenous administration of 0.2 mL/kg (0.1 mmol/kg) gadopentetate dimeglumine or 0.2 mL/kg (0.1 mmol/kg) gadodiamide.

The mean and longest follow-up durations were 16.1 months ± 11.3 (standard deviation) and 41.2 months, respectively; all patients were followed up for at least 6 months.

#### Image and Data Analyses

The follow-up MR images acquired within the first 6 months after ablation were evaluated for ablation zone size, signal intensity characteristics, and evidence of recurrent or residual tumor, which was defined as either an area of hyperintense soft tissue within the ablation zone or along its margin on T2-weighted images or an area of abnormal contrast enhancement on contrast-enhanced T1-weighted images within the ablation zone (16).

The signal amplitudes of the uninvolved renal cortex (remote from the ablation site), the RF thermal ablation zone, and the perirenal fat were measured with each follow-up MR imaging sequence by defining the regions of interest (ROIs) on the free-standing MR imaging workstation. All measurements were performed by the same board-certified radiologist (S.G.N.), who had undergone fellowship training in MR imaging. The sizes of the ROIs used to measure the signal intensity amplitude in tissue ranged from 37 to 899 mm² and were chosen in homogeneous, artifact-free areas of the tissue being measured. The standard deviation of the noise also was measured by using noise ROIs ranging in size from 535 to 676 mm². The ROIs used for noise measurements were oval and had a long axis perpendicular to the phase-encoding direction. The noise measurement ROIs were positioned anterior to the abdominal wall (in front of the treated kidney) at transverse MR imaging and lateral to the treated kidney at coronal MR imaging (17). Each signal intensity amplitude value was calculated as the average value for three separately sampled ROIs. Con-
TABLE 1
Mean CNRs Calculated at Follow-up MR Imaging Immediately after Renal RF Thermal Ablation

<table>
<thead>
<tr>
<th>MR Imaging Sequence</th>
<th>CNR Between RF Ablation Zone and Renal Cortex</th>
<th>CNR between RF Ablation Zone and Perirenal Fat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast SE T2 weighted</td>
<td>14.7 ± 6.1 (3.1–27.0)</td>
<td>18.5 ± 6.1 (9.0–34.6)</td>
</tr>
<tr>
<td>SE T1 weighted</td>
<td>1.0 ± 6.4 (–7.4–16.0)</td>
<td>16.1 ± 6.8 (2.5–23.2)</td>
</tr>
<tr>
<td>Gadolinium-enhanced SE T1 weighted</td>
<td>12.5 ± 5.1 (4.2–21.9)</td>
<td>18.8 ± 7.0 (6.7–35.5)</td>
</tr>
</tbody>
</table>

Note.—Data are mean CNRs ± standard deviations. Numbers in parentheses are ranges.

RESULTS

Size of Thermal Ablation Zone

The mean tumor size before ablation was 5.3 cm² (range, 0.7–11.2 cm²). After an initial increase in thermal ablation zone size within the first 2 weeks, involution of the zones was observed during the remainder of the MR imaging follow-up period. The mean sizes of the thermal ablation zones in biplanar dimensions were 6.8 cm² ± 3.3 (standard deviation) (range, 0.8–12.7 cm²), 7.0 cm² ± 4.3 (range, 1.5–15.4 cm²), 6.1 cm² ± 3.6 (range, 1.2–14.6 cm²), and 4.7 cm² ± 3.0 (range, 1.5–11.4 cm²), respectively, at immediate, 2-week, 3-month, and 6-month follow-up MR imaging examinations.

Signal Intensity Characteristics

At T2-weighted fast SE MR imaging performed at 0.2 T immediately after RF ablation, the ablation zone in all cases appeared as a round or ovoid hypointense region that replaced the intermediate- or high-signal-intensity tumor seen on the preablation image. The hypointense thermal ablation zone was surrounded by a faintly bright rim with a well-defined inner border and an ill-defined outer border. The ablation zones did not have a consistent appearance on the unenhanced T1-weighted SE images: Compared with the intact renal cortex, the ablation zone appeared isointense, with a difference in signal intensity of less than 10%, in 33% of the cases; hyperintense, with a difference in signal intensity of more than 10%, in another 33% of the cases; and hypointense, with a difference in signal intensity of more than 10%, in the remaining 33% of the cases. Thin rim enhancement was noted on all contrast-enhanced MR images obtained immediately after ablation. The uninvolved perirenal fat had a higher signal intensity than the ablation zone at all postprocedural MR imaging examinations. Table 1 shows the mean CNRs between the thermal ablation zone and the surrounding tissue—either the renal cortex or the perirenal fat—calculated at immediate follow-up MR imaging.

For the high-field-strength (1.5-T) follow-up MR imaging examinations performed 2 weeks, 3 months, and 6 months after RF thermal ablation, the T1-weighted SE sequence was replaced by a T1-weighted GRE breath-hold sequence. Thermal ablation zones continued to appear hypointense on the T2-weighted half-Fourier RARE images, similar to their appearance on the images obtained immediately after ablation (Fig 1) (Tables 2, 3). However, the surrounding reactive bright rim that was seen immediately after ablation resolved gradually over time and was barely detectable after the 3-month examination.

The RF thermal ablation zones had variable appearances on unenhanced T1-weighted GRE MR images. The zones appeared hyperintense compared with the renal cortex in the majority of cases: after 13 (62%) of the 21 procedures at 2-week follow-up and after 14 (67%) procedures at the 3- and 6-month follow-up examinations (Fig 2). The remaining thermal ablation zones appeared either isointense (after three [14%] procedures at the 2-week and 6-month follow-up examinations, after four [19%] procedures at 3-month follow-up) or hypointense (after five [24%] procedures at 2-week follow-up, after three [14%] procedures at 3-month follow-up, after four [19%] procedures at 6-month follow-up) compared with the uninvolved renal cortex. In three cases, the thermal ablation zone appeared hyperintense, even when compared with the perirenal fat (Fig 2). Rim enhancement was noted on all contrast-enhanced postablation images, but it resolved gradually over time and was barely detectable, as was the bright rim on T2-weighted fast SE images, after the 3-month examination (Fig 3).

Statistical Analyses

Statistical analyses were performed by using the paired Student t test (with Microsoft Excel and Microsoft Windows XP Professional software; Microsoft, Redmond, Wash) to test the null hypothesis that CNRs calculated by using various MR imaging sequences are the same. P values were derived from comparisons of the uninvolved renal cortex–to–thermal ablation zone CNRs and the perirenal fat–to–thermal ablation zone CNRs calculated with different follow-up MR imaging sequences immediately, 2 weeks, 3 months, and 6 months after ablation. The uninvolved perirenal fat had a higher signal intensity than the ablation zone at all postprocedural MR imaging examinations. Table 1 shows the mean CNRs between the thermal ablation zone and the surrounding tissue—either the renal cortex or the perirenal fat—calculated at immediate follow-up MR imaging.

For the high-field-strength (1.5-T) follow-up MR imaging examinations performed 2 weeks, 3 months, and 6 months after RF thermal ablation, the T1-weighted SE sequence was replaced by a T1-weighted GRE breath-hold sequence. Thermal ablation zones continued to appear hypointense on the T2-weighted half-Fourier RARE images, similar to their appearance on the images obtained immediately after ablation (Fig 1) (Tables 2, 3). However, the surrounding reactive bright rim that was seen immediately after ablation resolved gradually over time and was barely detectable after the 3-month examination.

The RF thermal ablation zones had variable appearances on unenhanced T1-weighted GRE MR images. The zones appeared hyperintense compared with the renal cortex in the majority of cases: after 13 (62%) of the 21 procedures at 2-week follow-up and after 14 (67%) procedures at the 3- and 6-month follow-up examinations (Fig 2). The remaining thermal ablation zones appeared either isointense (after three [14%] procedures at the 2-week and 6-month follow-up examinations, after four [19%] procedures at 3-month follow-up) or hypointense (after five [24%] procedures at 2-week follow-up, after three [14%] procedures at 3-month follow-up, after four [19%] procedures at 6-month follow-up) compared with the uninvolved renal cortex. In three cases, the thermal ablation zone appeared hyperintense, even when compared with the perirenal fat (Fig 2). Rim enhancement was noted on all contrast-enhanced postablation images, but it resolved gradually over time and was barely detectable, as was the bright rim on T2-weighted fast SE images, after the 3-month examination (Fig 3).

Statistical Analyses

A detailed listing of all P values is provided in Table 4. At immediate follow-up MR imaging performed with the low-field-strength (0.2-T) system, the T2-weighted fast SE sequence yielded the highest CNR values between the uninvolved renal cortex and the ablation zone. Differences in CNR between the T2-weighted fast SE sequence and the T1-
weighted SE sequence before as well as after gadolinium-based contrast material administration were statistically significant (Table 4).

With use of high-field-strength MR imaging systems 2 weeks, 3 months, and 6 months after RF ablation, the CNRs between the uninvolved renal cortex and the ablation zone were higher on the contrast-enhanced T1-weighted images than on the unenhanced T2- and T1-weighted images. These differences were statistically significant at all analyses except one: that in which the CNRs calculated at 6-month contrast-enhanced T1-weighted GRE MR imaging follow-up were compared with those calculated at 6-month T2-weighted half-Fourier RARE MR imaging follow-up (Table 4). T2-weighted imaging was superior to unenhanced T1-weighted imaging in the comparison of uninvolved renal cortex-to-ablation zone CNRs (Table 4).

Perirenal fat-to-ablation zone CNRs were not significantly different at immediate-follow-up low-field-strength MR imaging (Table 4). During the further course of the follow-up period with 1.5-T MR imaging, both the T2-weighted half-Fourier RARE sequence and the contrast-enhanced T1-weighted GRE sequence yielded significantly higher CNRs between the perirenal fat and the ablation zone than the unenhanced T1-weighted GRE sequence (Table 4).

### Residual and Recurrent Tumors

The two patients in this series with the largest and most central tumors (maximal diameters, 3.6 and 4.0 cm) were treated with the primary aim of debulk-
DISCUSSION

The high incidence of small renal tumors has forced a reevaluation of radical nephrectomy as the primary therapeutic option for treatment of RCC, and nephron-sparing surgery is increasingly being advocated, especially when renal function is of prime importance (5,18).

Parallel to nephron-sparing surgery, minimally invasive therapeutic modalities such as image-guided percutaneous RF thermal ablation also have been refined (3,7). The monitoring protocol after these procedures usually consists of dedicated contrast-enhanced CT of the kidneys performed within 1 month after the ablation and then at 3 and 6 months. Additional surveillance CT scans are usually obtained every 6–12 months (7). Thermal ablation areas are poorly visualized on contrast-enhanced CT surveillance images, however, and the diagnosis of residual or recurrent tumor is based only on contrast enhancement characteristics (19,20). Although the use of CT as the primary imaging modality is justified because of cost and availability issues, a substantial number of eligible patients cannot be exposed to iodine-containing contrast agents owing to preexisting allergies or impaired renal function, with creatinine levels higher than 2.0 mg/dL (176.8 μmol/L) (7). These patients are usually referred for contrast-enhanced MR imaging of the kidneys (7).

