HydroFrame®
HydroCoil® Embolic System

HydroFill®
HydroCoil® Embolic System

HydroSoft®
HydroCoil® Embolic System

HYDROGEL ADVANCED COILS
The clinical benefits of Hydrogel NOW engineered with the softness of bare platinum on the V-Trak® Advanced Coil System.

• CLINICALLY STUDIED HYDROGEL
• EFFECTIVE¹,²,³
• ESTABLISHED SAFETY PROFILE⁴
• SOFTER DESIGN⁵

MicroVention, Inc.
Worldwide Headquarters
1311 Valencia Avenue
Tustin, CA 92780 USA
PH +1.714.247.8000

MicroVention UK Limited
PH +44 (0) 191 258 6777

MicroVention Europe, S.A.R.L.
PH +33 (1) 39 21 77 46

MicroVention Deutschland GmbH
PH +49 211 210 798-0

microvention.com
The Trevo Retriever is indicated for use to restore blood flow in the neurovasculature by removing thrombus for the treatment of acute ischemic stroke to reduce disability in patients with a persistent, proximal anterior circulation, large vessel occlusion, and smaller core infarcts who have first received intravenous tissue plasminogen activator (IV t-PA). Endovascular therapy with the device should start within 6 hours of symptom onset.

NEW
Indication for Trevo® Retrievers
A New Standard of Care in Stroke

***
FIRST
mechanical thrombectomy device indicated to reduce disability in stroke.*

FIRST
new treatment indication for stroke in 20 years.

Trevo XP
PROVUE RETRIEVER

Success accelerated.

*The Trevo Retriever is indicated for use to restore blood flow in the neurovasculature by removing thrombus for the treatment of acute ischemic stroke to reduce disability in patients with a persistent, proximal anterior circulation, large vessel occlusion, and smaller core infarcts who have first received intravenous tissue plasminogen activator (IV t-PA). Endovascular therapy with the device should start within 6 hours of symptom onset.
MAGIC catheters are designed for general intravascular use. They may be used for the controlled, selective regional infusion of therapeutic agents or embolic materials into vessels.¹

¹ Magic Catheters IFU – Ind 19
MKTG-068 Rev. A
18 Technology Drive #169, Irvine Ca 92618
P 949.788.1443 | F 949.788.1444
The brain at your fingertips
Welcome and Greetings

Please join us in Vancouver, CANADA for the 56th Annual Meeting of the American Society of Neuroradiology on June 2–7, 2018 at the Vancouver Convention Centre East. Surrounded by the coastal mountains and located on the waterfront, you can enjoy these spectacular views in the heart of downtown Vancouver. With its undeniable charm and friendly atmosphere, Vancouver is known around the world as both a popular tourist attraction and one of the best places to live.

ASNR enthusiastically presents Neuroradiology: Adding Value and Improving Healthcare at the Symposium of the Foundation of the ASNR, as well as the common thread throughout the Annual Meeting. Implementing a value-based strategy in imaging has grasped the attention of nearly every healthcare provider; in particular with Radiologists understanding that the future will demand their imaging practices deliver better value. Value in healthcare is typically defined as those imaging strategies that yield improved outcomes, lower costs, or both. As payment transitions from a fee-for-service to a value-based system, thus creating a fundamentally different marketplace dynamic, measuring good outcomes are at the center of this changeover. At this time of uncertainty what little remains clear is that without a well-defined knowledge of their outcomes, no medical specialty will be able to succeed in the future value-based system. The Symposium will feature how Neuroradiology, in its many subspecialty areas, adds value to clinical care pathways by directing healthcare practice towards better outcomes. The annual meeting programming will continue on this theme emphasizing imaging that improves health outcomes, while considering costs, thus adding value. Our discussions will incorporate many innovative approaches to how neuroimaging currently does and will continue to improve overall healthcare performance.

As the Program Chair for ASNR 2018, it is my pleasure and honor to welcome you to Vancouver, CANADA for our annual meeting! Vancouver is known for being a very walkable city with a compact downtown core hosting many places to enjoy. So pack your comfortable walking shoes and let’s tour together with our colleagues and friends!

Pina C. Sanelli, MD, MPH, FACR
ASNR 2018 Program Chair/President-Elect
WE’VE MOVED TO FALL 2017!

11th Annual Meeting of the American Society of FUNCTIONAL NEURORADIOLOGY

October 9-11, 2017 • the Nines • Portland, Oregon

Optional Hands-on BOLD fMRI Workshop • October 8, 2017

PLEASE CONTACT EDUCATIONAL SYMPOSIA AT 813-806-1000 or ASFNR@edusymp.com or visit www.ASFNR.org for additional information.
KCR 2017
The 73rd Korean Congress of Radiology
And Annual Delegate Meeting of
The Korean Society of Radiology

2017.10.25 Wed – 2017.10.28 Sat
Coex, Seoul, Korea

Important Dates

- Abstract Submission by May 31 (Wed)
- Abstract Acceptance Notice on July 5 (Wed)
- Presenter Registration by July 31 (Mon)
- Early Registration by July 31 (Mon)
- Pre-registration by August 31 (Thu)
- Hotel Reservation by August 31 (Thu)

Hosted by The Korean Society of Radiology

www.kcr4u.org
We’re Inside Every Great Neuroradiologist!

Join the leaders in neuroradiology today!
Learn more at www.asnr.org/join

American Society of Neuroradiology
800 Enterprise Dr., Suite 205, Oak Brook, IL 60523 • (630)574-0220 • membership@asnr.org • www.asnr.org
51st Annual Meeting
American Society of
Head & Neck Radiology

Head and Neck Imaging in the City of Lights

Caesars Palace • Las Vegas, NV
September 16 – 20, 2017
(Saturday - Wednesday)

Please contact Educational Symposia at 813-806-1000 or ASHNR@edusymp.com or visit www.ASHNR.org for additional information.
**INDICATIONS FOR USE**

1. The Trevo® Retriever is indicated for use to restore blood flow in the neurovasculature by removing thrombus for the treatment of acute ischemic stroke to reduce disability in patients with a persistent, proximal anterior circulation occlusion. It is also indicated for use as a tortuous, small or smaller cross-debranching microcatheter through which a patient has been previously received intravenous tissue plasminogen activator (IV-TPA). Endovascular therapy with the device should start within 6 hours of symptom onset.

2. The Trevo® Retriever is intended for use to restore blood flow in the neurovasculature by removing thrombus in patients experiencing ischemic stroke within 6 hours of symptom onset. Patients who are eligible for intravenous tissue plasminogen activator (IV-TPA) or fail IV-TPA therapy are candidates for treatment.

**COMPICATIONS**

Procedures requiring percutaneous catheter introduction should not be attempted by physicians unfamiliar with possible complications which may be encountered during or after the procedure. Possible complications include, but are not limited to, a) following an embolic, hemitranstn or hemorrhage at puncture site; infection; distal embolization; pain; head/embolus, vessel spasm, thrombosis, dissection, or perforation; emboli; acute occlusion; ischemia; intracranial hemorrhage; false aneurysm formation; neurological deficits including stroke; and death.

**COMPATIBILITY**

Software versions are compatible with Trevo® Pro 14 Micromotors (REF 90231) and Trevo® Pro 18 Micromotors (REF 90283). 40x26mm retrievers are compatible with Trevo® Pro 18 Micromotors (REF 90283). 40x26mm retrievers are compatible with Excelsior® XT-27® Micromotors (150cm X 6cm straight (REF 275081) and Trevo® Pro 18 Micromotors (REF 90283). 4x26mm Retriever is compatible with Catalyst® 1X™ (150cm X 6mm straight (REF 275281). Compatibility of the Retriever with other microcatheters has not been established. The establishment of the Retriever device may be impacted if a different microcatheter is used.

Ballion Guide Catheters (such as Merci® Ballion Guide Catheter and FlowWire® Ballion Guide Catheter) are recommended for use during thrombectomy procedures.

Retrievers are compatible with the Abbott Vascular DCC® Guide Wire Extension (REF 22060).

Retrievers are compatible with Boston Scientific RAH (REF 621742).

**SPECIFIC WARNINGS FOR INDICATION 1**

- The safety and effectiveness of the Trevo® Retriever is reducing disability has not been established in patients with large core infarcts (i.e., ASPECTS ≤ 7). There may be increased risks, such as intracerebral hemorrhage, in these patients.

- The safety and effectiveness of the Trevo® Retriever in reducing disability has not been established to some extent with patients with occlusions in the posterior circulation (e.g., basilar or vertebral arteries) or for more distal occlusions in the anterior circulation.

**WARNINGS APPLIED TO BOTH INDICATIONS**

- Administration of IV-PA must be performed using a standard IV-PA approved window (within 6 hours of stroke onset) or as otherwise directed by the treating physician.

- Contents supplied STERILE, using an ethylene oxide (EO) process. Nonpyrogenic.

- To reduce risk of vessel damage, adhere to the following recommendations:
  - Take care to appropriately size Retriever to vessel diameter at intended site of deployment.
  - Do not use Retriever-sized (4x6) retrievals in attempts to achieve using Retriever devices.
  - Maintain Retriever position in vessel when removing or exchanging Microcatheter.
  - To reduce risk of linking/fracture, adhere to the following recommendations:
    - Immediately after unthreading Retriever, position Microcatheter tip marker just proximal to selected shape section. Maintain Microcatheter tip marker just proximal to shape section of Retriever during manipulation and withdrawal.
    - Do not resterilize or reuse. Structural integrity and/or function may be impaired by reuse or cleaning.
  - The Retriever is a delicate instrument and should be handled carefully. Before use and when possible during procedure, inspect device carefully for damage. Do not use a device that may prevent device from functioning and may cause complications.

**PRECAUTIONS**

- Prescription – only – device restricted to use by or on the order of a physician.

- Store in cool, dry, dark place.

- Do not use open or damaged packages.

- Use by "Use By" date.

- Expose to temperatures above 54°C (130°F) may damage device and accessories. Do not autoclave.

- Do not expose Retriever to solvents.

- Do not resterilize and reuse. Structural integrity and/or function may be compromised by reuse or cleaning.

- Advancing the delivery wire beyond the target microcatheter delivery wire marker may result in a stretched coil or premature detachment of the coil from the delivery wire, which could result in vessel rupture.

- Verify there is no coil loop protrusion into the parent vessel after coil placement and prior to coil detachment. Coil loop protrusion after coil placement may result in thromboembolic events if the coil is detached.

- Do not advance or withdraw Retriever against resistance or significant vasospasm.

- Advancing or torqueing against resistance or significant vasospasm may result in damage to vessel or device. Assess cause of resistance using fluoroscopy and if needed rethread the device to withdraw.

- The Retriever is difficult to withdraw from the vessel, do not torque Retriever.

- Advance Microcatheter distally, gently pull Retriever back into Microcatheter, and remove Retriever and Microcatheter as a unit. If unstable resistance is met when withdrawing the Retriever into the Microcatheter, consider extending the Retriever using the Abbott Vascular DCC® guide wire extension (REF 227261) so that the Microcatheter can be left in the patient with an extra catheter such as a DCA® catheter. Gently withdraw the Retriever into the larger diameter catheter.

- Advise anti-coagulation and anti-platelet medications per standard institutional guidelines.

**CONTRAINdications**

- Prescription – only – device restricted to use by or on the order of a physician.

- Store in cool, dry, dark place.

- Do not use open or damaged packages.

- Use by "Use By" date.

- Expose to temperatures above 54°C (130°F) may damage device and accessories. Do not autoclave.

- Do not expose Retriever to solvents.

- Do not resterilize and reuse. Structural integrity and/or function may be compromised by reuse or cleaning.

- Advancing the delivery wire beyond the target microcatheter delivery wire marker may result in a stretched coil or premature detachment of the coil from the delivery wire, which could result in vessel rupture.

- Verify there is no coil loop protrusion into the parent vessel after coil placement and prior to coil detachment. Coil loop protrusion after coil placement may result in thromboembolic events if the coil is detached.

- Do not advance or withdraw Retriever against resistance or significant vasospasm.

- Advancing or torqueing against resistance or significant vasospasm may result in damage to vessel or device. Assess cause of resistance using fluoroscopy and if needed rethread the device to withdraw.

- If resistance is difficult to withdraw from the vessel, do not torque Retriever.

- Advance Microcatheter distally, gently pull Retriever back into Microcatheter, and remove Retriever and Microcatheter as a unit. If unstable resistance is met when withdrawing the Retriever into the Microcatheter, consider extending the Retriever using the Abbott Vascular DCC® guide wire extension (REF 227261) so that the Microcatheter can be left in the patient with an extra catheter such as a DCA® catheter. Gently withdraw the Retriever into the larger diameter catheter.

- Advise anti-coagulation and anti-platelet medications per standard institutional guidelines.

**PRECAUTIONS**

- Prescription – only – device restricted to use by or on the order of a physician.

- Store in cool, dry, dark place.

- Do not use open or damaged packages.

- Use by "Use By" date.

- Expose to temperatures above 54°C (130°F) may damage device and accessories. Do not autoclave.

- Do not expose Retriever to solvents.

- Do not resterilize and reuse. Structural integrity and/or function may be compromised by reuse or cleaning.

- Advancing the delivery wire beyond the target microcatheter delivery wire marker may result in a stretched coil or premature detachment of the coil from the delivery wire, which could result in vessel rupture.

- Verify there is no coil loop protrusion into the parent vessel after coil placement and prior to coil detachment. Coil loop protrusion after coil placement may result in thromboembolic events if the coil is detached.

- Do not advance or withdraw Retriever against resistance or significant vasospasm.

- Advancing or torqueing against resistance or significant vasospasm may result in damage to vessel or device. Assess cause of resistance using fluoroscopy and if needed rethread the device to withdraw.

- If resistance is difficult to withdraw from the vessel, do not torque Retriever.

- Advance Microcatheter distally, gently pull Retriever back into Microcatheter, and remove Retriever and Microcatheter as a unit. If unstable resistance is met when withdrawing the Retriever into the Microcatheter, consider extending the Retriever using the Abbott Vascular DCC® guide wire extension (REF 227261) so that the Microcatheter can be left in the patient with an extra catheter such as a DCA® catheter. Gently withdraw the Retriever into the larger diameter catheter.

- Advise anti-coagulation and anti-platelet medications per standard institutional guidelines.
The V-Trak® Advanced Coil System, the next generation to power the performance of our most technically advanced line of coils. Offering the optimal combination of support and flexibility.
Smooth and stable.

Target Detachable Coils deliver consistently smooth deployment and exceptional microcatheter stability. Designed to work seamlessly together for framing, filling and finishing. Target Coils deliver the high performance you demand.

For more information, please visit www.strykerneurovascular.com/Target or contact your local Stryker Neurovascular sales representative.
CREDO® Stent only available within ASSISTENT

Unique treatment concept comprising a flexible self-expanding stent deliverable through a compliant PTA balloon catheter for gentle, efficient and timesaving treatment of intracranial stenosis.