Thermal Ablation Zone Size and Involution over Time

Thermal ablation zones—defined as hypointense areas—could be clearly delineated on all postprocedural T2-weighted MR images obtained in our study. After an average initial increase in the RF thermal ablation zone size of approximately 10% at bidimensional measurements obtained within the first 2 weeks after ablation, the involution of these zones—by an average of approximately 30%—was observed during the following 6 months of the MR imaging follow-up period. This involution most likely reflected the elimination of coagulation necrosis by macrophages and other components of the human immune system and was consistent with the involution pattern that has been seen in the liver, pancreas, and kidneys in animal studies (21–23).

Signal Intensity Characteristics of Thermal Ablation Zones

RF thermal ablation involves a spectrum of tissue damage processes, including deactivation of enzymes, cell membrane rupture and alteration of tissue structure, protein denaturation and aggregation, and vasoconstriction and intravascular coagulation (24). These effects manifest during different temperature elevations and heating durations, and probably only a subset of these effects are observable as obvious MR imaging signal intensity alterations. In addition, various tissue types have varied responses after thermal treatment in terms of MR imaging signal intensity characteristics. Graham et al (24) found that various tissue types can be classified into four groups: fat, blood, neural tissue such as muscle, liver, and kidney, and fibrous or glandular tissue such as muscle, liver, and kidney.

Fat predominantly consists of triglycerides that undergo reversible effects during the thermal ablation process (24). This phenomenon explains the lack of permanent MR imaging signal intensity alterations after thermal treatment in perirenal fat tissue. In short, thermal ab-

the intraprocedural intermittently acquired T2-weighted fast SE images as hypointense tissue that was well contrasted against the hypointense ablation zone. In terms of depiction on the MR images obtained immediately and at subsequent follow-up examinations after ablation, residual RCC was seen best on the T2-weighted images—as hypointense tissue—and it was seen as enhancing tissue on the contrast-enhanced T1-weighted images (Fig 4).

<table>
<thead>
<tr>
<th>MR Imaging Sequences Compared</th>
<th>Immediate Follow-up CNR&lt;sub&gt;F&lt;/sub&gt;</th>
<th>Immediate Follow-up CNR&lt;sub&gt;F&lt;/sub&gt;</th>
<th>Two-week Follow-up CNR&lt;sub&gt;F&lt;/sub&gt;</th>
<th>Two-week Follow-up CNR&lt;sub&gt;F&lt;/sub&gt;</th>
<th>Three-month Follow-up CNR&lt;sub&gt;F&lt;/sub&gt;</th>
<th>Three-month Follow-up CNR&lt;sub&gt;F&lt;/sub&gt;</th>
<th>Six-month Follow-up CNR&lt;sub&gt;F&lt;/sub&gt;</th>
<th>Six-month Follow-up CNR&lt;sub&gt;F&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2 vs T1</td>
<td>&lt;.01&lt;sup&gt;†&lt;/sup&gt;</td>
<td>.63</td>
<td>&lt;.01&lt;sup&gt;†&lt;/sup&gt;</td>
<td>&lt;.01&lt;sup&gt;†&lt;/sup&gt;</td>
<td>&lt;.01&lt;sup&gt;†&lt;/sup&gt;</td>
<td>&lt;.01&lt;sup&gt;†&lt;/sup&gt;</td>
<td>&lt;.01&lt;sup&gt;†&lt;/sup&gt;</td>
<td>&lt;.01&lt;sup&gt;†&lt;/sup&gt;</td>
</tr>
<tr>
<td>T2 vs gadolinium-enhanced T1</td>
<td>.01&lt;sup&gt;†&lt;/sup&gt;</td>
<td>.75</td>
<td>.02&lt;sup&gt;†&lt;/sup&gt;</td>
<td>.38&lt;sup&gt;†&lt;/sup&gt;</td>
<td>.04&lt;sup&gt;†&lt;/sup&gt;</td>
<td>.55</td>
<td>.10</td>
<td>.51</td>
</tr>
<tr>
<td>T1 vs gadolinium-enhanced T1</td>
<td>&lt;.01&lt;sup&gt;†&lt;/sup&gt;</td>
<td>.72</td>
<td>&lt;.01&lt;sup&gt;†&lt;/sup&gt;</td>
<td>&lt;.01&lt;sup&gt;†&lt;/sup&gt;</td>
<td>&lt;.01&lt;sup&gt;†&lt;/sup&gt;</td>
<td>&lt;.04&lt;sup&gt;†&lt;/sup&gt;</td>
<td>&lt;.01&lt;sup&gt;†&lt;/sup&gt;</td>
<td>.49</td>
</tr>
</tbody>
</table>

Note.—Data are <it>P</it> values calculated by using the Student <it>t</it> test. CNR<sub>F</sub> = CNR between uninvolved renal cortex and thermal ablation zone, CNR<sub>FE</sub> = CNR between perirenal fat and thermal ablation zone.

*<it>T</it> = 0.2-T SE and 1.5-T GRE T1 weighted, <it>T</it> = 0.2-T fast SE and 1.5-T half-Fourier RARE T2 weighted.

†Statistically significant difference.
lation zones extending into the perirenal fat appear bright on T1- and T2-weighted MR images.

In contrast, fibrous or glandular tissue such as the renal parenchyma demonstrates irreversible effects, including denaturation and shrinkage of proteins such as collagen and increased hydrophobic interactions, that result in the extrusion of water. These irreversible effects most likely represent the underlying cause of the shortened T2 relaxation time after thermal ablation, which ultimately leads to the uniform hypointense appearance of the ablation zones on T2-weighted MR images. This hypointense appearance on T2-weighted MR images obtained in humans coincides with the experimental data acquired in animal models after renal thermal ablation (16,25); it also resembles the signal intensity pattern seen in other human organs such as the liver and brain.

On the other hand, the appearances of thermal ablation zones on unenhanced T1-weighted MR images have a high degree of variability compared with the uniform hypointense appearance of these zones on T2-weighted MR images. In the current study, although the appearances of the thermal ablation zones as hypointense, isointense, or hyperintense relative to the uninvolved renal cortex were equally common at low-field-strength unenhanced T1-weighted SE MR imaging performed immediately after ablation, the thermal ablation zones appeared hyperintense compared with the uninvolved renal cortex in the majority of cases at further follow-up MR imaging performed with T1-weighted GRE sequences and high-field-strength systems.

The level of reduction in T1 relaxation time during thermal ablation correlates with the degree of tissue vascularity (24). This correlation explains why RF thermal ablation zones appear slightly hyperintense on unenhanced T1-weighted GRE MR images in the majority of cases in the liver and kidney. It also explains why renal RF thermal ablation zones should appear slightly brighter than hepatic RF thermal ablation zones: The degree of vascularity in the kidney is higher than the degree of vascularity in the liver (26).

However, this correlation does not explain why renal RF thermal ablation zones have variable appearances—from hypointense to markedly hyperintense compared with uninvolved renal cortex—on unenhanced T1-weighted MR images; these findings usually are not seen following RF thermal ablation of focal hepatic lesions. Procedure-related diffuse hemorrhage within the thermal ablation zone may be the most likely explanation for this discrepancy. Reasons for this hemorrhage are probably twofold: First, the risk of bleeding during renal biopsy is higher than that during hepatic biopsy owing to the purely arterial blood supply of the kidneys compared with the mainly portal venous blood supply of the liver. Second, RCC is generally
a more hypervascular tumor than hypovascular colorectal metastases to the liver. Blood itself exhibits an abrupt decrease in T1 and T2 relaxation times at temperatures higher than 60°C, which results in a hyperintense appearance on T1-weighted MR images (24). Diffuse hemorrhage may also explain why the hyperintense appearance of renal RF thermal ablation zones is more often appreciated on high-field-strength GRE MR images than on low-field-strength SE MR images.

At contrast-enhanced T1-weighted MR imaging, no substantial enhancement was observed within the RF thermal ablation zone. However, rim enhancement that gradually resolved over time and was barely detectable after the 3-month examination was noted on all contrast-enhanced postablation images.

Residual or Recurrent Tumor

The major reason for performing surveillance imaging after renal RF thermal ablation is the early detection of residual or recurrent tumor. Although contrast-enhanced MR imaging findings are quite similar to the findings seen at CT, the RF thermal ablation zone also is very well depicted on unenhanced T2-weighted MR images. This additional information is very helpful, as seen in the two cases of residual tumor following RF thermal ablation in the current series. In these cases, residual RCC was best seen on the unenhanced T2-weighted and contrast-enhanced T1-weighted MR images.

Study Limitations

The major limitations of our study were the relatively small number of patients examined and the lack of a true reference standard, such as results of gross pathologic and/or microscopic analysis of the RF ablation zone, for comparison with the MR imaging findings. The minor limitations of our study were the two contrast agents and two magnetic field strengths used. However, the statistical analyses were performed by using data acquired at the same follow-up imaging session only and not by using data from different follow-up imaging sessions. Thus, these two minor limitations probably did not substantially affect our quantitative data analyses.

RF thermal ablation zones in the kidneys have the same pattern as RF thermal ablation zones in the liver in terms of the temporal evolution of the ablation zone size. After an initial increase in size within the first 2 weeks, involution occurs during the remainder of the MR imaging follow-up period. Although the signal intensity characteristics of the kidneys on T2-weighted MR images obtained after RF thermal treatment are the same as those of the liver—both organs appear hypointense—the appearance of the renal RF thermal ablation zone on T1-weighted MR images is brighter than that of the hepatic RF thermal ablation zone. This appearance is best appreciated on GRE MR images and most likely reflects diffuse hemorrhage within the thermal ablation zone. The high vascularity of RCCs and the purely arterial blood supply of the kidneys compared with the mainly portal venous blood supply of the liver may explain this finding.

References

Multistage Ethanol Sclerotherapy of Soft-Tissue Arteriovenous Malformations: Effect on Pulmonary Arterial Pressure

PURPOSE: To retrospectively investigate how repeat injections of absolute ethanol in therapeutic doses, required for multisession sclerotherapy of large high-flow soft-tissue arteriovenous malformations (AVMs) in patients with normal cardiopulmonary function, affect pulmonary arterial pressure (PAP).

MATERIALS AND METHODS: Study received approval and waiver of informed consent by institutional review board and was conducted in 16 male and 16 female patients with AVMs who underwent repeat sclerotherapy (142 sessions total) with absolute ethanol from July 1997 to December 2003. PAPs were monitored during first session in all patients. In subsequent sessions, PAP was monitored with pulmonary catheter when predicted single dose of ethanol exceeded 3 mL and total amount exceeded 0.25 mL/kg. PAP was measured in 104 sessions. Serum ethanol levels from blood samples obtained at end of each session were reviewed. PAP parameters were analyzed at beginning and end of each session, and highest value was recorded to assess any increase after repeat therapy. Difference between initial and highest PAP values recorded in a session (Δmax) was noted to determine any increase during repeat sessions. Possible relationship was reviewed between this value and amount of ethanol used. For sessions without PAP monitoring, mixed model was used for statistical analysis.