CREDO® Stent with NeuroSpeed® PTA Balloon Catheter

GENTLE VARIABILITY

VARIABLE. EFFICIENT. GENTLE.

www.acandis.com
Abnormalities in congenital Zika syndrome with and without arthrogryposis

Thrombus fragmentation during endovascular stroke treatment

Imaging characteristics of diffuse intracranial dolichoectasia
Perspectives

J.H. Rees

852 Review Article
Benign Spine Lesions: Advances in Techniques for Minimally Invasive Percutaneous Treatment
A. Tomasian, et al.

General Contents

862 Brain Perivascular Spaces as Biomarkers of Vascular Risk: Results from the Northern Manhattan Study
J. Gutierrez, et al.

868 Site and Rate of Occlusive Disease in Cervicocerebral Arteries: A CT Angiography Study of 2209 Patients with Acute Ischemic Stroke
D.C. Rotzinger, et al.

875 Ipsilateral Prominent Thalamostriate Vein on Susceptibility-Weighted Imaging Predicts Poor Outcome after Intravenous Thrombolysis in Acute Ischemic Stroke
X. Zhang, et al.

882 Restriction Spectrum Imaging Improves Risk Stratification in Patients with Glioblastoma
A.P. Krishnan, et al.

890 A Multiparametric Model for Mapping Cellularity in Glioblastoma Using Radiographically Localized Biopsies
P.D. Chang, et al.

899 Multiparametric Evaluation in Differentiating Glioma Recurrence from Treatment-Induced Necrosis Using Simultaneous 18F-FDG-PET/MRI: A Single-Institution Retrospective Study
A. Jena, et al.

905 Classification of High-Grade Glioma into Tumor and Nontumor Components Using the Support Vector Machine
D.T. Blumenthal, et al.

915 Clinical and Imaging Characteristics of Diffuse Intracranial Dolichoectasia
W. Brinjikji, et al.

923 Detection of Focal Longitudinal Changes in the Brain by Subtraction of MR Images
N. Patel, et al.

928 Acute Cytotoxic and Vasogenic Edema after Subarachnoid Hemorrhage: A Quantitative MRI Study

935 Progressive Cortical Neuronal Damage and Extracranial-Intracranial Bypass Surgery in Patients with Misery Perfusion
H. Yamauchi, et al.
Cognitive Implications of Deep Gray Matter Iron in Multiple Sclerosis
E. Fujiwara, et al.

Enhancing Brain Lesions during Acute Optic Neuritis and/or Longitudinally Extensive Transverse Myelitis May Portend a Higher Relapse Rate in Neuromyelitis Optica Spectrum Disorders
G. Orman, et al.

Can MRI Visual Assessment Differentiate the Variants of Primary-Progressive Aphasia?
S.A. Sajjadi, et al.

Entorhinal Cortex: Antemortem Cortical Thickness and Postmortem Neurofibrillary Tangles and Amyloid Pathology
A.A. Thaker, et al.

Combined Diffusion Tensor Imaging and Apparent Transverse Relaxation Rate Differentiate Parkinson Disease and Atypical Parkinsonism
G. Du, et al.

Performance Assessment for Brain MR Imaging Registration Methods
J.S. Lin, et al.

Quantitative Assessment of Variation in CT Parameters on Texture Features: Pilot Study Using a Nonanatomic Phantom
K. Buch, et al.

Quantifying Intracranial Internal Carotid Artery Stenosis on MR Angiography
H. Baradaran, et al.

Risk of Thrombus Fragmentation during Endovascular Stroke Treatment
J. Koemacher, et al.

Selective-versus-Standard Poststen Dilation for Carotid Artery Disease: A Systematic Review and Meta-Analysis
O. Petr, et al.

Presurgical Brain Mapping of the Ventral Somatomotor Network in Patients with Brain Tumors Using Resting-State fMRI
N. Yahyavi-Firouz-Abadi, et al.

Prognostic Predictors of Visual Outcome in Open Globe Injury: Emphasis on Facial CT Findings
U.K. Bodanapally, et al.

MRI-Based Texture Analysis to Differentiate Sinonasal Squamous Cell Carcinoma from Inverted Papilloma
S. Ramkumar, et al.

Trochlear Groove and Trochlear Cistern: Useful Anatomic Landmarks for Identifying the Tentorial Segment of Cranial Nerve IV on MRI
P.M. Bunch, et al.

Hindbrain Herniation in Chiari II Malformation on Fetal and Postnatal MRI

Brain Development in Fetuses of Mothers with Diabetes: A Case-Control MR Imaging Study
F.C. Denison, et al.

Spectrum of Spinal Cord, Spinal Root, and Brain MRI Abnormalities in Congenital Zika Syndrome with and without Arthrogryposis
M.F.V.V. Aragao, et al.

Intraforaminal Location of Thoracolumbar Radicular Arteries Providing an Anterior Radiculomedullary Artery Using Flat Panel Catheter Angiography
L. Gregg, et al.

Quantitative Measurement of CSF in Patients with Spontaneous Intracranial Hypotension
H.-C. Chen, et al.

35 YEARS AGO IN AJNR
ONLINE FEATURES

LETTERS

E30 Regarding “MR Imaging of the Cervical Spine in Nonaccidental Trauma: A Tertiary Institution Experience”  
X. Wu, et al.

E31 Reply  
T.N. Booth, et al.

E32 Intracranial Arteriovenous Shunting Detection with Arterial Spin-Labeling and Susceptibility-Weighted Imaging: Potential Pitfall of a Venous Predominant Parenchymal Arteriovenous Malformation  
S.A. Nabavizadeh

E33 Reply  
J. Hodel, et al.

E34 3D Ultrasound for Imaging and Quantifying Carotid Ulcers  
J.D. Spence

E37 Reply  
J. Yuan, et al.

BOOK REVIEWS R.M. Quencer, Section Editor

Please visit www.ajnrblog.org to read and comment on Book Reviews.

ADC, FLAIR, T1 C+ images (left) are utilized to estimate tumor cellularity (middle). Histopathology (right) shows variation in cellularity around the area of tumor enhancement.
Title: Big Hole. Big Hole is a sidewater loop of the Myakka River southeast of Sarasota, Florida in the Myakka River State Park. For unknown reasons, it is an extremely popular area for "alligator lounging" with as many as 70–80 gators often visible at any one time. This, coupled with its natural beauty, makes the 2.5-mile hike popular among those seeking to refresh themselves “far from the things of man!”

John H. Rees, Chief, Neuroradiology, Partners Imaging Center, Bradenton, Florida
Benign Spine Lesions: Advances in Techniques for Minimally Invasive Percutaneous Treatment

A. Tomasian, A.N. Wallace, and J.W. Jennings

ABSTRACT

SUMMARY: Minimally invasive percutaneous imaging-guided techniques have been shown to be safe and effective for the treatment of benign tumors of the spine. Techniques available include a variety of tumor ablation technologies, including radiofrequency ablation, cryoablation, microwave ablation, alcohol ablation, and laser photocoagulation. Vertebral augmentation may be performed after ablation as part of the same procedure for fracture stabilization or prevention. Typically, the treatment goal in benign spine lesions is definitive cure. Painful benign spine lesions commonly encountered in daily practice include osteoid osteoma, osteoblastoma, vertebral hemangioma, aneurysmal bone cyst, Paget disease, and subacute/chronic Schmorl node. This review discusses the most recent advancement and use of minimally invasive percutaneous therapeutic options for the management of benign spine lesions.

ABBREVIATIONS: PVA = percutaneous vertebral augmentation; RF = radiofrequency; RFA = radiofrequency ablation

The role of minimally invasive percutaneous imaging-guided techniques for the treatment of benign tumors of the spine has increased during the past decade. Although previously limited by the complex anatomy of the vertebral column and the proximity of neural elements, new technologies, including radiofrequency ablation (RFA), cryoablation, microwave ablation, alcohol ablation, and laser photocoagulation, now provide an attractive alternative or adjunct therapeutic options for the treatment of benign spinal tumors beyond medical pain management, surgery, radiation therapy, and standard vertebral augmentation.

Benign tumors compose only 4%–13% of spinal lesions and are treated with curative intent. Most commonly, the pain secondary to benign spine lesions is managed with nonsteroidal anti-inflammatory drugs and opioids titrated to achieve pain relief while attempting to minimize side effects. In the absence of neurologic deficits and spinal instability, radiation therapy may be the standard of care when medical management is insufficient. Although described in malignant spine lesions, pain relief following radiation therapy may be delayed and transient. Surgical interventions are invasive and result in increased morbidity, which, in the absence of pending neurologic compromise and instability, is undesirable for these patients.

Minimally invasive, percutaneous, image-guided interventions for the management of benign spine lesions are indicated when pharmacologic therapy (analgesics and nonsteroidal anti-inflammatory drugs) is inadequate or contraindicated (such as side effects and pregnancy), radiation therapy is contraindicated or not desired by the patient, and surgical intervention is not recommended (absence of spinal instability and neurologic compromise). The contraindications include spinal instability with or without concomitant pathologic fracture, focal neurologic deficit, advanced coagulopathy, and systemic infection. Although no published literature exists directly comparing an operation with percutaneous ablation, these minimally invasive percutaneous procedures may be more cost-effective compared with traditional surgical interventions. This review details the armamentarium available and the most recent advances in minimally invasive, image-guided percutaneous techniques for the treatment of benign spinal lesions.

General Considerations, Procedural Setup, and Patient Care

Preprocedural consultation with the patient is mandatory and should include an explanation of the benefits, risks, potential complications, and alternative treatment options. In addition, the patient’s clinical status and procedural indications should be discussed with the referring clinicians. Initial preprocedural work-up includes a history and physical examination to confirm focal pain/tenderness at the lesion site and neurologic examina-
tion to assess potential focal neurologic deficits. Laboratory tests are used to evaluate the coagulation status (platelet count of >50,000 per microliter and international normalized ratio of <1.5 are generally acceptable) and, at times, the possibility of systemic infection (complete blood count, erythrocyte sedimentation rate, C-reactive protein level, and cultures, if clinically indicated). Thermal ablation of bone and soft-tissue tumors is painful and requires moderate sedation or general anesthesia, in particular with technically challenging spinal ablations, to minimize the risk of voluntary patient movement and potential subsequent complications.

Preprocedural imaging includes dedicated weight-bearing spine radiographs, CT, and/or MR imaging. Radiographs are used to determine the presence and extent of vertebral compression fractures, kyphosis, scoliosis, the integrity of the osseous central canal, and the suitability of fluoroscopy for imaging guidance. CT is used to determine the bone lesion density (osteolytic, osteoblastic, or mixed), which has implications for the choice of thermal ablation technique, and to identify cortical discontinuities due to tumor erosion and pathologic fracture clefts, which are sources of pain and sites where cement is most likely to leak during postablation cementoplasty. MR imaging is important for delineating the extent of the tumor, including the presence of marrow replacement and extraosseous tumor extension that may be occult on CT. Pretreatment imaging also provides information regarding procedural planning in terms of the choice of thermal ablation technology, approach, probe placement, and the need for thermal protection and cementation. Vertebral augmentation is often performed under fluoroscopic guidance, which is readily available and allows near-real-time monitoring of cement distribution. Thermal ablation of benign lesions of the spine is most commonly performed under CT guidance. Postprocedural care should include a follow-up physical examination and questionnaire to determine the efficacy of treatment and identify potentially recurrent symptoms. For benign spine lesions, contrast-enhanced MR imaging is the technique of choice for posttreatment imaging and is performed at the discretion of the referring physician to evaluate treatment adequacy or when symptoms recur. Postprocedural imaging should be performed 6–8 weeks after treatment to allow ablation-related inflammation to subside.

**Minimally Invasive Percutaneous Techniques**

**Vertebral Augmentation.** Percutaneous vertebral augmentation (PVA) encompasses several techniques aimed at internal vertebral body stabilization with bone cement. Traditional vertebroplasty involves instillation of methyl methacrylate cement directly into the vertebral body, while kyphoplasty involves first creating a cavity or several cavities with inflatable balloons, bone tamps, or osteotomies to attempt more controlled cement delivery and improved cement interdigitation.8–11 Although both vertebroplasty and kyphoplasty offer durable pain relief, improved quality of life, and improved spine alignment, no randomized clinical trials exist for direct comparison of these techniques.12 Indirect comparison of these techniques based on systematic reviews shows no substantial difference in patient outcome.12 However, the rate of asymptomatic cement leakage is higher in vertebroplasty.12 Bone cement is delivered under real-time conventional fluoroscopy or CT-fluoroscopic guidance to minimize cement leakage by using a transpedicular approach. Several randomized controlled trials have demonstrated the safety, efficacy, and durability of PVA for the management of benign and malignant spine compression fractures.8–11 It is the position of multiple radiologic and neurosurgical societies that PVA remains an appropriate therapy for the treatment of painful vertebral compression fractures refractory to nonoperative medical therapy and for vertebrae weakened by neoplasia when performed for the medical indications outlined in the published standards.9

Recently, radiofrequency (RF)-induced high-viscosity cement has been used for PVA to potentially reduce intervertebral disc and venous cement leakage.13,14 Anselmetti et al13 performed PVA in 60 patients (190 vertebrae) and reported asymptomatic venous leak in 8.2% versus 41.3% of patients and intervertebral disc cement leakage in 6.1% versus 13% of patients (high-viscosity cement versus standard viscosity cement). Several mechanisms have been postulated to contribute to pain relief following PVA, including stabilization of vertebral body macro- and microfractures, exothermic reaction associated with the setting of cement to thermally ablate nerve endings, and an possible embolization effect with reduction of vascular congestion due to cement filling of enlarged bone marrow spaces.

**Radiofrequency Ablation.** In RFA, a RF generator is used to deliver high-frequency, alternating current (375–600 kHz) to the patient through an RF probe. The current passes through the exposed active tip of the probe and results in oscillation of charged tissue molecules (ions) within the ablation zone, producing frictional heat. The thermal effect depends on the electrical conducting properties of the treated tissue and the characteristics of the RF probe.15 When local tissue temperature between 60°C and 100°C is reached, there is protein denaturation and immediate coagulative necrosis.15 Several RFA probe-design technologies exist, attempting to achieve an improved ablation zone.15 Unipolar systems use dispersing grounding pads placed on patient skin near the ablation site to serve as the receiving limb of the electrical circuit to prevent potential skin burns while bipolar systems use built-in transmitting and receiving electrical elements within the probe; this feature eliminates the need for grounding pads and the possibility of skin burns.15 Internally cooled probes decrease tissue char (carbonization) by maintaining a lower temperature surrounding the active probe tip.15

More recently navigational bipolar RFA probes with built-in thermocouples have become available, which allow real-time monitoring of the ablation zone size by measuring the temperatures along the periphery of the ablation zone during the procedure.16 The navigating tip of the probe can be articulated in different orientations through the same entry site; this procedure is beneficial for accessing lesions in challenging locations and achieving larger ablation zones.16 The choice of the RF probe depends in large part on the volume of tissue to be ablated and the proximity to vital structures.15,16 RFA is typically used for the treatment of vertebral lesions with no or small extraosseous components. The main advantage of RFA is precise determination of the geometry of the ablation zone beyond which tissues are safe from thermal injury.15,16 Intact cortical bone also serves as a boundary for undesired RF energy propagation.15,16 RFA is pri-
marily used for the treatment of lesions that are mainly osteolytic because the higher intrinsic impedance of sclerotic bone lesions prevents the radiofrequency circuit from generating sufficiently high temperatures to ensure cell death and renders RFA ineffective.\textsuperscript{17} Limitations of RFA include nonvisualization of the ablation margin with CT, pain associated with the procedure, and, frequently, increased pain during the immediate posttreatment period.