RESULTS: Total ethanol used was variable. In 43 sessions, highest mean PAP exceeded 25 mm Hg. Incidence of sustained pulmonary hypertension (mean PAP > 25 mm Hg) at end of each session was 30.8% (32 of 104 sessions). Initial PAP parameters did not increase or decrease during repeat sessions. No significant changes in Δmax of systolic and mean PAP were observed with increasing number of sessions. Rather, Δmax of diastolic PAP was reduced after repeat sessions (P = .03). There was no significant correlation between serum ethanol level and PAP parameters at end of sessions. Relationships between Δmax values of systolic, mean, and diastolic PAP and total ethanol used were not significant.

CONCLUSION: High incidence of acute pulmonary hypertension was observed in each sclerotherapy session without lasting effect on PAP.
Radiology

Sclerotherapy of large high-flow soft-tissue AVMs, which is required for multisession therapy, and the effect of repeat injection of absolute ethanol on the pulmonary arterial pressure (PAP). Thus, the purpose of our study was to retrospectively investigate whether repeat injections of absolute ethanol in therapeutic doses, which is required for multisession sclerotherapy of large high-flow soft-tissue AVMs in patients with normal cardiovascular function, have an effect on PAP.

**MATERIALS AND METHODS**

### Patients

A retrospective review, which received full approval and a waiver of informed consent by our institutional review board, was conducted in 34 patients diagnosed as having soft-tissue AVMs who underwent ethanol sclerotherapy at our center from July 1997 to December 2003. The patients included in this study were referred to interventional radiologists as a result of the inoperable state of their AVM lesions. They received sclerotherapy with absolute ethanol (99.6%) and required PAP monitoring more than once. Two patients were excluded from the study because they had dilated cardiomyopathy and moderate pulmonary hypertension before the start of sclerotherapy. The remaining 32 patients did not show any other disorder or disease at the beginning of sclerotherapy.

PAPs were monitored during the first session in all patients. In subsequent sessions, PAP was monitored by using a pulmonary catheter only when the predicted single injection dose of absolute ethanol was greater than 3 mL and the total amount exceeded 0.25 mL per kilogram of body weight. Thirty-two patients with AVMs (16 males and 16 females; mean age, 29.6 years; age range, 11–53 years) underwent a total of 142 sessions of ethanol sclerotherapy with general anesthesia. PAP was measured during 104 sessions for systolic, mean, and diastolic values by using a pulmonary artery catheter (7.5-F Swan-Ganz catheter [931HF-75; Baxter Healthcare, Irvine, Calif]). The mean number of therapy sessions per patient was 4.4 (range, 2–15) (Table 1).

### Procedure

All sclerotherapy procedures that involved the use of absolute ethanol were performed with general anesthesia. Ethanol sclerotherapy was performed by two interventional radiologists (Y.S.D., H.S.B., with 7 and 10 years of experience in interventional radiology, respectively), who had been performing ethanol sclerotherapy since 1996. Cardiopulmonary monitoring was used by introducing a pulmonary artery catheter into the right internal jugular vein. Staged ethanol embolization was directed against the nidus itself, not against the vascular feeders; the goal was to embolize all or part of the nidus until the desired clinical result was achieved. Routes of vascular access to attack the nidus were chosen after an initial angiogram was obtained. Transarterial and transvenous catheterization with the use of a coaxial catheter and/or percutaneous direct puncture were required to reach the nidus for embolization. In some cases, proximal inflow occlusion was performed during the ethanol injection. To achieve vascular stasis, an intravascular occlusion balloon catheter was used for the trunk and pelvis area, and external pneumatic blood pressure cuffs were used for the extremities.

Among the 104 total sessions of ethanol sclerotherapy, 43 sessions were performed with the direct puncture method, 49 sessions with the transcatheter method, and the remaining 12 sessions with both methods. Arteriograms were acquired to determine the exact flow characteristics of the AVMs. Moreover, to determine the volume of ethanol used during embolization and the rate of injection, test injections of the contrast material were performed with fluoroscopic monitoring. The amount of ethanol used was based on the amount of contrast material required to fill the fistula without opacifying the normal vessels. After ethanol injections, we waited for 10 minutes and then acquired arteriograms to determine whether the therapy was successful. Complete embolization of at least one compartment of the AVMs required meticulous repetition of the previously de-

---

**TABLE 1**

Demographic Data for Patients with Soft-Tissue AVMs

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. of Patients*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>16</td>
</tr>
<tr>
<td>Female</td>
<td>16</td>
</tr>
<tr>
<td>Age (y)</td>
<td>29.6 ± 10.8 (11–53)</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>164.8 ± 9.1 (145.0–181.0)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>62.2 ± 13.8 (39.2–96.0)</td>
</tr>
<tr>
<td>Mean no. of sclerotherapy sessions</td>
<td>4.4 (2–15)</td>
</tr>
<tr>
<td>Mean duration of multisession sclerotherapy (mo)</td>
<td>14.3 ± 17.0 (2.0–55.0)</td>
</tr>
<tr>
<td>Mean duration of procedure (min)</td>
<td>207.8 ± 49.7 (105–325)</td>
</tr>
</tbody>
</table>

Note.—Values in parentheses are ranges.

* Unless specified otherwise.

† Data are mean ± standard deviation.
scribed technique (3). The single ethanol dose per each injection during the sessions ranged from 2 to 10 mL. The total amount of absolute ethanol used per session was less than 1.01 mL/kg. PAP values were recorded at the start and end of each procedure, and the highest PAP value was recorded after injection of absolute ethanol or after deflation of an intravascular occlusion balloon or external pneumatic cuffs.

When the mean PAP exceeded 25 mm Hg, nitroglycerine was administered as either a bolus injection (50–100 μg) or by means of continuous infusion (0.3–3.0 μg/kg/min) intravenously. When increased PAP was sustained at the end of the session, the patient was kept at the intensive care unit for close PAP monitoring and for continuous administration of nitroglycerine.

Data Acquisition

The initial, maximum, and final PAP values (systolic, mean, and diastolic) were investigated and compared with the total amount of absolute ethanol used at each session and the serum ethanol level measured from a blood sample acquired at the end of each session. We also investigated the incidence where the highest mean PAP exceeded 25 mm Hg, where nitroglycerine was required to control the increased PAP during the sessions, and where there was sustained pulmonary hypertension (mean PAP > 25 mm Hg) at the end of each session. The complications that occurred in the 142 sessions were reviewed. Data acquisition was performed by two of the authors (B.S.S., Y.S.D.).

Data Analysis

We analyzed the data to determine whether the initial, maximum, and final PAP values (systolic, mean, and diastolic), as well as the difference between the initial-session PAP measurement and corresponding maximum PAP measurement (Δmax) for each session, increased at the repeat sessions (M.H.K., G.S.K.). We compared the relationship between the Δmax values and the total amount of ethanol used per session. We reviewed whether the highest PAP occurred during the session or at the end of the session. The correlation between the serum ethanol level and PAP parameters (systolic, mean, and diastolic) at the end of the sessions was also investigated. Data analysis was performed by two of the authors (M.H.K., G.S.K.).

### Table 2: Sclerotherapy Data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean total ethanol used (mL)</td>
<td>32.2 ± 14.3 (6.0–69.0)</td>
</tr>
<tr>
<td>Mean ethanol volume (mL/kg)</td>
<td>0.51 ± 0.25 (0.14–1.01)</td>
</tr>
<tr>
<td>Mean serum ethanol level at end of session (mmol/L)</td>
<td>9.29 ± 6.19 (1.2–25.8)</td>
</tr>
<tr>
<td>No. of sessions with highest mean PAP in excess of 25 mm Hg and nitroglycerine infusion required</td>
<td>43 (41.3%)</td>
</tr>
</tbody>
</table>

Note.—Data are mean ± standard deviation, and numbers in parentheses are ranges, unless specified otherwise.

### Table 3: Sclerotherapy Data according to AVM Location

<table>
<thead>
<tr>
<th>AVM Location</th>
<th>No. of Patients</th>
<th>Mean No. of Sessions</th>
<th>Mean Serum Ethanol Levels at End of Sessions (mmol/L)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face</td>
<td>4</td>
<td>7.0 (2–15)</td>
<td>11.5 ± 6.6</td>
</tr>
<tr>
<td>Upper extremity</td>
<td>12</td>
<td>3.8 (2–10)</td>
<td>7.8 ± 6.5</td>
</tr>
<tr>
<td>Trunk</td>
<td>5</td>
<td>2.4 (2–3)</td>
<td>9.9 ± 4.4</td>
</tr>
<tr>
<td>Lower extremity</td>
<td>8</td>
<td>5.5 (3–10)</td>
<td>9.3 ± 6.5</td>
</tr>
<tr>
<td>Pelvis and right thigh</td>
<td>1</td>
<td>6</td>
<td>11.9 ± 6.2</td>
</tr>
<tr>
<td>Right buttock and back</td>
<td>1</td>
<td>2</td>
<td>10.0 ± 5.5</td>
</tr>
<tr>
<td>Pelvic cavity and inguinal area</td>
<td>1</td>
<td>4</td>
<td>12.0 ± 3.1</td>
</tr>
</tbody>
</table>

Note.—Numbers in parentheses are ranges.

* Values are mean ± standard deviation.

Statistical Analysis

To consider the bias caused by the sessions without PAP monitoring, a mixed model was used for statistical analysis with SAS version 8.1 software (SAS Institute, Cary, NC) to analyze the changes in PAP during each session, since individual patients underwent different numbers of sessions. The primary hypothesis in this study was that the difference between the highest values of the systolic, mean, and diastolic PAP and the initial PAP values are positively correlated with the number of sessions. If we suppose that the coefficient of correlation is greater than 0.5 (beyond a moderate correlation), 30 patients were required under the significant level of .05 and a power of .8. By using a mixed model, the relationship between the Δmax values and the total amount of ethanol used per session and the correlation between the serum ethanol level and PAP parameters (systolic, mean, and diastolic) at the end of the sessions were also analyzed. A P value of less than .05 was considered to indicate a statistically significant difference. All statistical analyses were performed by one author (M.H.K.) and a statistician.

RESULTS

The total amount of absolute ethanol used to control the AVM nidus varied: 6.0–69.0 mL was used per session (range, 0.14–1.01 mL/kg), with a mean of 32.2 mL. The mean serum ethanol level at the end of the sessions was 42.8 mg/dL (9.3 mmol/L) (range, 5.4–118.7 mg/dL [1.2–25.8 mmol/L] (Tables 2, 3). Serum ethanol level greater than the legal intoxication level at the end of the session (>80 mg/dL [17.4 mmol/L]) occurred in 11.5% of sessions (12 of 104). The initial mean values of systolic, mean, and diastolic PAPs were 22.6 mm Hg ± 6.4 (standard deviation), 16.3 mm Hg ± 5.1, and 11.6 mm Hg ± 4.4, respectively. The highest mean values of systolic, mean, and diastolic PAPs were 33.3 mm Hg ± 11.1, 24.8 mm Hg ± 8.2, and 18.4 mm Hg ± 7.5, respectively. The mean PAP values at the end of the sessions were 28.9 mm Hg ± 9.2, 21.8 mm Hg ± 7.0, and 16.4 mm Hg ± 6.4, respectively.

Increases of more than 10 mm Hg for systolic, mean, and diastolic PAP immediately after injection occurred in 49.0% of sessions (51 of 104), 37.5% of sessions (39 of 104), and 25.0% of sessions (26 of 104), respectively. The highest PAP values were recorded within 3 minutes of ethanol injection. In 30.8% of sessions (32 of 104), the highest PAP value was recorded at the end of the session, and in 64.4% of sessions (92 of 104), it was recorded during the sessions.
In 43 sessions, the highest mean PAP exceeded 25 mm Hg. Nitroglycerine was required to control pulmonary hypertension in these patients. The incidence of sustained pulmonary hypertension (mean PAP > 25 mm Hg) at the end of each session was 30.8% (32 of 104 sessions). These patients were transferred to the intensive care unit (20 of 32 patients; range, one to four times), and PAP levels returned to baseline levels within a maximum of 10 hours.

Repeat Session Findings

There was no increase or decrease in the initial PAP parameters (systolic, mean, diastolic) during repeat sessions (P = .26, .09, and .61, respectively) (Fig 1). No increase in the \( \Delta_{\text{max}} \) values of the highest and initial systolic and mean PAPs were observed when increasing the number of sessions (P = .52 and .21, respectively) (Fig 2). The relationship between the \( \Delta_{\text{max}} \) values of initial systolic, mean, and diastolic PAPs and the total amount of absolute ethanol used was not statistically significant (P = .38, .22, and .22, respectively). Similar to the findings of Mason and colleagues (13), we found statistically significant correlations between injected serum ethanol levels acquired at the end of session and total amount of ethanol injected (P < .05), but there was no significant correlation between the serum ethanol level and PAP parameters (systolic, mean, and diastolic) at the end of sessions (P = .30, .26, and .29, respectively) (Fig 3).

Complications

Twenty-three (16.2%) complications occurred in 142 procedures. Nineteen (19 of 142 procedures, 13.4%) minor complications (17 skin blisters or necrosis and two transient peripheral nerve injuries) occurred in 14 patients (43.8%). All of these minor complications responded to conservative treatment. There were five major complications (five of 142 procedures, 3.5%) in five patients (15.6%): an infection caused by severe skin necrosis, acute renal failure due to muscle necrosis, permanent median nerve injury of the arm, focal urinary bladder necrosis, etc.
eral oxygen saturation (SpO2) decreased to 90% with end-stage renal disease was reported in a 3-month-old boy. A fatal outcome in a 3-month-old boy, with pulmonary hypertensive crisis accompanying a bronchospasm.

DISCUSSION

Currently, many agents such as absolute ethanol, sodium tetradecyl sulfate, N-butyl cyanoacrylate, and isobutyl 2-cyanoacrylate are used for the treatment of AVM. These are applied by using either a direct puncture technique into the nidus or a superselective catheterization technique to deliver embolotherapy or sclerotherapy agent selectively to the nidus of the lesions (14). Of these agents, absolute ethanol has a unique ability to induce protein denaturation of the endothelial cells with subsequent vessel wall demudation and interruption, which results in the complete obliteration of the vessel lumen rather than simple obstruction. Ethanol does not allow recanalization as a result of permanent damage to the endothelium of the AVM nidus (15–17).

Therefore, a multistage approach by using absolute ethanol, which has the lowest recurrence rate of available techniques, is considered to be more effective than methods involving the use of other embolotherapy or sclerotherapy agents currently available (14,18). Also, ethanol is readily available, has a long shelf life, and is inexpensive.

However, many clinicians who have used absolute ethanol have been alarmed by the various complications as a result of its chemical toxicity, even though the method has produced excellent results with promising outcomes and an increased chance of cure (3,13,19). These events were the result of various conditions, such as dysrhythmia, cardiac collapse, and/or acute and severe pulmonary vasospasm induced by the ethanol. A fatal outcome in a 3-month-old boy with end-stage renal disease was reported after arterial injection of 4.5 mL of absolute ethanol for renal ablation (20). In one of our cases involving right-thigh and large pelvic AVM lesion, the PAP values abruptly increased to 72/49/40 mm Hg (systolic/mean/diastolic), mechanical ventilation became impossible as a result of high resistance of the airway, and peripheral oxygen saturation (Spo2) decreased to 60% after injection of 10 mL of absolute alcohol. PAP values were promptly restored to 24/19/10 mm Hg, however, and mechanical ventilation and Spo2 returned to normal after four repeated bolus injections of nitroglycerine in conjunction with continuous infusion and hyperventilation with 100% oxygen. Even though we have experienced only one case of pulmonary hypertensive crisis among a total of 142 sessions, we do not consider it an insignificant incidence, because the patient with AVM must undergo multiple sessions of sclerotherapy that can result in life-threatening complications.

Pulmonary vasospasm caused by absolute ethanol, even though it is diluted by the time it reaches the pulmonary vasculature, results in acute pulmonary hypertension at the precapillary level, which in turn increases right ventricular afterload and leads to right ventricular failure (21). The right ventricle begins to fail when the systolic pressure is acutely increased by a factor of two—that is, by more than 50 mm Hg. This results in decreased left-sided heart filling, decreased left-sided cardiac output, systemic hypotension, and finally cardiopulmonary collapse (15,21).

Therefore, the rapid correction of pulmonary hypertension should be absolutely warranted before the cascade of events reaches the final stage of pulmonary hypertension—cardiopulmonary collapse. In case of high-flow AVM, the importance of having PAP cautiously monitored by an experienced anesthesiologist for the prompt prevention of serious pulmonary hypertension cannot be overemphasized to prevent such a disastrous outcome.

Hypoxia, acidosis, and hypercapnia can aggravate pulmonary hypertension by means of hypoxic pulmonary vasoconstriction. Cockrill et al (22) noted that supplemental oxygen induced a mild degree of pulmonary vasodilatation and reduced PAP. Therefore, adequate ventilatory management is mandatory during ethanol sclerotherapy. In addition, an adequate depth of anesthesia is needed, because sympathetic stimulation caused by severe pain during injection of absolute ethanol contributes to the increase in PAP (23).

Nitroglycerine, prostaglandin E1, prostaglandin I2, and inhaled nitric oxide are highly selective pulmonary vasodilators with low systemic effects (24–27). In our study, the incidence of pulmonary hypertension was high. Nitroglycerine was used in 43 (41.3%) of the 104 sessions by following our regimen described previously. The incidence of sustained pulmonary hypertension (mean PAP increase > 25 mm Hg) at the end of each session was 30.8% (32 of 104 sessions). Although the incidence of pulmonary hypertension in our study was high, our findings show that repeated sessions of absolute ethanol injections did not elevate initial PAP. From these findings, we can derive the conclusion that repeated ethanol injections within the limited dosage of less than 0.5–1.0 mL/kg absolute ethanol per session and less than 10 mL of single ethanol dose per injection do not seem to cause increasing tendencies of PAP in multisession sclerotherapy.

As Yakes et al (3) recommended, it is important to perform ethanol sclerotherapy within the maximum volume of less than 0.5–1.0 mL/kg absolute ethanol per session. By following this recommendation, our results did not show any statistically meaningful relationship between PAP increase (between highest and initial values of systolic, mean, and diastolic) and the total amount of absolute ethanol used. Furthermore, in 30.8% of sessions, the highest PAP value occurred at the end of sessions, and in 64.4%, it occurred during the procedure. It is important to realize that in two-thirds of the interventions, the highest PAP occurred during the session because the highest single absolute ethanol dose was administered during and not at the end of ethanol sclerotherapy.

In our experience, increases in PAP were higher in subsequent sessions than in previous sessions in some patients, and we questioned whether the highest PAP increases with the repetition of the sessions. We investigated the difference (Δmax) between the initial and the highest PAP in each session and whether it increased during repeat sessions. Statistical analysis showed that there were no increasing or decreasing tendencies, which meant that the reactivity of pulmonary vasculature to the injected ethanol did not increase during multisite ethanol sclerotherapy.
The main limitation of this study is that, even though we used a mixed model to take into account the sessions the volunteers and control group, the bias caused by the 38 sessions without PAP monitoring could not be ruled out completely. In addition, these results failed to show any statistically meaningful relationship between PAP increase and the total amount of absolute ethanol used within the maximum volume of less than 0.5–1.0 mL/kg absolute ethanol per session. It was assumed that the highest PAP may be primarily influenced by a single injection dose rather than the total amount of absolute ethanol. However, the correlations between a single ethanol injection dose and PAP increase were not examined as a result of the limitation of our retrospective study. Also, we could not determine the effect of high PAP increase on cardiac output. Further study may be required to find these correlations.

In conclusion, a high incidence of acute pulmonary hypertension was observed in each session without lasting effect on PAP during multistage ethanol sclerotherapy. Since a tendency toward significantly increasing or decreasing PAP was not shown, we believe that the risk of pulmonary hypertensive crisis may still exist in subsequent sessions.

Acknowledgment: The authors thank Seon Woo Kim, PhD, of Samsung Biomedical Research Institute, Seoul, Korea, for her assistance with statistical analysis.

References

Percutaneous Transhepatic Balloon Dilation of Portal Venous Stenosis in Patients with Living Donor Liver Transplantation

**PURPOSE:** To retrospectively evaluate the long-term effectiveness of percutaneous transhepatic balloon dilation of portal venous stenosis in patients who have undergone living donor liver transplantation.

**MATERIALS AND METHODS:** Institutional review board approval and informed consent were not required. From June 1996 to August 2003, obstructed portal venous blood flow was diagnosed in 45 patients (21 male, 24 female) with a history of living donor liver transplantation; patients ranged in age from 9 months to 61 years (mean, 9.2 years). All stenoses occurred in the extrahepatic portal vein near the anastomosis of the portal vein. All dilation procedures were performed with percutaneous transhepatic puncture of the intrahepatic portal vein and subsequent balloon dilation of the stenosis. Patients who experienced recurrent stenosis underwent another balloon dilation session. Intravascular metallic stents were not deployed because of the possible need for repeated transplantation. The authors used paired *t* tests to compare patients successfully treated with one venoplasty procedure and those requiring repeated venoplasty, with regard to age and stenosis diameter percentages before and after the initial procedure.

**RESULTS:** Percutaneous balloon dilation was technically successful in 35 of 45 patients. In the remaining 10 patients, portal venous thrombotic occlusion precluded access to the mesenteric side of the portal vein. Twenty-five patients were successfully treated with a single session of balloon dilation (group 1). Results at follow-up ultrasonography revealed restenosis in 10 of 35 patients. Recurrent stenosis was resolved by means of repeated balloon dilation in nine patients (group 2). There were no significant differences between groups 1 and 2 in age (*P* = .87) or in stenosis diameter percentages before (*P* = .053) or after (*P* = .95) the initial procedure.