**Cryoadablation.** In closed cryoadablation systems used in percutaneous tumor ablation, a liquid gas, commonly argon, is used to rapidly cool the tip of the cryoprobe (taking advantage of the Joule-Thomson effect—that is, pressurized gas, when allowed to expand, results in a drop in temperature), forming an enlarging ice ball with time followed by a “thawing” phase, commonly achieved with helium gas, resulting in an osmotic gradient.\textsuperscript{15,18} Formation of extracellular ice results in a relative imbalance of solutes between the intra- and extracellular environment, subsequent intracellular water extraction by osmosis, and cellular dehydration. The increase in intracellular concentration of solutes results in damage to both the enzymatic machinery of the cell and the cell membrane.\textsuperscript{15,18} The formation of intracellular ice crystals damages the cellular organelles. Thawing results in an osmotic gradient and consequent cell membrane injury.\textsuperscript{15,18} A temperature of \(-40^\circ\text{C}\) or lower is necessary to ensure complete cell death.\textsuperscript{15} At present, the cryoprobes of 2 major manufacturers are used in spines with diameters of 1.2 mm (17-ga) and 1.7 mm (13-ga), generating predictable ablation zones using at least 1 freeze/active thaw/freeze cycle (typically 1/5/10 minutes).\textsuperscript{19} Cryoadablation is typically used for benign lesions with large soft-tissue components or large lesions involving the posterior vertebral elements and is preferred to RFA for the management of osteoblastic lesions. Advantages of cryoadablation include formation of a hypoattenuating ice ball, which is readily identified by CT,\textsuperscript{20} beyond which tissues are safe from thermal injury; decreased intraprocedural and postprocedural pain compared with RFA or microwave ablation; and the ability to use multiple probes in various orientations to achieve additive overlapping ablation zones. An interprobe distance of 1.5–2 cm is typically recommended with probe tips at bone-tumor or tumor–soft-tissue interfaces.\textsuperscript{19} In addition, the outer margin of the ice ball corresponds to 0°C, which is usually not sufficient for permanent tumor destruction. Reliable cell death is achieved at 3 mm from the outer edge; therefore, the ice ball should extend beyond the tumor margins.\textsuperscript{18} MR imaging–compatible cryoprobes offer an alternative imaging technique, which may permit safe ablation of spinal lesions, given the proximity of the neural elements.\textsuperscript{21}

**Laser Ablation.** Laser ablation uses optical fibers to transmit infrared light energy into a tumor to produce rapid temperature elevation, protein denaturation, and coagulative necrosis. A continuous wave semiconductor diode laser with an 805-nm wavelength delivers energy to the tumor by using a flexible single-use bare-tipped 400-μm optical fiber with polymer cladding placed coaxially.\textsuperscript{22–25} The amount of energy to be delivered is calculated according to the formula \([\text{Nidus Size (mm)} \times 100 \text{ Joules} + 200 \text{ Joules}]\), and the duration of the ablation is typically 200–600 seconds, depending on the nidus size.\textsuperscript{22–25} The size of the induced ablation zone depends on the laser wavelength, thermal and optical properties of the tissue, total duration of energy delivery, laser power, diameter of the laser fiber, and the number of fibers used.\textsuperscript{26} The maximum coagulation effect is reached at 1000–1200 J, and more energy at the same location does not increase the volume of coagulation.\textsuperscript{26} Spine laser ablation has been mainly used in the treatment of osteoid osteomas and osteoblastomas.\textsuperscript{22–25} and an energy amount of 1200 J is usually adequate. However, for larger lesions, a greater amount of energy may be necessary.\textsuperscript{22} The advantages of laser ablation include a predictable size of the ablation in proportion to the energy delivered, lack of the need to use a neutral electrode, lack of interaction with stimulators and pacemakers, and the relatively low cost of disposable laser fiber.\textsuperscript{22–25} A disadvantage is the lack of visualization of the ablation zone with CT.

**Microwave Ablation.** Microwave ablation uses antennae to deliver electromagnetic microwaves (approximately 900 MHz) to target tissue, which result in agitation of ionic molecules and frictional heat and, subsequently, tissue coagulative necrosis. It is theorized that microwave ablation is less influenced by variable tissue impedance and perfusion-mediated tissue cooling; this feature potentially results in higher intratumoral temperatures and creates a larger, more uniform ablation zone and faster ablation times by using a single probe.\textsuperscript{27} Additional advantages of microwave ablation include efficacy in the management of osteoblastic lesions due to less susceptibility to increased impedance of attenuated bone, diminished heat sink phenomena, and lack of the need for grounding pads and thus diminished risk of skin burns.\textsuperscript{27–29} The antenna shaft cooling system implemented in the latest generations of microwave ablation equipment eliminates the risk of back-heating phenomena.\textsuperscript{28,29} The ability to rapidly deliver high amounts of power (up to 100 Watts) to large ablation zones might be a disadvantage when applied to spinal lesions because surrounding overheating could potentially lead to injuries to adjacent neural elements.\textsuperscript{28,29} Although hypoattenuating ablated tissue is often identified on CT, the margins of the ablation zone are not well-defined; this feature is considered a disadvantage of spinal microwave ablation. The literature regarding the efficacy and safety of spine microwave ablation is sparse, and this technology should be used with caution for spinal applications. While 2 retrospective studies have shown its safety and efficacy for the management of spine metastases,\textsuperscript{28,29} this technology has not been used for the management of benign spine lesions.

**Alcohol Ablation.** Alcohol ablation is a relatively inexpensive percutaneous tumor ablation technique that causes tumor necrosis directly through cellular dehydration and indirectly through vascular thrombosis and tissue ischemia.\textsuperscript{30} Instillation of iodinated contrast or venography, before alcohol ablation, to delineate the extent of ethanol diffusion has been described as a strategy for reducing the risk of the vascular intravasation and potential injury to adjacent structures.\textsuperscript{30–32} In alcohol ablation, volumes of 3–30 mL are directly injected into the target tissue.\textsuperscript{30–32} The most important limitation of alcohol ablation is the poorly predictable size and configuration of the ablation zone due to poorly reproducible diffusion of alcohol through tumoral (particularly cortical) bone and peritumoral tissues.\textsuperscript{30}
Benign Spinal Lesions

Osteoid Osteoma. Osteoid osteoma is a benign painful bone-forming lesion that typically occurs in patients younger than 30 years of age with a male predilection (2–4:1). In the spine, these lesions classically involve the posterior elements and in 60% of cases are located in the lumbar spine. A central nidus with or without central mineralization (typically <15 mm) and surrounding osseous sclerosis are the typical imaging manifestations. Osteoid osteomas have been traditionally treated with surgical excision, which has substantial morbidity, including spinal instability, nerve injury, infection, and blood loss as well as cost burden. Spinal osteoid osteomas account for 10%–20% of cases, and although spontaneous regression of these tumors has been reported, 70% of untreated patients develop painful scoliosis. The entire nidus of osteoid osteoma must be ablated to ensure complete treatment. Due to the proximity of neural elements, percutaneous ablation of spinal osteoid osteoma may require logistic considerations. However, thermal protective effects of intact cortical bone, flow of CSF, and small vessels in the epidural space have been hypothesized.

Given the small size of spinal osteoid osteomas (typically <15 mm), RFA and laser ablation are recommended as treatment options with no need for cementation. Both RFA and laser ablation have similar efficacy and safety profiles, and the choice is at the discretion of the operator. Imaging guidance is typically achieved with CT with thin sections (1–2 mm) and 3D reconstructions to allow precise positioning of the probe at the center of the nidus. Since the initial reports of RFA for the treatment of osteoid osteoma, this technique has almost completely supplanted surgical resection. If RFA is to be used, monopolar noncooled straight electrodes are recommended with application of RF energy to achieve temperature of 90°C for 5–6 minutes (Fig 1). Recently, navigational bipolar RFA probes have been successfully used to treat spinal osteoid osteomas (Fig 2). These probes allow real-time monitoring of the ablation zone size by using built-in thermocouples and are beneficial for larger lesions and challenging locations due to the articulating tip, which may be placed in various orientations by using a single skin-entry site, obviating grounding pads. Adequate ablation is achieved in 2–3 minutes. Investigators have successfully used RFA for management of spinal osteoid osteomas by using thermoprotective techniques for lesions as close as 1 mm to neural elements without complications (Figs 1 and 2). During RFA, patients under general anesthesia have a fairly predictable response with elevated heart (average increase, 40%) and respiratory rates (average increase, 50%) during both the biopsy and ablation portions of the treatment.

Morassi et al treated 13 patients with spinal osteoid osteomas (11 in the posterior elements and 2 in the vertebral bodies) by using a non-cool-tip unipolar system with a 5-mm active tip (90°C for 6 minutes) using thermoprotection, achieving pain relief in 11 patients with no complications. In case of laser ablation, typically a single flexible bare-tip laser fiber is adequate; it is placed in the center of the nidus coaxially through an 18-ga spinal needle, with a delivered energy of approximately 1200 J for 200–600 sec-
onds. However, for larger lesions, 2 fibers and larger energy amounts may be necessary to ensure the adequacy of the ablation. Tsoumakidou et al reported successful laser ablation of spine osteoid osteomas in 57 patients (vertebral body, n = 18, and posterior elements/facet joints, n = 36) by using thermoprotection with most lesions located closer than 5 mm to the neural structures. Sixty-one ablations (mean delivered energy, 1271 J) were performed with a technical success rate of 100% and a primary clinical success rate of 98.2% (at 1 month) with no major complications. Cryoablation has been successfully used for management of a spine osteoid osteoma.

Osteoblastoma. Osteoblastomas are rare benign tumors with striking histologic similarity to osteoid osteomas. There is a male predilection (2.5–1). They are typically larger (>2 cm) and expansile with less sclerotic components compared with osteoid osteomas with thin peripheral sclerosis and may be associated with aneurysmal bone cyst. On MR imaging, they demonstrate avid osseous and extraosseous enhancement. Spinal osteoblastomas compose approximately 40% of cases and often involve the posterior elements and are located in the cervical spine in 10%–40% of cases. Most untreated lesions lead to painful scoliosis, and surgical excision has traditionally been the treatment of choice for spinal osteoblastomas, which is associated with morbidity, especially given the size of the osseous defect. The entire osseolytic component and the soft-tissue component (if present) must be ablated for definitive cure. The imaging guidance (CT) and thermoprotective measures are similar to those in osteoid osteoma ablation. Typically, there is no need for postablation vertebral augmentation. Given the larger size of osteoblastomas and particularly with involvement of posterior elements and potential soft-tissue components, cryoablation could be safely and efficiently performed with simultaneous placement of multiple probes with visualization of the hypoattenuating ice ball.

Both RFA and laser ablation are also recommended for management of these lesions. In case of RFA, considering lesion size, ≥2 ablations with straight unipolar probes should be performed to cover the entire lesion (90°C for 6 minutes each). Alternatively, a navigational bipolar RFA probe could be used to articulate the probe tip in different orientations through a single entry site, with precise determination of the size of the ablation zone by using built-in thermocouples. Although the RF ablation zone could not be visualized on CT, the intact surrounding cortical bone serves as a barrier for undesired RF energy propagation. Similarly, with laser ablation, at least 2 probes and more energy deposition are required to achieve a cure (Fig 3). Substantial inflammatory reaction may occur following thermal ablation of osteoblastomas, and nonsteroidal anti-inflammatory drugs may be administered for management. Weber et al successfully managed 2 spinal osteoblastomas with a unipolar cool-tip RFA system (90°C for approximately 400 seconds) by using >1 ablation to cover the entire nidus with no complications.

Aneurysmal Bone Cyst. Aneurysmal bone cyst is a benign expansive lesion of uncertain etiology comprising numerous blood-filled channels. It occurs in patients younger than 20 years of age in 80% of cases. Aneurysmal bone cyst involves the vertebral column in 20%–30% of cases (typically the posterior elements), of which about 40% involve the vertebral body. Aneurysmal bone cysts are expansive osteolytic lesions with fluid-fluid levels on MR imaging and thin peripheral/septal enhancement. These lesions may result in pain and neurologic compromise, prompting treatment. Standard surgical management depends on the lesion size and includes curettage, bone grafting, and internal fixation as well as en bloc or wide excision. Although no large studies exist to evaluate the efficacy of thermal ablation for the treatment of spinal aneurysmal bone cysts, given their larger size, typical involvement of posterior elements, and possible soft-tissue components, cryoablation is suggested as the thermal ablation technique of choice, and thermoprotection is recommended in all cases (Fig 4). Preablative embolization is also suggested to reduce the risk of hemorrhage and heat sink effect in large lesions. In case of pathologic fracture or extensive involvement of the vertebral body or pedicles, postablation cementation is recommended for stabilization. Griaudze et al reported the case of a spinal aneurysmal bone cyst involving the posterior elements of C6–T1 and the right C8 nerve root managed by cryoablation 4 days following embolization with N-butyl cyanoacrylate and lipiodol. The patient developed sensory neuropathy, which improved at 16-month follow-up, and imaging demonstrated no evidence of recurrence.

Hemangioma. Vertebral hemangiomas are the most common benign vertebral neoplasms, with a slight female predilection, with most located in the thoracic spine. Most of these lesions are asymptomatic and incidentally identified on imaging. Classic imaging features include a polka-dotted appearance on CT and T1
Hyperintensity and enhancement on MR imaging. Painful vertebral hemangiomas are typically due to impingement of the central canal or neuroforamina or concomitant pathologic fractures. Surgical intervention for vertebral hemangiomas is reserved for patients with neurologic compromise and spinal instability, and radiation therapy is used when immobilization fails to resolve neurologic symptoms or in cases of lesion progression following immobilization.47,48 PVA has been successfully used for minimally invasive management of symptomatic vertebral hemangiomas in the absence of neurologic compromise and spinal instability.47,48 PVA is associated with substantially less morbidity compared with an operation and radiation therapy and is more cost-effective. Liu et al,48 treated 33 symptomatic vertebral hemangiomas in 31 patients with PVA (mean follow-up, 15.8 months) and evaluated the clinical effects by using the Visual Analog Scale and Roland-Morris Disability Questionnaire at 1 week, 1 month, and final follow-up. The authors reported substantial improvement in the Visual Analog Scale and Roland-Morris Disability Questionnaire scores and no major complications. Aggressive spinal hemangiomas, which are associated with extraosseous epidural and paravertebral components, compose approximately 1% of spinal hemangiomas and can be successfully treated with PVA.49,50 In a case series of 16 patients, Cloran et al49 treated 4 patients with symptomatic aggressive vertebral hemangiomas without neurologic compromise by using PVA, achieving complete pain resolution in all patients with no major complications.

Alcohol ablation followed by PVA has also been successfully used for management of symptomatic vertebral hemangiomas, including aggressive hemangiomas (Fig 5).30-32 As indicated previously, preablation venography with iodinated contrast should be performed by placement of a needle within the vertebral body to delineate the extent of ethanol diffusion to reduce the risk of the vascular intravasation and potential injury to adjacent structures. Doppman et al32 treated 11 patients with vertebral hemangiomas and neurologic symptoms using alcohol ablation with no immediate complications. Radiculopathy improved in 4 of 5 patients. Two patients who received the largest volumes of ethanol, 42 and 50 mL, developed pathologic compression fractures 4 and 16 weeks following treatment, respectively.32

**Paget Disease.** Approximately 35%–50% of patients with Paget disease have spinal involvement, half of whom experience back pain.51 It can be seen in up to 10% of patients older than 80 years of age. The classic features of vertebral Paget disease include a "picture frame" appearance and squaring of vertebral bodies. Findings on MR imaging can be variable on the basis of the disease phase. Conservative management of painful vertebral Paget disease includes a combination of bisphosphonate therapy and analgesics.51-54 Patients with spinal instability and concomitant neurologic compromise are managed by surgical intervention. Patients refractory to conservative management, in the absence of

---

**FIG 4.** A 17-year-old boy with low back pain due to a right L4 aneurysmal bone cyst. A, Axial T2-weighted image demonstrates an expansile lesion in the right L4 transverse process, pedicle, and posterior vertebral body with multiple fluid-fluid levels, most compatible with aneurysmal bone cyst. B, Axial CT image during cryoablation of the transverse process portion of the lesion. Note the hypoattenuating ice ball (white arrows) extending beyond portions of the lesion into the soft tissues. A spinal needle (black arrow) was placed into the epidural space for temperature monitoring and carbon dioxide injection. C, Axial contrast-enhanced CT scan demonstrates an aggressive T3 hemangioma. A, Axial contrast-enhanced CT scan demonstrates an aggressive T3 hemangioma with intraosseous and extraosseous components. B, Axial prone intraprocedural CT scan during a venogram performed via an 18-gauge needle in the left pedicle/vertebral body. Note extensive vascularity in the vertebral body and in the soft tissues surrounding the vertebral body, including the epidural space. C, Sagittal maximum-intensity-projection CT of a T3 venogram before alcohol ablation demonstrates a vascular lesion with epidural flow. Ethanol ablation was followed by vertebral augmentation for stabilization of the vertebral body. D, Post-alcohol ablation and vertebral augmentation CT demonstrates cement filling the vertebral body and the hyperattenuating ablated extraosseous component posterior to the vertebral body (white arrow).

**FIG 5.** A 46-year-old man with upper back pain due to an aggressive T3 hemangioma. A, Axial T2-weighted image 3 months post-treatment demonstrates near-complete resolution of the expansile component of the tumor (white asterisk) and hypointense cement (white arrow) in the right pedicle and posterior vertebral body.
spinal instability and neurologic deficit, have been successfully managed by PVA with no major complications (Fig 6).51-54 Pedicelli et al51 treated a patient with painful L4 Paget disease refractory to opioid analgesics and bisphosphonates with PVA and achieved substantially improved Visual Analog Scale pain and Roland-Morris Disability Questionnaire scores at day 1 and 6-month follow-up, with no complications.

Schmorl Nodes. Schmorl nodes are common in all ages and may be associated with trauma. Symptomatic Schmorl nodes reflect acute or subacute vertebral endplate fracture, typically along the posterior margin, allowing vertical disc herniation and nuclear migration.55 Schmorl nodes as pain generators and their management remain a controversial topic. In cases of symptomatic Schmorl nodes, fluid-sensitive MR imaging sequences typically demonstrate hyperintensity along the node and vertebral endplate, which has been hypothesized as the source of pain.55 PVA may be an alternative option for the management of painful Schmorl nodes and degenerative Modic changes of the spine refractory to analgesics and conservative management (Fig 7).55,56 Masala et al55 managed 23 patients with painful Schmorl nodes refractory to medial and conservative management by using PVA. Pain improvement was achieved in 18 patients as evaluated 4 hours after the procedure (mean Visual Analog Scale score, 8.4 versus 2.3). The authors reported no cement leakage into the Schmorl nodes or intervertebral disc space (Fig 8).55 In addition, a Rami communicans nerve block has been successfully used for the management of symptomatic Schmorl nodes.57

Other Benign Spine Lesions. Several other benign spinal lesions have been successfully treated by using minimally invasive percutaneous thermal ablation and PVA with no major complications (Fig 9).58-62 However, the available literature is scant and limited to case reports, including PVA for the management of eosinophilic granuloma,58,59 RFA for the management of intraosseous glomus tumor and hibernoma,60,61 and cryoablation for the management of epithelioid hemangioendothelioma.62

Complications
The most important potential complication of percutaneous thermal ablation of the spine is injury to the spinal cord and nerve roots due to the proximity of the ablation zone to neural elements. In addition, there remains a risk of thermal skin injury. Furthermore, in ablation of vertebral body or pedicle lesions or large tumors, there is a potential risk of ablation-related fracture, which may be minimized with cementation.

Posttreatment Imaging
As stated earlier, contrast-enhanced MR imaging is the technique of choice for posttreatment imaging of benign spine lesions, which is performed at the discretion of the referring physician, if symptoms recur. The necrotic ablated volume shows no enhancement on postcontrast imaging, with variable signal intensity on T1- and T2-weighted sequences, depending on the relative amounts of sclerotic bone, residual vascular and yellow marrow, and hemorrhagic products.63 The ablated volume is surrounded...
by a T1-hypointense, T2-hyperintense rim that enhances after contrast administration and corresponds histologically to granulation tissue or vascular fibrosis. Thermal insulation can be achieved by hydrodissection or instillation of warm or cool liquid, which actively modifies the temperature surrounding the structure at risk. Hydrodissection in RFA procedures should be performed by using nonionic solutions such as dextrose 5% in water. Saline solutions should be avoided ther treatment should be based on clinical evaluation.

Thermoprotection and Thermal Monitoring

Percutaneous thermal tumor ablation in the spine poses an inherent risk of injury to the spinal cord and nerve roots due to the proximity of the ablation zone to the susceptible neural elements, which is the most important potential complication of these procedures. Numerous parameters affect the extent and severity of neurologic thermal injury, including absolute temperature, duration of thermal effect, distance from margins of ablation zone, presence or absence of intact osseous cortex, and type of nerve fiber. A slight motor function loss becomes evident at 10°C, and mild sensory loss, at 7°C, while both functions disappear between 5°C and 0°C.

Current practice to prevent thermal injury during thermal spine ablation procedures involves the use of thermal insulation, and temperature and neurophysiologic monitoring. Thermal insulation can be achieved by hydrodissection or instillation of warm or cool liquid, which actively modifies the temperature surrounding the structure at risk. Hydrodissection in RFA procedures should be performed by using nonionic solutions such as dextrose 5% in water. Saline solutions should be avoided with RFA because the electrical conductivity may result in expansion of the ablation zone and creation of a plasma field. Carbon dioxide insufflation of the epidural space or neuroforamina can also be used to dissect and actively insulate the neural structures. In addition to insulation, continuous real-time and precise temperature monitoring may be undertaken during spine ablations by placing thermocouples close to the threatened structures, typically within the neuroforamina. In clinical practice, active thermoprotection (Figs 1–4) is initiated once the temperature reaches 45°C (heat) and 10°C (cold).

Investigators have implemented neurophysiologic monitoring and nerve electrostimulation during spine thermal ablation procedures by using estimations of motor- and somatosensory-evoked potential amplitudes. Substantial reduction in the amplitude of evoked potential amplitudes affords early detection of impending neurologic injury, which should prompt active thermoprotection or modification in the ablation procedure. Skin injury is another potential complication of percutaneous thermal ablation. Careful assessment of the boundaries of the ablation zone minimizes the risk of skin injury. Active skin thermoprotection such as surface application of warm saline during cryoablation should be implemented to minimize skin injury. In the case of RFA, using a bipolar system inherently obviates the risk of skin burn, and use of grounding pads with unipolar systems decreases the risk of skin injury.

Summary

Continuously evolving image-guided percutaneous vertebral augmentation and spine thermal ablation procedures have been proved safe and effective tools in minimally invasive management of selected patients with benign spine lesions (Table). With the progressively increasing role of these procedures in clinical practice, radiologists should be familiar with the indications, techniques, potential complications, and most recent advances of these procedures for optimal patient care. Special attention to the details of procedural techniques, including the choice of treatment technology; thermoprotection; and the adequacy of treatment will translate into improved patient outcomes.

TABLE 3.

Summary of benign spine lesions and preferred treatment modalities

<table>
<thead>
<tr>
<th>Benign Spine Lesion</th>
<th>Recommended Treatment Modality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Osteoid osteoma</td>
<td>Radiofrequency ablation, laser ablation</td>
</tr>
<tr>
<td>Osteoblastoma</td>
<td>Cryoablation [if large or coexisting soft-tissue component]</td>
</tr>
<tr>
<td>Aneurysmal bone cyst</td>
<td>Cryoablation [due to involvement of posterior elements, soft-tissue component, and large size]</td>
</tr>
<tr>
<td>Hemangioma</td>
<td>Vertebroplasty augmentation, alcohol ablation</td>
</tr>
<tr>
<td>Paget disease</td>
<td>Vertebral augmentation</td>
</tr>
<tr>
<td>Schmorl node</td>
<td>Vertebral augmentation</td>
</tr>
</tbody>
</table>

Disclosures: Adam N. Wallace—UNRELATED: Support for Travel to Meetings for the Study or other Purposes: Dfine.* Comments: Dfine paid my travel to a meeting to discuss a prospective clinical trial funded by the company. Other: Dfine.* Galil Medical.* Medtronic.* Comments: These companies donated ablation probes for an animal research study I am conducting. Jack W. Jennings—UNRELATED: Grant: Dfine, Comments: Principal Investigator for the STARIT trial. Consulting Fee or Speaking and Teaching: DFINE, Other: Toshiba, Comments: Engaged in Expert Testimony. Other: Galil Medical, Comments: Consultant, Engaged in Expert Testimony. Other: Medtronic, Comments: Consultant, Engaged in Expert Testimony. Other: B. Braun, Comments: Consultant, Engaged in Expert Testimony.
REFERENCES


Brain Perivascular Spaces as Biomarkers of Vascular Risk: Results from the Northern Manhattan Study

J. Gutierrez, M.S.V. Elkind, C. Dong, M. Di Tullio, T. Rundek, R.L. Sacco, and C.B. Wright

ABSTRACT

BACKGROUND AND PURPOSE: Dilated perivascular spaces in the brain are associated with greater arterial pulsatility. We hypothesized that perivascular spaces identify individuals at higher risk for systemic and cerebral vascular events.

MATERIALS AND METHODS: Stroke-free participants in the population-based Northern Manhattan Study had brain MR imaging performed and were followed for myocardial infarction, any stroke, and death. Imaging analyses distinguished perivascular spaces from lesions presumably ischemic. Perivascular spaces were further subdivided into lesions with diameters of ≤3 mm (small perivascular spaces) and >3 mm (large perivascular spaces). We calculated relative rates of events with Poisson models and hazard ratios with Cox proportional models.

RESULTS: The Northern Manhattan Study participants who had MR imaging data available for review (n = 1228; 59% women, 65% Hispanic; mean age, 71 ± 9 years) were followed for an average of 9 ± 2 years. Participants in the highest tertile of the small perivascular space score had a higher relative rate of all deaths (relative rate, 1.38; 95% CI, 1.01–1.91), vascular death (relative rate, 1.87; 95% CI, 1.12–3.14), myocardial infarction (relative rate, 2.08; 95% CI, 1.01–4.31), any stroke (relative rate, 1.79; 95% CI, 1.03–3.11), and any vascular event (relative rate, 1.74; 95% CI, 1.18–2.56). After we adjusted for confounders, there was a higher risk of vascular death (hazard ratio, 1.06; 95% CI, 1.01–1.11), myocardial infarction (hazard ratio, 2.22; 95% CI, 1.12–4.42), and any vascular event (hazard ratio, 1.04; 95% CI, 1.01–1.08) with higher small perivascular space scores.

CONCLUSIONS: In this multiethnic, population-based study, participants with a high burden of small perivascular spaces had increased risk of vascular events. By gaining pathophysiologic insight into the mechanism of perivascular space dilation, we may be able to propose novel therapies to better prevent vascular disorders in the population.

ABBREVIATIONS: HR = hazard ratio; LPI = lesions presumably ischemic; MI = myocardial infarction; NOMAS = Northern Manhattan Study; PP = pulse pressure; PVS = perivascular space; RR = relative rate; WMHV = white matter hyperintensity volume

On entering the skull, large arteries of the circle of Willis become encircled by a layer of leptomeninges and pial cells that accompany them through their course in the subarachnoid space and into the brain parenchyma, creating a potential perivascular space (PVS). These perivascular spaces are heterogeneous in their anatomic and physiologic characteristics. For example, basal ganglia PVSs are anatomically distinct from cortical PVSs, and arterial PVSs differ from venous PVSs. These anatomic differences may have implications for interstitial fluid drainage rates in different brain regions.

An increased prevalence of PVSs has been associated with multiple sclerosis, which may suggest altered clearance of inflammatory cells and exudates from the PVSs. PVSs have also been associated with pulsatile blood hemodynamics and hypertension. It may be that through these associations, PVSs relate to carotid atherosclerosis, imaging biomarkers of brain small-artery disease such as white matter hyperintensities, and lacunar stroke, as well as brain atrophy. The coexistence with other imaging biomarkers of small-artery disease, such as lacunar infarcts and white matter disease, which themselves are associated with hypertension, may suggest a shared physiopathology.