**CONCLUSION:** Percutaneous transhepatic balloon dilation seems to be an effective method for treatment of portal venous stenosis after living donor liver transplantation.

Liver transplantation is an important option in managing the end stages of liver diseases. Living donor liver transplantation has recently been introduced to solve the problem of liver graft shortages. Improvements in surgical techniques and immunosuppression have contributed to improved outcomes after transplantation, but biliary and vascular complications are still important causes of graft failure in liver transplant recipients (1,2). Percutaneous transluminal angioplasty is an established, safe, and effective method for treating anastomotic vascular stenosis after liver transplantation (3–5). In this study, our purpose was to retrospectively evaluate the long-term effectiveness of percutaneous transhepatic balloon dilation of portal venous stenosis in patients who have undergone living donor liver transplantation.
MATERIALS AND METHODS

Patients

Our institutional review board did not require its approval or patients’ informed consent for this retrospective study. Between June 1996 and August 2003, 58 patients with a history of liver transplantation were suspected of having obstructed portal venous blood flow because of clinical signs, such as hematemesis or ascites, or findings at routine Doppler and B-mode ultrasonography (US), computed tomography (CT), and/or magnetic resonance imaging. In 45 patients, portal venous stenosis was confirmed by means of percutaneous transhepatic portography followed by interventional procedures. All stenoses occurred in the extrahepatic portal vein near the anastomosis of the portal vein and/or the interposition venous graft.

The 45 patients ranged in age from 9 months to 61 years (mean, 9.2 years) and included 21 male and 24 female patients. The underlying disease most frequently observed in the recipients was biliary atresia (n = 32). Other diseases included fulminant hepatitis (n = 2), Alagille syndrome (n = 2), hepatoblastoma (n = 2), liver cirrhosis (n = 2), Wilson disease (n = 1), hepatocellular carcinoma (n = 1), liver fibrosis (n = 1), cholangiectasis (n = 1), and glucose storage disease (n = 1). The interval between transplantation and initial intervention ranged from 9 days to 98.3 months (mean, 18.0 months). Lateral segments of the left lobe had been used as grafts in 36 patients, including one patient who received an auxiliary partial orthotopic liver transplant, while nine patients had received right lobe grafts. Portal venous anastomosis was performed with end-to-end anastomosis. All stenoses occurred in the extrahepatic portal vein near the anastomosis of the portal vein and/or the interposition venous graft.

The underlying disease most frequently observed in the recipients was biliary atresia (n = 32). Other diseases included fulminant hepatitis (n = 2), Alagille syndrome (n = 2), hepatoblastoma (n = 2), liver cirrhosis (n = 2), Wilson disease (n = 1), hepatocellular carcinoma (n = 1), liver fibrosis (n = 1), cholangiectasis (n = 1), and glucose storage disease (n = 1). The interval between transplantation and initial intervention ranged from 9 days to 98.3 months (mean, 18.0 months). Lateral segments of the left lobe had been used as grafts in 36 patients, including one patient who received an auxiliary partial orthotopic liver transplant, while nine patients had received right lobe grafts. Portal venous anastomosis was performed with end-to-end anastomosis. In four patients the remaining portal vein was short, and a donor’s vein (ovarian or external iliac) was used as an interposition venous graft.

Balloon Dilation

Before treatment, informed consent was obtained from all patients or their parents. In pediatric patients (age, <18 years), the procedure was conducted with use of general anesthesia, while in adults (age, ≥18 years), the procedure was conducted after administration of local anesthetic and/or intravenous administration of 15 mg pentalozine (Sosegone; Yamanouchi Pharmaceuticals, Tokyo, Japan).

Two authors performed percutaneous transhepatic balloon dilation. One author (K.I.) had 16 years of experience in interventional radiology, and the other (Toyomichi Shibata) had 6 years of experience. The portal vein was punctured with an 18-gauge needle (Hanako, Saitama, Japan) under US and fluoroscopic guidance. After the portal vein was entered, a sheath was inserted into the vein with a 7-F interventional sheath introducer (Brite Tip; Cordis, Roden, the Netherlands). A 0.032-inch angled hydrophilic guidewire (Terumo, Tokyo, Japan) and a 5-F catheter with a hockey stick–shaped tip (Cook, Bloomington, Ind) were then used to traverse the stenotic segment.

After passage through the stenosis, venograms of the portal and superior mesenteric veins were obtained. To confirm the site and degree of stenosis, balloon dilation was attempted with a percutaneous transluminal angioplasty catheter.
The balloon had a diameter of 6–12 mm, which was altered according to the diameter of the mesenteric side of the stenosis, and a length of 40 mm. Sixty seconds of balloon inflation with an atmospheric pressure of 10 atm was exerted on the stenotic segment. Dilation was performed twice, after which a second portal venogram was obtained to evaluate the effectiveness of the procedure. On completion of these interventional procedures, the sheath tract was occluded with collagen material (Avitene; Zeria Pharmaceuticals, Tokyo, Japan), and prophylactic anticoagulation was commenced. Warfarin (Eisai, Tokyo, Japan) was administered to produce an international normalized ratio of 1.0–1.5.

In all cases, the stenosis diameter percentage was measured before and after balloon dilation, and stenoses of more than 50% were considered hemodynamically significant. In 16 cases, pressure gradients in the stenosis were obtained before and after balloon dilation. In many of the initial cases, we performed the interventional procedures without obtaining a pressure gradient. Pressure gradients of more than 3 mm Hg were defined as significant.

**Effectiveness of Dilation**

Doppler and B-mode abdominal US was performed every month after completion of the interventional procedure. All US examinations were performed by one of the board-certified physicians in our institution, who all had more than 5 years of experience in performing vascular US. US was performed with a real-time scanner and a 3.5-MHz transducer (model SSD-5500; Aloka, Tokyo, Japan). The shape of the portal vein was assessed with B-mode US, and the blood flow was assessed with Doppler US.

CT was performed with a Hispeed Advantage scanner (GE Medical Systems, Milwaukee, Wis) in patients who had clinical signs of portal venous stenosis and in whom US failed to depict the portal vein because of the intervening intestine. At contrast material–enhanced CT, portal venous phase scans were obtained with a 3.5-MHz transducer (model SSD-5500; Aloka, Tokyo, Japan). The shape of the portal vein was assessed with B-mode US, and the blood flow was assessed with Doppler US.

CT scans were interpreted by three abdominal radiologists (T.K., Y.M., and Toshiya Shibata), each with more than 5 years of experience in abdominal imaging.

Patients with recurrent stenosis underwent another balloon dilation session performed by the aforementioned authors. Intravascular metallic stents were not deployed because of the possible need for repeated transplantation. The follow-up period ranged from 1.0 to 73.9 months (mean, 24.8 months). We divided the patients into two groups to evaluate the effectiveness. Group 1 included patients who were successfully treated with a single session of balloon dilation, and group 2 included patients who needed to undergo another balloon dilation session because of recurrent stenosis.

**Statistical Analysis**

We compared groups 1 and 2 in terms of age and the stenosis diameter percentage before and after the initial procedure. For multiple comparisons, a paired t test was performed with statistical software (SPSS version 11.0; SPSS, Chicago, Ill). Compensation of multiplex nature was performed by using the Dunnett method.

**TABLE 2**

<table>
<thead>
<tr>
<th>Patient No./Age (y)/Sex</th>
<th>Side of Transplanted Graft</th>
<th>Transplantation-Dilation Interval (d)*</th>
<th>Length of Patency (d)</th>
<th>Stenosis (%)</th>
<th>Preprocedural</th>
<th>Postprocedural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/13/M</td>
<td>Left</td>
<td>35</td>
<td>1131</td>
<td>80</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>2/33/F</td>
<td>Right</td>
<td>674</td>
<td>438</td>
<td>90</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3/1/F</td>
<td>Left</td>
<td>277</td>
<td>613</td>
<td>100</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4/8/M</td>
<td>Left</td>
<td>2949</td>
<td>963</td>
<td>60</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>5/8/F</td>
<td>Left</td>
<td>945</td>
<td>1054</td>
<td>60</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>6/4/F</td>
<td>Left</td>
<td>128</td>
<td>1159</td>
<td>90</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>7/29/F</td>
<td>Right</td>
<td>882</td>
<td>1124</td>
<td>90</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>8/60/M</td>
<td>Left</td>
<td>114</td>
<td>109</td>
<td>100</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>9/8/F</td>
<td>Left</td>
<td>1792</td>
<td>1453</td>
<td>100</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>10/1/M</td>
<td>Left</td>
<td>455</td>
<td>613</td>
<td>70</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>11/2/F</td>
<td>Left</td>
<td>107</td>
<td>158</td>
<td>100</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>12/1/F</td>
<td>Left</td>
<td>138</td>
<td>116</td>
<td>90</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>13/23/F</td>
<td>Right</td>
<td>157</td>
<td>393</td>
<td>50</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>14/2/F†</td>
<td>Left</td>
<td>457</td>
<td>1383</td>
<td>80</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>15/1/F</td>
<td>Left</td>
<td>238</td>
<td>445</td>
<td>80</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>16/16/M</td>
<td>Right</td>
<td>138</td>
<td>347</td>
<td>80</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>17/39/F</td>
<td>Right</td>
<td>412</td>
<td>263</td>
<td>60</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>18/3/F</td>
<td>Left</td>
<td>397</td>
<td>489</td>
<td>90</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>19/1/M</td>
<td>Left</td>
<td>496</td>
<td>438</td>
<td>80</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>20/4/F‡</td>
<td>Left</td>
<td>866</td>
<td>74</td>
<td>80</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>21/3/M</td>
<td>Left</td>
<td>14</td>
<td>1670</td>
<td>80</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>22/3/F</td>
<td>Left</td>
<td>1176</td>
<td>2013</td>
<td>100</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>23/5/F</td>
<td>Left</td>
<td>233</td>
<td>1593</td>
<td>80</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>24/1/M</td>
<td>Left</td>
<td>592</td>
<td>127</td>
<td>100</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>25/1/M</td>
<td>Left</td>
<td>170</td>
<td>144</td>
<td>100</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

* Interval between transplantation and initial session of balloon dilation.
† In these patients, the pressure gradient decreased to less than 3 mm Hg.
‡ These patients received interposition venous grafts (ovarian vein).
P values less than .05 were considered to indicate a statistically significant difference.