Received October 11, 2016; accepted after revision January 6, 2017.
From the Departments of Neurology (J.G., M.S.V.E.) and Medicine (M.D.), Columbia University Medical Center, New York, New York; Department of Epidemiology (M.S.V.E.), Mailman School of Public Health, Columbia University, New York, New York; and Department of Neurology (C.D., T.R., R.L.S., C.B.W.), Miller School of Medicine, University of Miami, Miami, Florida.
This work was supported by National Institutes of Health 2R01NS029993–23A1. Please address correspondence to Jose Gutierrez, MD, MPH, 710 W 168th St, 6th floor, Suite 639, New York, NY 10032; e-mail: jg3233@cumc.columbia.edu; @drjosegc
http://dx.doi.org/10.3174/ajnr.A5129

Indicates open access to non-subscribers at www.ajnr.org

Indicates article with supplemental on-line tables.
Nonetheless, dilated PVSs have been associated with higher systolic blood pressure and pulse pressure (PP) to a greater extent than other imaging biomarkers of small-artery disease\textsuperscript{6,15}; this association suggests a differential role for greater pulsatility as a predisposing factor for these lesions. The validity of dilated PVSs as imaging biomarkers of risk has not been validated prospectively, however, to our knowledge.

Because of the links between PVSs and vascular disease, we tested the hypothesis that brain PVSs are independent imaging biomarkers of vascular risk and that this risk is greater in individuals with high PP and among those with evidence of established non-PVS-related small-artery disease in the Northern Manhattan Study (NOMAS) MR imaging substudy.

**MATERIALS AND METHODS**

The NOMAS MR imaging substudy represents a population-based sample of adults 50 years of age or older at the time of MR imaging who have prospective follow-up. A detailed report of the criteria for enrollment in the NOMAS cohort and the MR imaging substudy has been provided elsewhere.\textsuperscript{6} In brief, participants needed to be free of stroke symptoms at the time of enrollment and had to provide informed consent to undergo brain MR imaging. The study was approved by the local institutional review board (Columbia University and University of Miami). Demographic data were self-reported. Hypertension, diabetes, and hypercholesterolemia were defined at the time of MR imaging and during follow-up by either self-reported diagnoses, medication use to treat these vascular risks, or blood pressure and/or laboratory evidence of these vascular risks, as previously reported.\textsuperscript{6,9} Blood pressure measurements were taken twice (\textgreek{gamma} hour apart) at the time of the brain MR imaging. Smoking was defined as self-reported current smoking at the time of MR imaging and during follow-up.

**Brain MR Imaging Protocol and Postprocessing**

Imaging was performed on a dedicated 1.5T research MR imaging system (Philips Healthcare, Best, the Netherlands). The FLAIR image was acquired in the multisection turbo spin-echo mode with an FOV of 250 mm, rectangular FOV of 80\%, acquisition matrix of 192 $\times$ 133 scaled to $256 \times 256$ in reconstruction, 3-mm section thickness with no gap, a TE of 144 ms, a TR of 5500 ms, an inversion recovery delay of 1900 ms, and a flip angle of 90\°. Dilated PVSs have been subdivided into 4 types: type III lacunes, a to d. Dilated PVSs of small diameter (ie, <3 mm) have been called “type III lacunes a” or “cribrures,”\textsuperscript{5} and here we call them small PVSs. They are often numerous, and it is not practical to count them. Consequently, semiquantitative scores have been proposed.\textsuperscript{14} Small PVSs are defined here as brain parenchymal voids observed on axial T1 images of no more than 5 mm in any axial diameter (Fig 1).\textsuperscript{16,17} A semiquantitative score ranging from 0 to 2 was used to rate 13 separate anatomic areas in the brain, with a possible overall score thus ranging from 0 to 26. The intra- and interreader reliabilities were excellent (intraclass correlation coefficient = 0.90) and good (intraclass correlation coefficient = 0.73).\textsuperscript{5,9} For T1 voids of $\geq$5 mm, we noted their anatomic location and the associated FLAIR characteristics (ie, hyperintense rim) and obtained the longest axial diameter, a diameter perpendicular to that, and a vertical diameter (multiplying the number of sections by section thickness [1.3 mm in this case]) to account for the rostral path of perivascular spaces. The volume of each large T1 void was calculated by using the ABC/2 formula for ovoid bodies,\textsuperscript{18} which yielded a minimum volume-derived diameter of 3 mm. Motion artifacts on T1 were noted as none, mild-to-moderate, or severe. The measurement and visual ratings were performed with 3D Slicer open source software, Version 4.3.1 (http://www.slicer.org), which permits measurements below the pixel resolution of our scans.

We developed a previously published probabilistic method informed by anatomic-radiologic correlation studies that differentiate subclinical lesions presumably ischemic (LPI) and large PVSs (Fig 1).\textsuperscript{6} Large PVSs correspond to lacunes type IIb (subcortical or cerebellar) and type IIIc (infrapataminal).\textsuperscript{3} With the shape, anatomic location, and features of the T1 voids, we classified lesions as more likely LPI, large PVS, or uncertain. With these criteria, the FLAIR rim was the dominant feature in determining LPI versus large PVSs, except in the brain stem and the upper two-thirds of the basal ganglia, where pathologic evidence showed that voids in these areas are more likely to represent infarcts.\textsuperscript{17,19,20} Uncertain voids were found in 56 participants and were isolated from LPI or large PVSs in only 9 participants. Due to the lack of confirmatory testing about the nature of these voids, we left them out of this analysis. This method has proved to have moderate-to-good reliability.\textsuperscript{6,9} The total volume of large perivascular spaces was calculated as the sum of the individual volume of each large perivascular space noted per participant, and the large perivascular spaces score consisted of the number of large perivascular spaces per participant.

White matter hyperintensity volume (WMHV), total intracranial volume (ie, head size), and total cerebral volume were obtained automatically with a Quantum 6.2 package on an Ultra 5 workstation (Sun Microsystems, Santa Clara, California) as described before. Briefly, tracing the dura allowed the removal of brain elements from the scan, keeping anterior and middle cranial fossae but excluding the posterior fossa, resulting in total intracranial volume. Total cerebral volume was computed as the sum of whole-brain volume voxels from the T1 segmentation process, and white matter hyperintensity volume was calculated as the sum
of voxels of ≥3.5 SDs above the mean image intensity multiplied by pixel dimensions and section thickness using axial FLAIR images.\textsuperscript{21} For this study, the percentage of atrophy was obtained by using the formula \((\text{Total Intracranial Volume} - \text{Total Cerebral Volume}) / \text{Total Intracranial Volume} \times 100\).\textsuperscript{22}

**Longitudinal Follow-Up**

Participants in the NOMAS MR imaging substudy are screened annually with standardized telephone interviews and/or in-person visits if the participant was screened positive for a predefined outcome and/or censoring (ie, no outcome reported). The outcomes include any stroke, MI, or death (including cause of death). Any vascular event was defined as a composite of vascular death, any stroke, or myocardial infarction (MI), as described previously.\textsuperscript{23} Briefly, death and vascular death were adjudicated by NOMAS investigators with a negligible loss to follow-up. Most NOMAS participants are admitted to Columbia University Medical Center, we collect their medical records from the reported hospital of admission. Subsequently, stroke subtypes are adjudicated independently by 2 study vascular neurologists (blinded to the baseline MR imaging), and MI is adjudicated by a study cardiologist with the criteria from the Cardiac Arrhythmia Suppression Trial and the Lipid Research Clinics Coronary Primary Prevention Trial.\textsuperscript{24,25}

**Statistical Analysis**

Differences in continuous variables were tested with Student t tests, and differences in categoric variables were assessed with \(\chi^2\) tests. The 2 main independent variables were the small and large PVSs. Outcomes were all deaths, vascular death, MI, any stroke, and any vascular events. The small and large PVS scores were used continuously with a Poisson distribution or categorically with arbitrary cutoffs of tertiles for the small PVS score, and ≥2 large PVSs for the large PVS score. To build risk models, we adjusted for total intracranial volume as a covariate. We calculated incidence rates, relative rates, and their 95% confidence intervals with Poisson regression and hazard ratios (HRs) and their 95% CIs with Cox proportional hazard regressions, with robust sandwich error variance, adjusting for demographics, prevalent and incident vascular risk factors, and competing risks between death and the other outcomes.\textsuperscript{26} To verify that the effect of small or large PVSs was not confounded by pulsatility, we repeated fully adjusted models, substituting hypertension with PP, mean arterial pressure, and a variable for the use of antihypertensives. We tested whether an interaction existed between small and large PVSs with PP and separately with WMHV or LPI at the time of MR imaging and performed stratified models if indicated. The statistical analysis was performed with SAS software, Version 9.4 (SAS Institute, Cary, North Carolina).

**RESULTS**

**Description of the Cohort**

After excluding those without available T1 sequences, 1228 NOMAS participants were included in this analysis (On-line Table 1). Participants were followed for an average of 9 ± 2 years (median, 9 years; range, 0.5–12.0 years), contributing a total of 10,905 person-years of follow-up with negligible loss to follow-up.

Small PVSs were found in 91% of the sample (per-participant mean, 5.5 ± 3.8; median, 5; range, 0–22), while large PVSs were present in 42% (20% had only 1 large PVS; 22% had ≥2 large PVSs). Small PVSs were more often found in the basal ganglia (83%), followed by subcortical white matter (79%) and the posterior fossa (17%). Large PVSs were more often infraputaminal (29%) than subinsular (13%) or subcortical (7%). With linear regression with total intracranial volume as the outcome and adjusted for age, sex, and ethnicity, a small PVS score (\(\beta = 0.0001, P = .02\)) and large perivascular spaces total volume (\(\beta = 0.05, P = .04\)) were associated with total intracranial volume, but a large PVS score was not (\(\beta = 0.60, P = .80\)).

In adjusted models, hypertension was associated with small (\(\beta = 0.14 ± 0.05, P = .001\)) and large PVS scores (\(\beta = 0.52 ± 0.11, P < .001\)). Men (\(\beta = -0.13 ± 0.05, P = .01\)) and older age (\(\beta = 0.05 ± 0.01\) per year, \(P = .001\)) were associated with a small PVS score, and a lower body mass index was associated with a large PVS score (\(\beta = -0.03 ± 0.01, P = .001\), On-line Table 2). There was no association between brain atrophy and small PVS (\(\beta = -0.01, P = .34\)) or large PVS scores (by number \(\beta = 0.01, P = .64\); or by volume \(\beta = 0.001, P = .21\)).

**Incidence and Risk of Outcome Events**

Compared with participants in the lowest tertile of small PVSs, those in the highest tertile had a higher relative rate (RR) of all deaths (RR, 1.38; 95% CI, 1.01–1.91), vascular death (RR, 1.87; 95% CI, 1.12–3.14), MI (RR, 2.08; 95% CI, 1.01–4.31), any stroke (RR, 1.79; 95% CI, 1.03–3.11), and any vascular event (RR, 1.74; 95% CI, 1.18–2.56). There were no significant differences in the rate of events by the number of large PVSs (On-line Table 3).

In adjusted Cox proportional models, participants with higher small PVS scores had a higher risk of vascular death, MI, or any vascular events but not of any stroke (Table 1). Adjusting for WMHV and LPI attenuated the association of small PVSs with the risk of vascular death (HR, 1.04; 95% CI, 0.98–1.09) and any vascular events (HR,1.02; 95% CI, 0.99–1.06) but not with the risk of MI (HR, 2.15; 95% CI, 1.07–4.33). Substituting PP, mean arterial pressure, and antihypertensive use at the time of MR imaging instead of hypertension did not change the significance of these associations.

Large PVSs were not predictors of outcome events in this sample, either by number, volume, or categorized as ≥2 large PVSs. Stratifying the small and large PVS scores by anatomic location demonstrated that participants with higher small PVS scores in the basal ganglia were at a higher risk of vascular death (HR, 1.12; 95% CI, 1.02–1.22) and any vascular event (HR, 1.06; 95% CI, 1.00–1.13), while a higher small PVS score in the subcortical regions conferred a higher risk of any stroke (HR, 1.08; 95% CI, 1.00–1.16) but not of MI or vascular death (Table 2).

**Interactions with Pulsatile Hemodynamics and Other Biomarkers of Small-Artery Disease**

Pulse pressure measured at MR imaging was an effect modifier of the association between small PVSs and vascular death (\(P = .011\),
any stroke \((P = .095)\), and any vascular event \((P = .019, \text{On-line Table 4})\). In a stratified analysis, a small PVS score was predictive of vascular death, any stroke, and any vascular event only among participants in the highest tertile of PP and/or systolic blood pressure at the time of their MR imaging (Table 3).

There were statistical interactions between small and large PVSs with either WMHV or LPI (On-line Table 5). In stratified models, the risk of all deaths, vascular death, any stroke, and any vascular event related to the small PVS score increased in a dose-effect manner with increasing tertiles of white matter hyperintensity volume (Tables 4 and 5). The risk of MI related to small and large PVSs was higher among participants with coexisting LPI than in those without LPI. Similarly, the risk of vascular death and any vascular events related to a high PVS score was higher among participants with coexisting LPI than in those without LPI.

**DISCUSSION**

In this population-based study of stroke-free participants, higher incidence rates of death, vascular death, MI, stroke, and any vascular event were noted among those with the highest burden of brain small PVSs but not large PVSs. The risk of events was enhanced when PVSs coexisted with other imaging biomarkers of cerebrovascular disease, similar to what has been reported in other cohorts.\(^{27,28}\) Also, small PVSs predicted risk among those in the highest tertile of PP and systolic blood pressure, which we interpret as epidemiologic evidence that small PVSs may be considered imaging biomarkers of high systemic pulsatility and may represent end-organ (brain) damage of arterial stiffness.

The discrepancy noted by anatomic location and by coexistence with other biomarkers of cerebrovascular disease suggests a different pathophysiology of proximal-versus-distal small PVSs.

---

**Table 1: Hazard ratios between perivascular spaces and outcome events\(^a\)**

<table>
<thead>
<tr>
<th>Outcome Events</th>
<th>All Deaths</th>
<th>Vascular Death</th>
<th>Myocardial Infarction</th>
<th>Any Stroke</th>
<th>Any Vascular Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small PVS score</td>
<td>1.01 (0.98–1.04)</td>
<td>1.06 (1.01–1.11)</td>
<td>1.03 (0.98–1.08)</td>
<td>1.04 (0.98–1.10)</td>
<td>1.04 (1.01–1.07)</td>
</tr>
<tr>
<td>Categorically</td>
<td>First tertile</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Second tertile</td>
<td>0.72 (0.53–0.96)</td>
<td>1.13 (0.68–1.87)</td>
<td>2.21 (1.10–4.45)</td>
<td>0.96 (0.54–1.72)</td>
<td>1.23 (0.86–1.75)</td>
</tr>
<tr>
<td>Third tertile</td>
<td>0.95 (0.71–1.27)</td>
<td>1.34 (0.80–2.23)</td>
<td>1.81 (0.89–3.71)</td>
<td>1.51 (0.88–2.57)</td>
<td>1.38 (0.97–1.98)</td>
</tr>
<tr>
<td>Large PVS score</td>
<td>Continuous</td>
<td>0.99 (0.90–1.10)</td>
<td>0.99 (0.89–1.10)</td>
<td>1.10 (0.93–1.04)</td>
<td>1.07 (0.88–1.28)</td>
</tr>
<tr>
<td>Categorically</td>
<td>≤1</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>≥2</td>
<td>1.02 (0.77–1.36)</td>
<td>1.03 (0.77–1.67)</td>
<td>1.40 (0.83–2.36)</td>
<td>0.99 (0.60–1.61)</td>
<td>1.07 (0.78–1.48)</td>
</tr>
</tbody>
</table>

\(^a\) Models are adjusted for head size, motion artifacts, age, sex, ethnicity, vascular risk factors at MRI and during follow-up (hypertension, hypercholesterolemia, diabetes, and smoking), prior cardiac disease, and body mass index. Data are hazard ratios and 95% CIs.