RESULTS

Initial Stenosis

Percutaneous balloon dilation was technically successful in 35 of 45 patients (Fig 1, Tables 1–3). In the remaining 10 patients, portal venous thrombotic occlusion precluded access to the mesenteric side of the portal vein. Eight of the 10 patients had left lobe grafts. Twenty-five (71.4%) of the 35 patients, ranging in age from 1 to 60 years (mean, 10.8 years), were successfully treated with a single session of balloon dilation (group 1) (Table 2, Figs 2 and 3). Among these 25 patients, three had postprocedural stenosis of 60%. We judged that balloon dilation was effective in these patients because the pressure gradient decreased to less than 3 mm Hg. Portal venous patency in these patients was maintained for 2.5–67.1 months (mean, 24.4 months). The mean stenosis diameter percentage was 81.6% before and 18.4% after the initial procedure. The interval between transplantation and balloon dilation ranged from 14 days to 98.3 months (mean, 18.5 months).

Recurrent Stenosis

Follow-up US results revealed restenosis in 10 (28.6%) of the 35 patients who underwent successful balloon dilation. In nine (25.7%) of these patients, this restenosis was resolved by means of repeated balloon dilation (group 2; Table 3, Fig 4); dilation was repeated once or twice. The age range of patients in this group was 9 months to 61 years (mean, 9.9 years). Six patients required two balloon dilation sessions to maintain patency, and three patients required three sessions. In group 2, all patients except one had received left lobe segment grafts, and interposition venous grafts had been necessary in two patients. Patency was maintained for 5.5–67.1 months (mean, 23.9 months). The mean stenosis diameter percentage was 93.9% before and 14.4% after the initial procedure. The interval between balloon dilation sessions ranged from 2.3 to 8.9 months (mean, 5.3 months). In one patient, the portal vein occluded despite two balloon dilation sessions. Because this patient had received an auxiliary partial orthotopic liver transplant, the native liver helped maintain liver function and the patient was treated conservatively. The differences between groups 1 and 2 in age (P = .87) and in the stenosis diameter percentage before (P = .053) and after (P = .95) the initial procedure were not significant (Table 1).

Portal Venous Thrombotic Occlusion

The 10 patients with portal venous thrombotic occlusions (eight with left lobe and two with right lobe grafts) were also treated by means of conservative methods (n = 7), surgical reanastomosis (n = 2), or repeated transplantation (n = 1). In seven patients, the intrahepatic portal vein remained patent with the help of cavernous transformation and stable liver functions; therefore, these patients were also treated conservatively, without surgical revision. Two patients needed to undergo surgical reanastomosis because of gastrointestinal bleeding and hypofunction of the liver graft, and one patient required another transplant because of graft failure that resulted from bile duct stenosis (Fig 1).

<table>
<thead>
<tr>
<th>Patient No./ Age (y)/Sex and Session No.</th>
<th>Transplantation-Dilation Interval (d)*</th>
<th>Session Interval (d)</th>
<th>Length of Patency (d)</th>
<th>Stenosis (%) Preprocedural</th>
<th>Postprocedural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2/M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>553</td>
<td>256</td>
<td>100</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>644</td>
<td>115</td>
<td>90</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>2/2/F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>464</td>
<td>382</td>
<td>100</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>541</td>
<td>305</td>
<td>90</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>660</td>
<td>186</td>
<td>95</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3/3/F†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>664</td>
<td>2217</td>
<td>100</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>798</td>
<td>2083</td>
<td>80</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>868</td>
<td>2013</td>
<td>60</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>4/3/M†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>987</td>
<td>1495</td>
<td>95</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1253</td>
<td>1229</td>
<td>50</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>5/12/M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>359</td>
<td>1527</td>
<td>80</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>503</td>
<td>1383</td>
<td>80</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>6/0/M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>115</td>
<td>1163</td>
<td>100</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>329</td>
<td>949</td>
<td>90</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>588</td>
<td>690</td>
<td>100</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>7/1/F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>196</td>
<td>585</td>
<td>100</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>458</td>
<td>323</td>
<td>100</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>8/61/F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>117</td>
<td>251</td>
<td>90</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>187</td>
<td>181</td>
<td>80</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>9/5/M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>745</td>
<td>484</td>
<td>80</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>938</td>
<td>291</td>
<td>90</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Note.—Patient 8 received a right lobe graft; all other patients received a left lobe graft.

* Interval between transplantation and initial session of balloon dilation.
† Patients 3 and 4 received interposition venous grafts (ovarian and external iliac veins, respectively).
Complications

Procedure-related complications occurred in two of the 45 patients examined. In a 39-year-old man, an intrahepatic pseudoaneurysm of the hepatic artery was detected 6 days after the procedure and was successfully treated with microcoil embolization. In a 33-year-old woman, an intrahepatic portal venous thrombus formed during balloon dilation. It was resolved by means of direct infusion of 60,000 IU of urokinase (Mitsubishi Pharma, Osaka, Japan).

DISCUSSION

Vascular complications after liver transplantation include occlusion or stenosis at the site of anastomosis in the hepatic artery, portal vein, or vena cava (2,6–10). Among these complications, portal venous stenosis is rarely observed but nevertheless involves risks of graft failure (10). In 1991, Raby et al (3) reported on portal vein venoplasty in children. It has since been established in many hospitals as the treatment of choice for posttransplantation portal venous stenosis. Reports of the long-term outcomes of portal venous venoplasty in transplant recipients without metallic stent placement are rare, to our knowledge. Buell et al (2) reported long-term venous complications after liver transplantation and mentioned that living donor related transplants result in a higher incidence of portal venous stenosis than cadaveric liver transplants. With living donor liver transplantation, left lobe segments are frequently used as grafts (11). During left lobe graft transplantation, the extrahepatic portal vein slants to the right posterior, and the portal vein rotates backward to the left anterior into the liver. Stenotic areas almost always exist near the flexion. As the liver graft grows, it might kink or compress the flexion.

In this study, no significant differences were observed between groups 1 and 2 with regard to stenosis diameter percentages or age. Nine of the 10 patients in whom restenosis occurred had left lobe grafts, as did eight of the 10 patients whose portal veins were already occluded. Twenty-five (71.4%) of the 35 patients with portal venous stenosis were treated by means of a single balloon dilation session. Considering that long-term patency (mean, 27 months) was maintained without stent placement in 34 (97.1%) of 35 patients, we have shown that percutaneous transhepatic balloon dilation is a successful method for treating complications of portal venous stenosis after liver transplantation.

Funaki et al (4) reported long-term results...
of portal venoplasty in 25 patients who had undergone balloon dilation and metallic stent placement. They showed the effectiveness of metallic stents in the treatment of recurrent or “elastic” stenosis. In our series, recurrent stenosis was successfully treated with repeated balloon dilation in nine patients. Fortunately, we have not yet encountered any cases of elastic stenosis, which does not respond to balloon dilation. We are prepared to use metallic stents in such cases.

Stent placement is known to be more effective than balloon dilation alone in maintaining patency during the treatment of arterial stenosis, such as coronary, iliac, or renal artery stenosis with arteriosclerosis obliterans (12–14). In such cases, adjacent organ growth does not occur as it does after liver transplantation. With arterial stenting, the regions of stent placement are usually linear, whereas posttransplantation stenotic regions are often angulated, especially when left lobe grafts are used. If metallic stents are used, there is a risk of stent-edge stenosis and occlusion as a result of graft growth. In cases of occluded metallic stents, including those with intimal hyperplasia, repeated interventions may be necessary (15,16). Moreover, if repeated transplantation is required, metallic stent placements make operations even more difficult. Therefore, the use of metallic stents in portal venous stenosis after living donor liver transplantation should be avoided if possible.

In 10 (22.2%) of the 45 patients examined in this study, portal venous thrombotic occlusion precluded access to the extrhepatic portal vein. All 10 patients required sustained treatment for collateral vessels, and in the most severe case, repeated transplantation was necessary. Compared with biliary obstructions, early portal venous stenosis is difficult to detect from clinical signs and symptoms alone. Furthermore, sometimes the portal venous anastomotic site cannot be seen with US because of intestinal artifacts, and CT cannot be performed so many times in pediatric cases. Despite these difficulties, the early diagnosis of portal venous stenosis is very important.

There were limitations in our study. First, there were too few cases to compare the effectiveness of interventional procedures between right lobe and left lobe grafts. The statistical significance of the results cannot be determined with such small numbers to calculate sensitivities and specificities. Further clinical studies are needed to confirm these preliminary results. Second, pressure gradients in the stenosis were not obtained before and after balloon dilation in all cases, and obtaining pressure gradients is a more accurate way to assess the clinical condition than a gross index, such as the stenosis diameter percentage. In the future, we need to obtain pressure gradients in all cases. To conclude, our findings suggest that percutaneous transhepatic balloon dilation is a safe and effective method for the treatment of portal venous stenosis after living donor liver transplantation.

References


Letters to the Editor

Vascular Endothelial Growth Factor–related Angiogenesis

From:
Ukihide Tateishi, MD, PhD
Division of Diagnostic Radiology, National Cancer Center Hospital
5–1–1 Tsukiji, Chuo-Ku, Tokyo 104-0045, Japan
e-mail: utateishi@ncc.go.jp

Editor:
In the October 2004 issue of Radiology, Dr Yi and colleagues (1) report their experience with dynamic contrast material–enhanced multi–detector row computed tomography (CT) and comparison of findings with vascular endothelial growth factor (VEGF) in solitary pulmonary nodules. The authors reported extent of VEGF-related angiogenesis in both malignant (n = 38) and benign (n = 16) pulmonary nodules compared with microvessel density by using pathologic specimens. Results from our own study (2) would support their observations with contrast-enhanced dynamic CT, but we suggested that such correlations between CT measurements and VEGF-related angiogenesis were evident in patients with malignant pulmonary neoplastic nodules.

Benign pulmonary nodules often contain active inflammatory process, which also shows angiogenic activity. Angiogenesis of benign lesions is attributable to increased concentration of various molecules released by epithelial and mesenchymal cells. Although lesion molecules with regard to inflammatory process also include VEGF, the pattern of angiogenesis and microvascular remodeling of benign lesions is transient and mainly found in the active phase during the course of disease. This evidence is fundamentally different from the pattern shown in malignant neoplasm, because tumor cells themselves can continuously produce VEGF-related peptides and can affect tumor proliferation and aggressiveness. Therefore, benign lesions should be evaluated and discussed separately from malignant lesions.

The authors used monoclonal VEGF antibodies for evaluation of tumor angiogenesis in the study. The VEGF family, consisting of VEGF-A, VEGF-B, VEGF-C, VEGF-D, and VEGF-E on the basis of DNA splicing, is known as a potent neoangiogenic factor involved in the development and growth of vascular endothelial cells. VEGF-A and VEGF-B are major mediators of both angiogenesis and vasculogenesis, though their receptor VEGFR-1 and VEGF-C and VEGF-D regulate lymphangiogenesis through VEGFR-3. VEGF-A is a strong lymphangiogenic factor that promotes proliferation of lymphatic microvessels and vascular permeability. If we select the monoclonal antibody for VEGF subtype in the evaluation of tumor angiogenesis, the result of the study will depend on the selection of VEGF agonist and will be causative of bias in study analysis. In addition, in the evaluation of in vitro tumor angiogenesis with monoclonal antibody for VEGF subtype, it is considered difficult to analyze the distinction between tissue angiogenesis and lymphangiogenesis solely by means of contrast-enhanced dynamic CT.

References

Drs Lee and Han respond:

We thank Dr Tateishi for expressing his interest in our article on dynamic enhanced multi–detector row CT and the comparison with VEGF and microvessel density in solitary pulmonary nodules (1). As he correctly indicated, angiogenesis in benign lesions is transiently present in the active phase and is presumably associated with VEGF among various molecules released by epithelial and mesenchymal cells. In contrast, angiogenesis in malignant tumors is persistent and influential in tumor proliferation and is spread as a result of continuous shedding of VEGF from tumor cells.

VEGF is a group of families that involve the regulation of endothelial cell growth and contains members of VEGF-A, VEGF-B, VEGF-C, and VEGF-D. The classification of the family mainly depends on its DNA slicing. VEGF-A plays an essential role in angiogenesis—thus, tumor growth. VEGF-C facilitates entry of tumor cells into the lymph vasculature by regulating lymphatic vessel growth. The roles of VEGF-B and VEGF-D are not clear at the moment. VEGF-B appears to be involved in lymph node metastasis and progression of adenocarcinoma of the lung in the setting of low levels of VEGF-D (2).

The VEGF polyclonal antibody used in our study (1) was the VEGF-A family [VEGF-A-20]: sc-152; Santa Cruz Biotech, Santa Cruz, Calif). Therefore, the extent of its staining might have represented tumor angiogenesis related to VEGF-A (the extent of microvessel density). As we expected, the extent of VEGF staining was higher in malignant nodules and showed significant positive correlation with peak attenuation values at dynamic CT.

The attenuation values of pulmonary nodules at dynamic CT reflect the amount of contrast medium transport within a nodule in the intravascular and interstitial spaces (3). In other words, contrast enhancement values at a given time are a summation of the intra- and extravascular concentrations of contrast medium. The attenuation value of a nodule in the early phase (wash-in phase, usually within 2 minutes) of dynamic study represents the extent of tumor vascularity, whereas that in the late phase (washout phase, usually after 5 minutes) depends on the extent of combination of cells, nonvascular mesenchymal tissue, and fibrosis (4,5). Tumors with high VEGF staining show high enhancement in the early phase of dynamic study. This was proved in our study.
(1), in which malignant nodules with higher VEGF staining showed higher enhancement, higher maximum relative enhancement ratio, and shorter time to peak enhancement than did benign nodules. However, care should be taken because benign tumors of high vascularity, such as pulmonary sclerosing hemangioma, hemangiopericytoma, leiomyoma, and highly vascular hamartoma, can show high enhancement in the early phase of dynamic study. In these tumors, the morphology at CT usually suggests benignity with well-defined margins.

As far as we know, there is no single imaging tool that enables the evaluation of exclusive lymphatic vessel growth extent into a nodule. Therefore, it is difficult to measure the extent of VEGF-C (factor involved in lymphatic vessel dilation) family activity in a nodule with an imaging study. Further study on the relationship between in vitro staining extent of lung cancer with the VEGF family (including all members of the family) and presence of hilar or mediastinal lymph node metastasis at imaging, including positron emission tomography (PET) and PET/CT, would be interesting.

References

Kyung Soo Lee, MD,* and Joungho Han, MD†
Department of Radiology and Center for Imaging Science* and Department of Pathology,† Samsung Medical Center, Sungkyunkwan University School of Medicine
50 Ilwon-Dong, Kangnam-Ku, Seoul 135–710, South Korea
e-mail: kyungs.lee@samsung.com
Errata


As a result of a restrospective audit of study records, the authors determined that insufficient information was archived to verify study eligibility for seven of 18 study participants with hemoglobin AS. Statistical analysis of the data, excluding these participants, revealed a mean percentage of hemoglobin S of 35.2 ± 4.8 (n = 7) in subjects with a tortuosity index of 0 at magnetic resonance (MR) angiography, compared with a mean percentage of hemoglobin S of 40.0 ± 1.5 (n = 4) in subjects with a tortuosity index of 1 or 2 at MR angiography. Thus, after removal of the potentially ineligible study participants from the analysis, statistical significance of the difference was reduced to \( P < .09 \). Therefore, the authors must amend their conclusion to state that the data show a trend toward (rather than a statistical significance for) a correlation between the percentage of hemoglobin S and the degree of cranial vasculopathy.

“Tumor Volume in Pharyngolaryngeal Squamous Cell Carcinoma: Comparison at CT, MR Imaging, and FDG PET and Validation with Surgical Specimen.” Radiology 2004; 233:93–100

Page 96, transverse positron emission tomographic (PET) image in Figure 1, a misalignment between the outline and the PET image occurred during processing of this figure. Image should appear as follows:


Page 113, first sentence of the second paragraph should read, “Retro-ocular involvement ranges from perineural infiltration to diffuse masses associated with exophthalmia (7).”

Background: The ability to accurately predict pathologic response to preoperative chemotherapy may have a significant impact on the treatment strategy for non-small cell lung cancer (NSCLC). The purpose of this study was to examine the accuracy of positron emission tomography (PET) scanning in predicting the pathologic response to preoperative chemotherapy in the primary tumor and draining lymph nodes. Methods: A total of 25 patients were enrolled in two separate phase II trials investigating induction chemotherapy for NSCLC. All patients underwent pre-treatment and post-treatment PET scans followed by surgical resection. A significant PET scan response was defined as a reduction in the standard uptake value by 50% or more. We defined a major pathologic response as either no disease or microscopic disease only in the primary tumor. The percentage change in standard uptake value was then calculated and correlated with pathologic response in the primary tumor. In addition, the presence or absence of nodal metastases as determined by the postchemotherapy PET scan was compared with final pathologic nodal stage. Results: The positive and negative predictive values for PET detection of major pathologic response in the primary tumor were 43% and 100%, respectively. Positron emission tomography did not accurately predict nodal status in 52% of patients. The positive and negative predictive values of PET to detect node-positive disease were 73% and 64%, respectively. For N2 disease the positive predictive value of PET scans was less than 20%. Conclusions: Positron emission tomography scanning does not reliably predict pathologic response to preoperative chemotherapy in NSCLC in either the primary tumor or the draining lymph nodes.

Authors' Abstract

Reason for selecting abstract:

• Surgical and pathologic correlation
• Comparison with CT

Selected by Sheila D. Davis, MD
New York Presbyterian Hospital, New York, NY

Hemothorax Due to Extramedullary Erythropoietic Masses. Stergios Tassiopoulos, Kosmas Konstantopoulos, Yannis Rombos, et al. Ann Thorac Surg 2004; 77:323–324. (S.T., 12 Kotieou str, Athens 11521, Greece; e-mail: tassiop@hol.gr)

We describe a 27-year-old male patient suffering from beta-thalassemia intermedia who presented with a nontraumatic spontaneous hemothorax due to extramedullary hemopoietic foci. In reviewing the literature, four similar reports were found. The details of this unusual entity are discussed.

Authors' Abstract

Reason for selecting abstract:

• Unusual complication

Selected by Sheila D. Davis, MD
New York Presbyterian Hospital, New York, NY

Breast

Growth Patterns and the Risk of Breast Cancer in Women. Martin Ahlgren, Mads Melbye, Jan Wohlfahrt, et al. N Engl J Med 2004; 351:1619–1626. (M.A., Department of Epidemiology Research, Danish Epidemiology Science Center, Statens Serum Institut, Artillerivej 5, DK-2300, Copenhagen S, Denmark; e-mail: ahlgren@ssi.dk)

Background: Adult height and body-mass index influence the risk of breast cancer in women. Whether these associations reflect growth patterns of the fetus or growth during childhood and adolescence is unknown. Methods: We investigated the association between adult height and breast cancer in a cohort of 117,415 Danish women. Birth weight, age at menarche, and annual measurements of height and weight were obtained from school health records. We used the data to model individual growth curves. Results: Growth on vital status, age at first childbirth, parity, and diagnosis of breast cancer was obtained through linkages to national registries. Results: During 3,333,359 person-years of follow-up, 3340 cases of breast cancer were diagnosed. High birth weight, high stature at 14 years of age, low body-mass index (BMI) at 14 years of age, and peak growth at an early age were independent risk factors for breast cancer. Height at 8 years of age and the increase in height during puberty (8 to 14 years of age) were also associated with breast cancer. The attributable risks of birth weight, height at 14 years of age, BMI at 14 years of age, and peak growth at an early age were independent risk factors for breast cancer. Height at 8 years of age and the increase in height during puberty (8 to 14 years of age) were also associated with breast cancer. The attributable risks of birth weight, height at 14 years of age, BMI at 14 years of age, and peak growth at an early age were independent risk factors for breast cancer.

Authors' Abstract

Reason for selecting abstract:

• Importance of birth weight
• Importance of growth peak

Selected by Anthony V. Proto, MD
School of Medicine, Virginia Commonwealth University, Richmond

Cardiovascular System


Abstracts of Current Literature

Thorax


Background: Despite extensive literature, the diagnostic role of d-dimer for deep venous thrombosis (DVT) or pulmonary embolism (PE) remains unclear, reflecting multiple d-dimer assays and concerns about differing sensitivities and variability. Purpose: To systematically review trials that assessed sensitivity, specificity, likelihood ratios, and variability among d-dimer assays. Data Sources: Studies across all languages were identified by searching PubMed from 1983 to January 2003 and EMBASE from 1988 to January 2003. Study Selection: The researchers selected prospective studies that compared d-dimer with a reference standard. Studies of high methodologic quality were included in the primary analyses; sensitivity analysis included additional weaker studies. Data Extraction: Two authors collected data on study-level factors: d-dimer assay used, cutoff value, and whether patients had suspected DVT or PE. Data Synthesis: For DVT, the enzyme-linked immunosorbent assay (ELISA) and quantitative rapid ELISA dominate the comparative ranking among D-dimer assays. Specificity, likelihood ratios, and variability across studies.

Conclusions: Findings are based largely on indirect comparisons of test performance characteristics across studies. Conclusion: The ELISAs in general dominate the comparative ranking among the d-dimer assays for sensitivity and negative likelihood ratio. For excluding PE or DVT, a negative result on quantitative rapid ELISA is as diagnostically useful as a normal lung scan or negative duplex ultrasonography finding.

Authors' Abstract

Reason for selecting abstract:

• Systematic review

Selected by Anthony V. Proto, MD
School of Medicine, Virginia Commonwealth University, Richmond

Hemothorax Due to Extramedullary Erythropoietic Masses. Stergios Tassiopoulos, Koas Konstantopoulos, Yannis Rombos, et al. Ann Thorac Surg 2004; 77:323–324. (S.T., 12 Kotieou str, Athens 11521, Greece; e-mail: tassiop@hol.gr)

We describe a 27-year-old male patient suffering from beta-thalassemia intermedia who presented with a nontraumatic spontaneous hemothorax due to extramedullary hemopoietic foci. In reviewing the literature, four similar reports were found. The details of this unusual entity are discussed.