**Table 2: Risk of vascular events by anatomic location\(^a\)**

<table>
<thead>
<tr>
<th>Anatomic Location</th>
<th>All Deaths</th>
<th>Vascular Death</th>
<th>Myocardial Infarction</th>
<th>Any Stroke</th>
<th>Any Vascular Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small PVS score</td>
<td>1.01 (0.95–1.08)</td>
<td>1.12 (1.02–1.22)</td>
<td>1.06 (0.94–1.19)</td>
<td>1.00 (0.90–1.12)</td>
<td>1.06 (1.00–1.13)</td>
</tr>
<tr>
<td>Subcortical</td>
<td>1.01 (0.95–1.07)</td>
<td>1.06 (0.99–1.14)</td>
<td>1.01 (0.94–1.12)</td>
<td>1.08 (1.00–1.16)</td>
<td>1.05 (0.99–1.14)</td>
</tr>
<tr>
<td>Posterior fossa</td>
<td>1.02 (0.80–1.36)</td>
<td>1.15 (0.79–1.76)</td>
<td>1.26 (0.83–1.91)</td>
<td>1.09 (0.70–1.67)</td>
<td>1.09 (0.83–1.44)</td>
</tr>
<tr>
<td>Large PVS score</td>
<td>0.68 (0.43–1.07)</td>
<td>0.81 (0.42–1.47)</td>
<td>1.26 (0.85–1.85)</td>
<td>1.32 (0.95–1.85)</td>
<td>1.08 (0.79–1.47)</td>
</tr>
<tr>
<td>Subcortical white matter</td>
<td>0.94 (0.80–1.10)</td>
<td>1.21 (0.92–1.63)</td>
<td>1.10 (0.84–1.43)</td>
<td>1.08 (0.80–1.45)</td>
<td>1.11 (0.93–1.30)</td>
</tr>
<tr>
<td>Subcortical supraptaminal</td>
<td>1.73 (1.07–2.77)</td>
<td>0.54 (0.22–2.25)</td>
<td>1.17 (0.31–4.37)</td>
<td>1.45 (0.61–3.43)</td>
<td>0.90 (0.40–1.91)</td>
</tr>
<tr>
<td>Basal ganglia infraputaminal</td>
<td>1.05 (0.88–1.25)</td>
<td>1.24 (0.93–1.66)</td>
<td>0.98 (0.63–1.46)</td>
<td>0.92 (0.70–1.22)</td>
<td>1.09 (0.89–1.33)</td>
</tr>
</tbody>
</table>

\(^a\) Models are adjusted for head size, motion artifacts, age, ethnicity, vascular risk factors at MRI and during follow-up (hypertension, hypercholesterolemia, diabetes, smoking), prior cardiac disease, and body mass index. Data are hazard ratios and 95% CIs.
For example, as the pulse-wave pressure dissipates centrifugally from the aorta, proximal arterial stiffness limits attenuation of the pulse-wave as it travels into branching arteries leading to end organs. As the pulse-wave enters the brain, it encounters first the lenticulostrate penetrating branches of the middle cerebral arteries supplying the substantia innominata and the basal ganglia and, posteriorly, the vertebrobasilar system that supplies the brain stem and the cerebellum. The abrupt change in caliber from large to penetrating arteries renders these smaller vessels susceptible to systemic hemodynamic stresses, and it may explain why small PVSs in these proximal locations better predict vascular outcomes other than stroke. For example, NOMAS participants with PVSs in the posterior fossa had a nominal 26% higher risk of MI per PVS noted, while the risk of MI with evidence of subcortical small perivascular spaces was only 1% higher. Some studies have shown that subcortical (distal) PVSs are more common in nonvascular dementias and multiple sclerosis, while basal ganglia (proximal) PVSs are common with vascular dementias and hypertension. Another possible explanation for the disparities in the rates of PVS by location may be related to the double meningeal wrapping of penetrating arteries in the basal ganglia compared with single-layered coating in other locations, which may affect the drainage efficacy of interstitial fluids. Growing interest has been paid to the interplay among glia (proximal) PVSs are common with vascular dementias and small perivascular spaces was only 1% higher. Some studies in the posterior fossa had a nominal 26% higher risk of MI per PVS noted, while the risk of MI with evidence of subcortical small perivascular spaces was only 1% higher. Some studies have shown that subcortical (distal) PVSs are more common in nonvascular dementias and multiple sclerosis, while basal ganglia (proximal) PVSs are common with vascular dementias and hypertension.

Table 5: Stratified risk models by models by LPI-SBI status

<table>
<thead>
<tr>
<th>Stratified by</th>
<th>No LPI</th>
<th>Any LPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small PVS score</td>
<td>Referent group</td>
<td>1.06 (1.01–1.11)</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>Referent group</td>
<td>1.34 (1.12–1.60)</td>
</tr>
<tr>
<td>Large PVS score</td>
<td>Referent group</td>
<td>1.27 (1.01–1.58)</td>
</tr>
<tr>
<td>Vascular death</td>
<td>Referent group</td>
<td>1.18 (1.04–1.35)</td>
</tr>
</tbody>
</table>

* Models are adjusted for head size, motion artifacts, age, sex, ethnicity, vascular risk factors at MRI and during follow-up (hypertension, hypercholesterolemia, diabetes, and smoking), prior cardiac disease, and body mass index. Data are hazard ratios and 95% CIs.

The results presented here prospectively validate small VPSs as imaging biomarkers of vascular risk in an unselected population. We have attempted to systematically differentiate PVSs from infarcts as other groups have, but we have incorporated pathologic information that we believe may improve the accuracy of our methods. We lack pathology of these brains, which limits the claim that our method is valid and accurate. The reliability of ratings of brain parenchymal lesions, particularly <3 mm, is less than ideal, even in large epidemiologic studies such as NOMAS, Atherosclerosis Risk in Communities, and others. The error in measurement may explain the lack of statistical significance in some of our models as it pertains to small PVSs. Because large PVSs are larger and easier to quantify, error is unlikely to be the explanation for the lack of association between large PVSs and outcomes. Automation of the rating of small and large PVSs and LPI should be a priority in this field, but it appears daunting, given the complexities of the brain anatomy and the conflicting data among studies. The use of PP and systolic blood pressure as biomarkers of arterial stiffness is not unprecedented. The lack of more accurate measurements of arterial stiffness, however, is a limitation of our methods.

**CONCLUSIONS**

We present evidence that individuals with a higher burden of small PVSs on brain MR imaging are at a higher risk of vascular events, especially if their PP or systolic blood pressure is elevated. Among individuals with subcortical but not deep small perivascular spaces who also have coexpression of white matter hyperintensity and LPI, the risk of stroke is higher. These results emphasize the need for an integrated approach to the study of brain arterial disease that takes into account systemic and intracranial anatomic variations among individuals. By gaining this pathophysiologic insight, we may be able to propose novel therapies for palliating the effects of cerebrovascular disorders in the general population.

REFERENCES
32. Hughes W. Hypothesis. Lancet 1965;286:19–21 CrossRef Medline
Site and Rate of Occlusive Disease in Cervicocerebral Arteries: A CT Angiography Study of 2209 Patients with Acute Ischemic Stroke

D.C. Rotzinger, P.J. Mosimann, R.A. Meuli, P. Maeder, and P. Michel

ABSTRACT

BACKGROUND AND PURPOSE: CTA can rapidly and accurately detect and localize occlusive disease in patients with ischemic stroke. We have used CTA to assess arterial stenosis and occlusion in an ischemic stroke population arriving at a tertiary stroke center within 24 hours of symptom onset in order to obtain a comprehensive picture of occlusive disease pattern, and to determine the proportion of eligible candidates for endovascular treatment.

MATERIALS AND METHODS: Data from consecutive patients with acute ischemic stroke admitted to a single center between 2003 and 2012, collected in the Acute Stroke Registry and Analysis of Lausanne database, were retrospectively analyzed. Patients with a diagnostic CTA within 24 hours of symptom onset were selected. Relevant extra- and intracranial pathology, defined as stenosis of ≥50% and occlusions, were registered and classified into 21 prespecified segments.

RESULTS: Of the 2209 included patients (42.1% women; median age, 72 years), 1075 (48.7%) had pathology in and 308 (13.9%) had pathology outside the ischemic territory. In the 50,807 arterial segments available for revision, 1851 (3.6%) abnormal segments were in the ischemic (symptomatic) territory and another 408 (0.8%) were outside it (asymptomatic). In the 1211 patients with ischemic stroke imaged within 6 hours of symptom onset, 40.7% had symptomatic large, proximal segment occlusions potentially amenable to endovascular therapy.

CONCLUSIONS: CTA in patients with acute ischemic stroke shows large individual variations of occlusion sites and degrees. Approximately half of such patients have no visible occlusive disease, and 40% imaged within 6 hours show large, proximal segment occlusions amenable to endovascular therapy. These findings show the importance of early noninvasive imaging of extra- and intracranial arteries for identifying occlusive disease, planning recanalization strategies, and designing interventional trials.

ABBREVIATIONS: AIS = acute ischemic stroke; ASTRAL = Acute Stroke Registry and Analysis of Lausanne

More than 80% of strokes are ischemic, usually caused by large-artery atherosclerosis, cardiac embolism, or cerebral microangiopathy.1 CTA or MR arterial neuroimaging is frequently used at admission2,3 to determine occlusion sites and clot extent and to plan acute recanalization strategies.4-7 CTA is widely available and allows rapid assessment of the entire arterial vasculature from the aortic arch to the vertex. In addition, it accurately depicts arterial occlusive disease with good interrater agreement.8 Most clinically relevant arterial occlusive disease is found in the extra- and intracranial arteries, supporting aortic arch-to-vertex CTA for patients with acute ischemic stroke (AIS).9 Potential drawbacks of CTA are iodinated contrast allergy and radiation exposure.

The purpose of this study was to obtain a comprehensive picture of cerebrovascular occlusive disease in a representative AIS population, using admission CTA, to determine the proportion of eligible candidates for endovascular treatment. Such patients were identified by looking for symptomatic proximal arterial occlusions readily accessible with endovascular devices. Moreover, we aimed at determining the proportion...
of patients with abnormal segments in ischemic and nonischemic territories and describing the distribution of arterial occlusive disease in extra- versus intracranial vessels, anterior-versus-posterior circulation, and serial (tandem) pathologies.

MATERIALS AND METHODS

Patients

We used the patients admitted with AIS between January 2003 and December 2012 from the prospectively constructed Acute Stroke Registry and Analysis of Lausanne (ASTRAL), which collects information on all patients with AIS referred to the stroke center and/or intensive care unit of Lausanne University hospital (Centre Hospitalier Universitaire Vaudois) within 24 hours of the last-known-well time. We included all consecutive patients who had a diagnostic CTA available for analysis. CTA studies were rated as diagnostic or nondiagnostic by at least 2 authors, and the latter studies were excluded when at least 1 author considered them nondiagnostic (On-line Fig 1). A diagnostic CTA provides a sharp delineation of head and neck vessels, allowing evaluation of partial or complete filling defects. Patients arriving later than 24 hours from symptom onset and presenting with transient ischemic attack, cerebral or subarachnoid hemorrhage, persistent retinal ischemia, amaurosis fugax, and spinal cord ischemia were excluded. Reasons for exclusion were collected.

CTA-based imaging information, demographic data, cardiovascular risk factors, and ischemic side and territory were extracted and analyzed retrospectively. Stroke was categorized according to the TOAST (Trial of ORG 10172 in Acute Stroke) classification, and we added dissection, likely patent foramen ovale–related stroke, and combinations of mechanisms.

Neuroimaging Protocol

All included patients underwent standard AIS imaging, including noncontrast head CT, cerebral perfusion CT, and cervicocranial CT angiography. In patients undergoing thrombolysis, CTA was acquired immediately before or immediately after the rtPA bolus; therefore, this process made it unlikely that thrombolysis influenced the arterial occlusive pathology. Imaging was performed on a 16–detector row CT scanner (LightSpeed 16 Advantage; GE Healthcare, Milwaukee, Wisconsin) until November 2005 and afterwards on a 64–detector row CT scanner (LightSpeed VCT 64; GE Healthcare). Technical parameters for cerebral and extracranial CTA were the following: multidetector-array in a helical mode, 120 kV (peak), 150–300 mAs, section thickness = 1.25 mm before 2006 and 0.63 mm thereafter, pitch = 0.9:1. Image-acquisition delay depended on the arrival time of a 20-mL contrast bolus test (range, 15–20 seconds). Acquisition was then performed after intravenous administration of 50 mL of iohexol (300 mg/mL of iode, Accupaque; GE Healthcare) at a rate of 5 mL per second with a power injector (Stellant D CT Injection System; Medrad, Indianola, Pennsylvania) in the arterial phase. Coverage extended from the origin of the aortic arch to the top of the corpus callosum. Delayed phase images were obtained on the brain only.

Image Interpretation

Only CTA images were used for this study. Raw CTA source images in the axial plane were analyzed for focal arterial occlusive disease. Maximum intensity projections were then reconstructed in axial, sagittal, and coronal planes. Curvilinear reconstructions were obtained routinely.

A senior board-certified vascular neurologist (P. Michel) or neuroradiologists (R.A.M. or P. Maeder) with >10 years’ experience analyzed the acute CTAs within 7 days after admission. In cases of discordance, images were reviewed by the vascular neurologist and at least 1 neuroradiologist to reach agreement. Interrater agreement was assessed on 100 consecutive patients in the data base by using κ statistics and comparing the CTA interpretations by 1 vascular neurologist (P. Michel) and 1 neuroradiologist (P.J.M.). Both independently assessed the presence or absence of extracranial and intracranial segmental occlusive disease in the proximal (carotid siphon, M1, basilar artery) and distal circulation in the ischemic territory.

Twenty-one arterial segments (counting left and right sides, listed below) were analyzed separately for each patient and graded as normal, stenotic, or occluded. Segments were the following: extracranial vertebral artery (V1–3 segments); intracranial vertebral artery (V4 segment); basilar artery; posterior cerebral artery, P1, P2, and P3 segments; extracranial internal carotid artery; intracranial carotid artery, noting whether T–occlusion was present or not; middle cerebral artery, M1, M2, M3 segments; and anterior cerebral artery, A1, A2, and A3 segments. Imaging data are continuously recorded in the ASTRAL data base.

The term “abnormal” was used for stenotic or occluded segments.

“Occlusion” was defined as the absence of contrast medium filling the examined arterial segment on initial acquisition.11 For extracranial arteries, stenosis was defined by using the NASCET criteria: caliber reduction of ≥50% for vertebral arteries and ≥70% for carotid arteries.12 For intracranial arteries, stenosis was defined by using the Warfarin-Aspirin Symptomatic Intracranial Disease method (ie, ≥50% caliber reduction).13 “Tandem patterns” are defined as arterial occlusive disease affecting both the extra- and intracranial circulation in the same vascular axis. Abnormal segments on CTA were further categorized as symptomatic if the stenosis or occlusion was ipsilateral and proximal to the acute ischemic territory, or asymptomatic, if they were in the nonacute ischemic territory.

To evaluate the number of possible acute mechanical revascularization procedures in readily accessible large arteries, we subjected all occlusions in the extracranial and intracranial carotid arteries, V1–3, V4, basilar artery, P1, A1, M1, and M2 segments to a specific group analysis.

Data Processing and Statistical Analysis

Univariate analysis was performed to compare the characteristics of included and excluded patients by using the Wilcoxon 2-sample test for continuous variables and the χ² test for categoric variables. P values < .05 were considered significant. All data were processed by using STATA statistical software (Version 13.1, October 30 2013; StataCorp, College Station, Texas). Odds ratios were obtained by using an on-line calculator.14

ASTRAL was approved for scientific use by the institutional ethical commission, which did not require that individual informed consent be obtained.
RESULTS

Of 5022 patients with AIS symptoms who arrived at our institution during the observation period, 2209 patients were enrolled. Reasons for exclusion from ASTRAL were vascular diagnoses other than AIS (n = 2370). These and radiologic reasons for exclusion from the study (n = 443) are detailed in Online Fig 1.

Baseline characteristics of the included (n = 2209) and excluded (n = 443) patients with AIS are described in Table 1. In univariate comparison, patients with AIS excluded for radiologic reasons were older and more often women and tended to have more hypertension and atrial fibrillation and lower admission NIHSS scores. Stroke onset-to-CT and door-to-CT delays were longer, as expected in this group, because a main reason for study exclusion was a CTA performed after 24 hours of onset.

Cervicocerebral CTA analysis of the 2209 enrolled patients yielded 50,807 analyzable arterial segments.

Interrater agreement in the ischemic territory was almost perfect for intracranial proximal occlusive disease (κ = 0.87) and substantial for intracranial distal and extracranial occlusive disease (κ = 0.61 and 0.64, respectively), similar to previously published data.8

One thousand two hundred twenty-six patients (55.5%) had any arterial occlusion or stenosis (ischemic and nonischemic territories combined), while 983 patients (44.5%) did not show relevant occlusive disease (for details see Tables 2 and 3). One thousand seventy-five patients (48.7% of all patients) had such major abnormalities in the ischemic (ie, symptomatic) territory. Three hundred eight patients (13.9% of all patients) had stenosis or occlusion in territories not related to ischemia (ie, asymptomatic abnormalities) (see Online Tables 1 and 2 for full details).

If one considered arterial segments, arterial occlusion or stenosis occurred in 2259 segments (4.5% of all segments) (for details, see Tables 2 and 3); 1851 segmental abnormalities (81.9% of all abnormal segments or 3.6% of all examined segments) were in the ischemic territory.

---

Table 1: Patient characteristics and univariate comparison with the patients excluded for radiologic reasons

<table>
<thead>
<tr>
<th></th>
<th>Study Population (N = 2209)</th>
<th>Excluded Patients (n = 443)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>% or IQR</td>
<td>No.</td>
</tr>
<tr>
<td>Female sex (%)</td>
<td>929</td>
<td>42.1</td>
<td>216</td>
</tr>
<tr>
<td>Age (yr) IQR</td>
<td>72</td>
<td>21</td>
<td>77</td>
</tr>
<tr>
<td>Risk factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypertension (%)</td>
<td>1268</td>
<td>66.9</td>
<td>318</td>
</tr>
<tr>
<td>Heart valves (%)</td>
<td>61</td>
<td>3.2</td>
<td>11</td>
</tr>
<tr>
<td>Coronary artery disease (%)</td>
<td>295</td>
<td>15.5</td>
<td>80</td>
</tr>
<tr>
<td>Dyslipidemia (%)</td>
<td>1325</td>
<td>69.5</td>
<td>270</td>
</tr>
<tr>
<td>Diabetes (%)</td>
<td>331</td>
<td>17.4</td>
<td>82</td>
</tr>
<tr>
<td>Atrial fibrillation (%)</td>
<td>480</td>
<td>25.2</td>
<td>141</td>
</tr>
<tr>
<td>Low ejection fraction (%)</td>
<td>91</td>
<td>4.7</td>
<td>24</td>
</tr>
<tr>
<td>Smoking (%)</td>
<td>807</td>
<td>42.3</td>
<td>160</td>
</tr>
<tr>
<td>Stroke mechanism</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atherosclerosis (%)</td>
<td>304</td>
<td>13.8</td>
<td>99</td>
</tr>
<tr>
<td>Undetermined (%)</td>
<td>553</td>
<td>25</td>
<td>92</td>
</tr>
<tr>
<td>Cardiac (%)</td>
<td>641</td>
<td>29</td>
<td>145</td>
</tr>
<tr>
<td>Lacunar/microangiopathic (%)</td>
<td>269</td>
<td>12.2</td>
<td>62</td>
</tr>
<tr>
<td>Dissections (%)</td>
<td>107</td>
<td>4.8</td>
<td>14</td>
</tr>
<tr>
<td>PFO/other determined/rare (%)</td>
<td>159</td>
<td>7.2</td>
<td>40</td>
</tr>
<tr>
<td>Multiple/coexisting (%)</td>
<td>118</td>
<td>5.3</td>
<td>21</td>
</tr>
<tr>
<td>NIHSS admission (IQR)</td>
<td>7</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>Onset-to-CT delay (IQR) (min)</td>
<td>275</td>
<td>556</td>
<td>420</td>
</tr>
<tr>
<td>Door-to-CT delay (IQR) (min)</td>
<td>183</td>
<td>469</td>
<td>380</td>
</tr>
</tbody>
</table>

Note: IQR indicates interquartile range; PFO, patent foramen ovale.

a Values are expressed as medians and IQR for continuous variables or absolute counts and percentage for categoric variables unless otherwise stated.

---

Table 2: Distribution of arterial occlusive disease in numbers and rates, given per patient (N = 2209)

<table>
<thead>
<tr>
<th>Circulation</th>
<th>Ischemic Territory</th>
<th>Subtotal for Ischemic Territory</th>
<th>All Abnormalities (Related to Ischemia or Not)</th>
<th>No Relevant Abnormalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extracranial abnormalities per segment</td>
<td>Occlusion</td>
<td>Stenosis</td>
<td>Occlusion and/or Stenosis</td>
<td>Occlusion and/or Stenosis</td>
</tr>
<tr>
<td>Extracranial ICA or CCA</td>
<td>206</td>
<td>9.3%</td>
<td>165</td>
<td>7.5%</td>
</tr>
<tr>
<td>Extracranial vertebral</td>
<td>51</td>
<td>2.3%</td>
<td>26</td>
<td>1.2%</td>
</tr>
<tr>
<td>Subtotal extracranial abnormalities</td>
<td>257</td>
<td>11.6%</td>
<td>190</td>
<td>8.6%</td>
</tr>
<tr>
<td>Isolated extracranial abnormalities</td>
<td>35</td>
<td>1.6%</td>
<td>103</td>
<td>4.7%</td>
</tr>
<tr>
<td>Intracranial abnormalities per patient</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anterior intracranial</td>
<td>429</td>
<td>19.4%</td>
<td>88</td>
<td>4.0%</td>
</tr>
<tr>
<td>Posterior intracranial</td>
<td>96</td>
<td>4.4%</td>
<td>51</td>
<td>2.3%</td>
</tr>
<tr>
<td>Subtotal intracranial abnormalities</td>
<td>523</td>
<td>23.7%</td>
<td>135</td>
<td>6.1%</td>
</tr>
<tr>
<td>Isolated intracranial abnormalities</td>
<td>519</td>
<td>23.5%</td>
<td>133</td>
<td>6.0%</td>
</tr>
<tr>
<td>Total extra- and intracranial circulation per patient</td>
<td>780</td>
<td>35.3%</td>
<td>325</td>
<td>14.7%</td>
</tr>
</tbody>
</table>

Note: CCA indicates common carotid artery.

a The totals and subtotals may be smaller than the sum of the columns because patients may have arterial occlusive disease simultaneously in different segments or on different levels.
The proximal MCA (24.7% of all symptomatic abnormalities) was most frequently affected by symptomatic occlusive disease, followed by the extracranial ICA (20.2%) and then the distal MCA (14.5%).

As seen in the tables, occlusion was more frequent than stenosis. Symptomatic tandem patterns (On-line Table 3) were predominantly observed in the anterior circulation (82.1%) in 302 patients (28.1% of all patients with symptomatic abnormalities, 13.7% of all patients in the study). A graphic representation of observed occlusive disease per segment (symptomatic and asymptomatic stenosis or occlusions) is available in Fig 1. A column plot representation of arterial occlusive disease according to localization and type (stenosis versus occlusion) as defined in the “Materials and Methods” section is available in On-line Fig 2. Of patients having a symptomatic intracranial abnormality, 40% had a coexisting extracranial abnormality (tandem pattern). Only 34 of the tandem lesions (10.1% of all tandem lesions) were asymptomatic.

In 743 (33%) of the patients with AIS imaged within 24 hours of symptom onset, we found at least 1 symptomatic occlusion in a large, proximal artery and therefore readily accessible to endovascular treatment. These numbers increased to 640 (38.1%) of the 1679 patients imaged within 12 hours, 531 (40.7%) of 1304 imaged within 6 hours, and 480 (42.0%) of 1143 patients imaged within 4.5 hours. Of the patients arriving within 6 hours, the following segments were found occluded on CTA: extracranial ICA (n = 150, 11.5% of patients); V1–V3 (n = 31, 2.4% of patients); intracranial ICA (n = 34, 2.6% of patients); M1 (n = 193, 14.8% of patients); M2 (n = 114, 8.7% of patients); A1 (n = 32, 2.5% of patients); P1 (n = 23, 1.8% of patients); basilar artery (n = 22, 1.7% of patients). This is a total of 610 segments, found in 531 patients.

### Discussion

In the largest series to date of patients with AIS undergoing CTA within 24 hours from symptom onset, we found relevant arterial abnormalities in 55.5% of patients, most of which were observed in the ischemic territory. Regarding individual cervical and cerebral arterial segments, only 4.5% were abnormal, again affecting predominantly the ischemic territory. Arterial segments with the highest blood supply (anterior > posterior circulation, larger > smaller arteries) had the most occlusive disease. Approximately half of the patients with AIS did not have any visible arterial abnormalities on CTA, increasing the chances of a better outcome and making endovascular treatment unnecessary. The fact that an overwhelming proportion of arterial seg-

---

**Table 3: Distribution of occlusive disease for each arterial segment**

<table>
<thead>
<tr>
<th>Circulation</th>
<th>Ischemic territory</th>
<th>Subtotal Ischemic</th>
<th>All Abnormalities (Related to Ischemia or Not)</th>
<th>No Relevant Abnormalities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Occlusion</td>
<td>Stenosis</td>
<td>Occlusion or Stenosis</td>
<td>Occlusion or Stenosis</td>
</tr>
<tr>
<td>Extracranial abnormalities per segment in 2209 patients (n = 50,807 segments)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extracranial ICA or CCA</td>
<td>207</td>
<td>4.7%</td>
<td>167</td>
<td>3.8%</td>
</tr>
<tr>
<td>V1–3</td>
<td>59</td>
<td>1.3%</td>
<td>29</td>
<td>0.7%</td>
</tr>
<tr>
<td>Subtotal extracranial</td>
<td>266</td>
<td>3.0%</td>
<td>196</td>
<td>2.2%</td>
</tr>
<tr>
<td>Intracranial abnormalities per segment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carotid siphon without T</td>
<td>85</td>
<td>1.9%</td>
<td>47</td>
<td>1.1%</td>
</tr>
<tr>
<td>Carotid siphon with T</td>
<td>124</td>
<td>2.8%</td>
<td>65</td>
<td>1.5%</td>
</tr>
<tr>
<td>MCA: M1</td>
<td>405</td>
<td>9.2%</td>
<td>54</td>
<td>1.2%</td>
</tr>
<tr>
<td>MCA: M2–3 only</td>
<td>226</td>
<td>5.1%</td>
<td>42</td>
<td>1.0%</td>
</tr>
<tr>
<td>ACA: A1</td>
<td>104</td>
<td>2.4%</td>
<td>11</td>
<td>0.3%</td>
</tr>
<tr>
<td>ACA: A2–3 only</td>
<td>9</td>
<td>0.2%</td>
<td>3</td>
<td>0.1%</td>
</tr>
<tr>
<td>Subtotal anterior intracranial</td>
<td>953</td>
<td>3.6%</td>
<td>157</td>
<td>0.6%</td>
</tr>
<tr>
<td>PCA: P1</td>
<td>41</td>
<td>0.9%</td>
<td>10</td>
<td>0.2%</td>
</tr>
<tr>
<td>PCA: P2–3 only</td>
<td>33</td>
<td>0.8%</td>
<td>13</td>
<td>0.3%</td>
</tr>
<tr>
<td>V4</td>
<td>73</td>
<td>1.7%</td>
<td>30</td>
<td>0.7%</td>
</tr>
<tr>
<td>BA</td>
<td>55</td>
<td>2.5%</td>
<td>26</td>
<td>1.2%</td>
</tr>
<tr>
<td>Subtotal posterior intracranial</td>
<td>202</td>
<td>1.3%</td>
<td>79</td>
<td>0.5%</td>
</tr>
<tr>
<td>Subtotal intracranial</td>
<td>1155</td>
<td>2.8%</td>
<td>236</td>
<td>0.7%</td>
</tr>
<tr>
<td>Subtotal anterior circulation</td>
<td>1160</td>
<td>3.6%</td>
<td>324</td>
<td>1.0%</td>
</tr>
<tr>
<td>Subtotal posterior circulation</td>
<td>261</td>
<td>1.3%</td>
<td>108</td>
<td>0.5%</td>
</tr>
<tr>
<td>Total extra- and intracranial abnormalities per segment</td>
<td>1421</td>
<td>2.8%</td>
<td>432</td>
<td>0.9%</td>
</tr>
</tbody>
</table>

Note: ACA indicates anterior cerebral artery; PCA, posterior cerebral artery; BA, basilar artery; CCA, common carotid artery.