Authors' Abstract

Reason for selecting abstract:

• Unusual complication

Selected by Sheila D. Davis, MD
New York Presbyterian Hospital, New York, NY

Breast

Growth Patterns and the Risk of Breast Cancer in Women. Martin Ahlgren, Mads Melbye, Jan Wohlfahrt, et al. N Engl J Med 2004; 351:1619–1626. (M.A., Department of Epidemiology Research, Danish Epidemiology Science Center, Statens Serum Institut, Artillerivej 5, DK-2300, Copenhagen S, Denmark; e-mail: ahlgren@ssi.dk)

Background: Adult height and body-mass index influence the risk of breast cancer in women. Whether these associations reflect growth patterns of the fetus or growth during childhood and adolescence is unknown. Methods: We investigated the association between adult height and breast cancer in a cohort of 117,415 Danish women. Birth weight, age at menarche, and annual measurements of height and weight were obtained from school health records. We used the data to model individual growth curves. Results: Growth on vital status, age at first childbirth, parity, and diagnosis of breast cancer was obtained through linkages to national registries. Results: During 3,333,359 person-years of follow-up, 3340 cases of breast cancer were diagnosed. High birth weight, high stature at 14 years of age, low body-mass index (BMI) at 14 years of age, and peak growth at an early age were independent risk factors for breast cancer. Height at 8 years of age and the increase in height during puberty (8 to 14 years of age) were also associated with breast cancer. The attributable risks of birth weight, height at 14 years of age, BMI at 14 years of age, and peak growth at an early age were independent risk factors for breast cancer. Height at 8 years of age and the increase in height during puberty (8 to 14 years of age) were also associated with breast cancer. The attributable risks of birth weight, height at 14 years of age, BMI at 14 years of age, and peak growth at an early age were independent risk factors for breast cancer.

Authors' Abstract

Reason for selecting abstract:

• Importance of birth weight
• Importance of growth peak

Selected by Anthony V. Proto, MD
School of Medicine, Virginia Commonwealth University, Richmond

Cardiovascular System

Abstracts of Current Literature

Radiology

Clinic Foundation, 9500 Euclid Ave, Cleveland, OH 44195; e-mail: yadav@cf.org

Background: Carotid endarterectomy is more effective than medical management in the prevention of stroke in patients with severe symptomatic or asymptomatic atherosclerotic carotid-artery stenosis. Stenting with the use of an emboli-protection device is a less invasive revascularization strategy than endarterectomy in carotid-artery disease. Methods: We conducted a randomized trial comparing carotid-artery stenting with the use of an emboli-protection device to endarterectomy in 334 patients with coexisting conditions that potentially increased the risk posed by endarterectomy and who had either a symptomatic carotid-artery stenosis of at least 50 percent of the luminal diameter or an asymptomatic stenosis of at least 80 percent. The primary end point of the study was the cumulative incidence of a major cardiovascular event at 1 year—a composite of death, stroke, or myocardial infarction within 30 days after the intervention or death or ipsilateral stroke between 31 days and 1 year. The study was designed to test the hypothesis that the less invasive strategy, stenting, was not inferior to endarterectomy. Results: The primary end point occurred in 20 patients randomly assigned to undergo carotid-artery stenting with an emboli-protection device (cumulative incidence, 12.2 percent) and in 32 patients randomly assigned to undergo endarterectomy (cumulative incidence, 20.1 percent; absolute difference, −7.9 percentage points; 95 percent confidence interval, −16.4 to 0.7 percentage points; P = 0.004 for noninferiority, and P = 0.053 for superiority). At one year, carotid revascularization was repeated in fewer patients who had received stents than in those who had undergone endarterectomy (cumulative incidence, 0.6 percent vs. 4.3 percent; P = 0.04). Conclusions: Among patients with severe carotid-artery stenosis and coexisting conditions, carotid stenting with the use of an emboli-protection device is not inferior to carotid endarterectomy.

Authors’ Abstract

Reason for selecting abstract:
- Protected versus unprotected stent placement

Selected by Anthony V. Proto, MD
School of Medicine, Virginia Commonwealth University, Richmond

Miscellaneous

Viral Community-Acquired Pneumonia in Nonimmunocompromised Adults. Andrés A. de Roux, Maria A. Marcos, Elisa García, et al. Chest 2004; 125:1343–1351. (Antoni Torres, Respiratory Intensive Care Unit, Institut Clinic de Pneumologia i Cirurgia Toracica, escalera 2, planta 3, Hospital Clinic, Villarreal 170, Barcelona, 08036, Spain; e-mail: atorres@medicina.ub.es)

Introduction: Viral community-acquired pneumonia (CAP) has been poorly studied and clinically characterized. Using strict criteria for inclusion, we studied this type of infection in a large series of hospitalized adults with CAP.

Materials and methods: All nonimmunocompromised adult patients with a diagnosis of CAP having paired serology for respiratory viruses (RVs) [338 patients] were prospectively included in the study from 1996 to 2001 at our 1,000-bed university teaching hospital, and subsequently were followed up. We compared patients with pure viral (PV), mixed viral (RV + bacteria), and pneumococcal CAP. RVs (ie, influenza, parainfluenza, respiratory syncytial virus, and adeno) were diagnosed by means of paired serology. Results: Sixty-one of 338 patients (18%) with paired serology had an RV detected, and in 31 cases (9%) it was the only pathogen identified. Influenza was the most frequent virus detected (39 patients; 64%). Patients with chronic heart failure (CHF) had an increased risk of acquiring PV CAP (8 of 26 patients; 31%) when compared to a mixed viral/bacterial etiology (2 of 26 patients; 8%; p = 0.035) or CAP caused by Streptococcus pneumoniae (1 of 44 patients; 2%; p = 0.001). Multivariate analysis revealed that CHF (odds ratio [OR], 15.3; 95% confidence interval [CI], 1.4 to 163; p = 0.024) and the absence of expectoration (OR, 0.14; 95% CI, 0.04 to 0.6; p = 0.006) were associated with PV pneumonia compared to pneumococcal CAP. CONCLUSION: RVs are frequent etiologies of CAP (single or in combination with bacteria). Patients with CHF have an increased risk of acquiring a viral CAP.

Authors’ Abstract

Reason for selecting abstract:
- Epidemiology of community-acquired pneumonia

Selected by Robert M. Steiner, MD
Temple University Hospital, Philadelphia


Context: The success of measures to restrict smoking in indoor environments and the intensity of enforcement vary among countries around the world. In 2001, the Pan American Health Organization (PAHO) launched the Smoke-Free Americas Initiative to build capacity to achieve smoke-free environments in Latin America and the Caribbean. Objective: To assess secondhand smoke concentrations in public places in the capital cities of Argentina, Brazil, Chile, Costa Rica, Paraguay, Peru, and Uruguay in conjunction with the Smoke-Free Americas Initiative. Design and setting: Multi-country assessment of vapor-phase nicotine concentrations using a common protocol in all 7 Latin American countries. A total of 633 sampling devices were placed for 7 to 14 days in 1 hospital, 2 secondary schools, 1 city government building, 1 airport (2 in Argentina), and restaurants and bars in each country. Main Outcome Measure: Concentrations of airborne nicotine. Results: Airborne nicotine was detected in most (94%) of the locations surveyed. By country, Argentina and Uruguay had the highest median concentrations in most environments (eg, in hospitals: 1.33 [interquartile range [IQR], 0.51–3.12] μg/m³ and 0.8 [IQR, 0.30–1.69] μg/m³, respectively). Overall, bars and restaurants had the highest median concentrations (3.65 [IQR, 1.55–5.12] μg/m³ and 1.24 [IQR, 0.41–2.48] μg/m³, respectively). Nicotine concentrations were also found in a number of key, sentinel buildings, including 95% (155/163) of hospital samples (in the physicians’ and nurses’ stations the median was 0.27 [IQR, 0.02–1.94] μg/m³), schools, government buildings, and/or airports in most countries. Conclusions: The finding of airborne nicotine in critical locations in Latin America provides a basis for enforcing smoke-free initiatives and for strengthening the protection of the public from unwanted exposure to secondhand smoke.

Authors’ Abstract

Reason for selecting abstract:
- Passive smoking

Selected by Anthony V. Proto, MD
School of Medicine, Virginia Commonwealth University, Richmond
This signed form (a photocopy is acceptable) must be submitted with all proffered works.
FAX this form to the Radiology Editorial office: 804-644-5672

Transfer of Copyright and Certifications Agreement

Complete copyright to the article entitled:

is hereby transferred to the Radiological Society of North America (for U.S. government employees: to the extent transferable), effective if and when the article is accepted for publication in Radiology, printed or electronic format. In the case of authors who are officers or employees of the U.S. government, the Radiological Society of North America recognizes that works prepared by them as part of their official government duties are in the public domain.

Authors reserve the following rights:
1. Other than copyright, all proprietary rights including patent rights relating to the subject matter of the article.
2. The license for personal use in lectures, lecture notes, and exhibits.

Please return this form as soon as possible. The signed statement must be received by the Editor before the manuscript can be processed for publication.

A. All authors certify that this material or similar material has not been and will not be submitted by them or by colleagues in their institution to any other publication prior to its appearance in Radiology.
B. All authors certify that they have read the Publication Information for Authors found in the most recently published issue of Radiology at the time the manuscript is submitted. They further certify their understanding of and compliance with the various requirements found therein, in particular those relating to authorship and redundant publication and potential sanctions for redundant publication.
C. All authors certify that they have submitted all required permissions and that all required statements have been inserted in the text where appropriate.
D. All authors certify that all individuals who have participated sufficiently in the conception and design of this work, in the analysis of the data, and in the writing of the manuscript have been listed as authors. All authors further certify that they have read the final version of the manuscript and approve it for submission for publication, agree with the author contributions listed on the Manuscript Categorization Terms form, and agree with the placement of authors in the author listing.
E. All authors certify that the Editor may correspond with the designated corresponding author regarding acceptance, rejection, revision of the manuscript and that the corresponding author is responsible for informing all other authors of the same.
F. All authors certify that responses, revisions received by the Editor from the designated corresponding author will be treated as if from all authors and that the corresponding author is responsible for informing all other authors of the same.
G. All authors certify that: (must check one) (Please see Publication Information for Authors for explanation.)
   ☐ they have no direct or indirect financial interest in the products under investigation or subject matter discussed in this manuscript, or
   ☐ one or more (please name) authors do have a direct or indirect financial interest in the products under investigation or subject matter discussed in this manuscript. A separate statement is attached that details this proprietary interest, and the authors hereby grant permission to publish this information or an appropriate summary thereof with the manuscript if the manuscript is published.

TO BE SIGNED BY ALL AUTHORS:

Name (print) (First) ___________________________ Signature ___________________________
Name (print) _______________________________ Signature ___________________________
Name (print) _______________________________ Signature ___________________________
Name (print) _______________________________ Signature ___________________________
Name (print) _______________________________ Signature ___________________________
Name (print) _______________________________ Signature ___________________________

020102