---

The ischemic territory (ie, symptomatic). Four hundred eight segmental abnormalities (18.1% of all abnormalities or 0.8% of all examined segments) were in the nonischemic territory (ie, asymptomatic) (see On-Line Tables 1 and 2 for full details).

Occlusive disease was most commonly seen in the anterior circulation (76.1% of all abnormal segments). This was more often intracranial than extracranial (30.9% versus 24.6% of all patients).

DISCUSSION

In the largest series to date of patients with AIS undergoing CTA within 24 hours from symptom onset, we found relevant arterial abnormalities in 55.5% of patients, most of which were observed in the ischemic territory. Regarding individual cervical and cerebral arterial segments, only 4.5% were abnormal, again affecting predominantly the ischemic territory. Arterial segments with the highest blood supply (anterior > posterior circulation, larger > smaller arteries) had the most occlusive disease. Approximately half of the patients with AIS did not have any visible arterial abnormalities on CTA, increasing the chances of a better outcome and making endovascular treatment unnecessary. The fact that an overwhelming proportion of arterial seg-

---

ments are patent in AIS (95.5%) is a reminder that there is a great potential for collateral blood supply into the ischemic territory, potentially allowing access of thrombolytic or neuroprotective drugs to the acutely ischemic brain.

Symptomatic arterial abnormalities were more often found in the anterior circulation (76.1% of all abnormal segments, occurring in 81.6% of patients with symptomatic abnormalities), which is consistent with the proportion of blood flow being directed to this part of the cerebral circulation. In addition, there is significantly more arterial occlusive disease in patients with anterior-versus-posterior circulation stroke (see odds ratios in Table 4). This may be related to posterior circulation strokes being more often of microangiopathic origin, with less detectable arterial pathology on noninvasive imaging.

Patients more often had symptomatic arterial abnormalities in the intra- rather than the extracranial circulation (28.6% versus 20.0% of all patients). This may be because most emboli from the heart and proximal extracranial arteries are too small to occlude the cervical arteries and get blocked only in smaller, intracranial vessels. The finding that the proportion of abnormal arterial segments is lower intracranially (3.8% versus 7.7%, Table 3) is because the intracranial circulation is divided into subsegments. Our results, therefore, show that it is crucial to examine both the extracranial arteries and the circle of Willis in patients with AIS.

Not surprising, symptomatic abnormalities are usually occlusive, whereas asymptomatic abnormalities are typically stenotic, likely explained by a reduced-but-sufficient blood flow through the stenosis. We found a low overall rate (0.7%) of symptomatic intracranial stenosis in our cohort, likely consisting of a combination of atherosclerotic plaques and partially resorbed emboli. A higher rate of intracranial stenosis has been described in Asian patients with ischemic stroke.

The 40% concurrent extracranial (tandem) pathology observed in patients with a symptomatic intracranial abnormality has practical implications because it may complicate endovascular access. This situation seems to be particularly frequent in the anterior circulation (odds ratio of 5.0 for anterior-versus-posterior tandem lesions in our population).

Rapid endovascular treatment by using predominantly stent retrievers has recently been shown to be more effective than IV thrombolysis alone in well-selected patients with proximal intracranial occlusions of the anterior circulation and is now considered the standard of care. This is because recanalization is one of the most important predictors of prognosis in AIS, and proximal occlusions insufficiently respond to IV thrombolysis in many cases. Endovascular stroke trials have mostly excluded patients with posterior circulation strokes. Although large posterior circulation occlusions are also likely to benefit from endovascular treatment, this benefit has to first be proved by a randomized controlled trial such as the Basilar Artery International Cooperation Study trial.

The proportion of 40.7% of such patients in this tertiary care setting imaged within 6 hours provides an estimate of the number of patients who could benefit from such interventions. When adding clinical and other radiologic criteria as well as contraindications, this number may be reduced by half in the real world, however. Still, improved prehospital identification and transport of patients with probable proximal intracranial occlusions are likely to increase the number of endovascularly treatable patients in comprehensive stroke centers.

The spontaneous recanalization rate of proximal occlusions

Table 4: Localization of arterial occlusive disease in the ischemic territory

<table>
<thead>
<tr>
<th></th>
<th>Per Patient (Table 2)</th>
<th>Per Segment (Table 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>95% CI</td>
</tr>
<tr>
<td>Anterior vs posterior circulation strokes</td>
<td>6.7</td>
<td>5.6–7.9</td>
</tr>
<tr>
<td>Intracranial segments</td>
<td>4.6</td>
<td>3.7–5.6</td>
</tr>
<tr>
<td>Extracranial segments</td>
<td>5.8</td>
<td>4.5–7.4</td>
</tr>
<tr>
<td>Intracranial vs extracranial localization</td>
<td>1.6</td>
<td>1.4–1.8</td>
</tr>
<tr>
<td>Anterior circulation</td>
<td>1.5</td>
<td>1.3–1.7</td>
</tr>
<tr>
<td>Posterior circulation</td>
<td>1.8</td>
<td>1.4–2.4</td>
</tr>
<tr>
<td>Tandem pattern: anterior vs posterior circulation</td>
<td>5.0</td>
<td>3.7–6.7</td>
</tr>
</tbody>
</table>

*a Results are expressed as odds ratios plus 95% confidence intervals.

FIG 1. Graphic representation of observed occlusive disease per segment. The area of the circle is proportional to the observed rate. A, Rate of symptomatic (right patient side, blue circles) and asymptomatic arterial occlusions (left patient side, green circles). B, Rate of symptomatic (right patient side, yellow circles) and asymptomatic arterial stenoses (left patient side, pink circles).
between 4.5 and 24 hours from symptom onset is approximately 9% based on the aforementioned proportions of patients with such occlusions. In contrast, early digital subtraction studies reported an estimated 17% spontaneous recanalization at 6–8 hours from stroke onset.28

Limitations are the retrospective, uncontrolled nature of the study with data from a single-center registry, which may not have a population representative of other settings for acute stroke care. We excluded a certain number of patients because acute CTA was not performed, mainly due to contrast contraindications. Images of head and neck vessels were analyzed in the arterial phase only, possibly overestimating the degree and extent of proximal occlusive disease. Also, the results are based on CTA, which is a reliable method for assessment of arterial pathology in AIS,3,29,30 but may not be generalizable to other angiographic methods. We did not analyze associations among arterial imaging, clinical variables, and functional outcome because this was the purpose of recent publications.29,31

On the other hand, the available data may represent the best possible approximation of a real-world AIS population, given that some patients will always be unable to undergo acute angiographic evaluations.

CONCLUSIONS
This study gives a comprehensive picture of the frequency and distribution of relevant arterial pathology on CTA obtained within 24 hours of symptom onset. Approximately half of patients had relevant arterial pathology in the ischemic territory, mostly in the intracranial anterior circulation; a significant number also had tandem pathology. The 40% of patients with AIS having proximal intracranial arterial occlusions in the first 6 hours give an estimate of the eligibility for acute endovascular therapy.


REFERENCES


ABSTRACT

BACKGROUND AND PURPOSE: The impact of deep cerebral veins on neurologic outcome after intravenous thrombolysis in patients with acute ischemic stroke is unclear. We investigated the relationship between the appearance of deep cerebral veins on susceptibility-weighted imaging and neurologic outcome in patients who underwent thrombolysis.

MATERIALS AND METHODS: We retrospectively analyzed 109 consecutive patients with acute ischemic stroke who had pretreatment SWI and received intravenous thrombolysis within 6 hours. We calculated the signal difference ratio (defined as the relative difference in signal intensity between the ipsilateral and contralateral veins) of the thalamostriate vein, septal vein, and internal cerebral vein on pretreatment SWI.

RESULTS: Only the signal difference ratio of the thalamostriate vein was significantly associated with poor outcome (3-month modified Rankin Scale score ≥ 2, P = .008). The optimal threshold was relative hypointensity of the ipsilateral vein of >4.8% (sensitivity of 53.7% and specificity of 80.9%). We defined a signal difference ratio of the thalamostriate vein of ≥5% as an ipsilateral prominent thalamostriate vein. Patients with an ipsilateral prominent thalamostriate vein were more likely to have poor outcome (OR = 3.66; 95% CI, 1.25–10.68; P = .02) and a lower rate of successful reperfusion (reperfusion rate of ≥70%; OR = 0.35; 95% CI, 0.13–0.92; P = .03), compared with those without an ipsilateral prominent thalamostriate vein. However, patients with an ipsilateral prominent thalamostriate vein were still less likely to experience poor outcome when reperfusion was achieved compared with when reperfusion did not occur (80.0% versus 44.4%, P = .04).

CONCLUSIONS: A pretreatment ipsilateral prominent thalamostriate vein was associated with reduced reperfusion after thrombolysis and poor outcome. More intensive reperfusion approaches may be required for patients with an ipsilateral prominent thalamostriate vein.

ABBREVIATIONS: AIS = acute ischemic stroke; ICV = internal cerebral vein; IPTSV = ipsilateral prominent TSV; IVT = intravenous thrombolysis; SDR = signal difference ratio; SV = septal vein; TSV = thalamostriate vein

Venous changes in the affected hemisphere after acute ischemic stroke (AIS) may play a crucial role in determining clinical outcome.1-3 Susceptibility-weighted imaging is useful for evaluating cerebral veins in AIS because it is extremely sensitive to paramagnetic substances such as deoxyhemoglobin, thus reflecting the oxygen extraction fraction and the cerebral metabolic rate of oxygen of the hypoxic tissue.1-4 Shortly after vascular occlusion, the uncoupling between oxygen supply and demand in the hypoperfused region leads to an elevated oxygen extraction fraction and subsequent increased level of deoxyhemoglobin in the vessel, which contributes to prominent hypointensity of the draining veins on SWI.5-7

Asymmetric cortical veins, which can be visualized on SWI after AIS and are considered an indicator of salvageable ischemic penumbra, were found to be associated with 3-month outcome.8 However, it was difficult to differentiate cortical veins from leptomeningeal collaterals because there were no obvious morphologic differences between them.8 Additionally, the anatomic variability of cortical veins also contributes to reduced interrater agreement when assessing the cortical veins. In contrast to...
cortical veins, leptomeningeal collaterals do not interfere with the assessment of deep cerebral veins, and anatomic variation in deep veins is less prevalent.

On the basis of SWI, it was demonstrated that prominent hypointense cortical veins and medullary veins predicted infarct growth and poor outcome after AIS. However, most of the included patients did not undergo revascularization therapy, and salvage of potentially penumbral regions was therefore unlikely. The imaging was also acquired up to 72 hours after AIS, which limits its applicability to hyperacute treatment decision-making. Furthermore, hypointense veins were only associated with prognosis in patients with large-vessel occlusion after AIS. Actually, slight perfusion abnormalities could even cause signal changes in the internal cerebral vein (ICV) and thalamostriate vein (TSV) on SWI.

Therefore, we developed a quantitative method to assess signal asymmetry in the major deep cerebral veins, including the ICA, TSV, and septal vein (SV), and hypothesized that this could sensitively reflect cerebral metabolic changes in hypoxic brain tissue after AIS and may be associated with clinical outcome after intravenous thrombolysis (IVT).

MATERIALS AND METHODS

Ethics Statement
The protocol of MR imaging–guided intravenous thrombolysis has been approved by our local human ethics committee. All clinical investigations were conducted according to the principles expressed in the Declaration of Helsinki. Informed consent was obtained for all patients.

Patient Selection
We retrospectively reviewed our consecutive patients with AIS who received intravenous alteplase (0.9 mg/kg, up to a maximum of 90 mg; 10% of the total dosage as a bolus and the rest infused for 1 hour) between June 2009 and July 2015. We then included patients under the following circumstances: 1) They received IVT within 6 hours of symptom onset; 2) underwent pretreatment multimodal MR imaging, including SWI, diffusion-weighted imaging, and perfusion-weighted imaging; 3) had hypoperfusion (time-to-maximum of >6 seconds) in the supratentorial brain; and 4) underwent follow-up DWI or NCCT at 24 hours after IVT. Patients were excluded for the following reasons: 1) Their prestroke modified Rankin Scale score was ≥2; 2) their MR imaging was degraded due to motion artifacts; or 3) there were bilateral perfusion lesions.

Imaging Acquisition
Multimodal MR imaging was performed on a 3T system (Signa Excite HD; GE Healthcare, Milwaukee, Wisconsin) equipped with an 8-channel phased array head coil. Foam pads were inserted into the space between the subject’s head and the MR imaging head coil to minimize head motion. The MR imaging protocol included axial isotropic diffusion-weighted imaging (TR = 4000 ms; TE = 69.3 ms; b-value = 1000 s/mm²; FOV = 240 mm; section thickness = 5 mm; section gap = 1 mm; acquisition matrix = 160 × 160; acquisition duration = 32 seconds); TOF-MRA (TR = 20 ms; TE = 3.2 ms; flip angle = 15°; acquisition matrix = 320 × 224; section thickness = 1.4 mm; 3 slabs, acquisition duration = 3 minutes 46 seconds); SWI (TE = 4.5 ms [first echo]; interecho spacing = 4.5 ms; TR = 58 ms; FOV = 24 × 24 cm²; matrix size = 256 × 256; flip angle = 20°; section thickness = 2.0 mm with no gap between sections; acquisition duration = 3 minutes 27 seconds); and perfusion-weighted imaging (FOV = 240 mm; 50 repetitions with TR = 1500 ms; TE = 30 ms; acquisition matrix = 128 × 128; section thickness = 5 mm; section gap = 1 mm, gadolinium dose = 15 mL; flow rate = 4–5 mL/s; acquisition duration = 1 minute 15 seconds). The entire duration of the MR imaging protocol was 15 minutes.

Imaging Analysis
Assessment of Deep Veins on Pretreatment SWI. To assess venous structures, we reconstructed SWI in a minimum intensity projection of 20-mm thickness with OsiriX Imaging Software (http://www.osirix-viewer.com) (Fig 1A). As indicated in Fig 1B–D, the mean signal-intensity value of ICV, TSV, and SV was measured by drawing a line (1 pixel in width) along the center of each vein. After obtaining the contralateral mean signal-intensity value ($S_{co-v}$) and ipsilateral mean signal-intensity value ($S_{ip-v}$) of each vein, we calculated the signal difference ratio (SDR) by using the following equation:

$$SDR = \left( \frac{S_{ip-v} - S_{co-v}}{S_{co-v}} \right) \times 100\%.$$  

Two stroke fellows (S.Z. and X.Z.) who were blinded to other imaging and clinical data independently assessed the SDR of SV,
Table 1: Univariate comparison between patients with good and poor outcomes

<table>
<thead>
<tr>
<th>Risk factors</th>
<th>Good Outcome (mRS ≤2)</th>
<th>Poor Outcome (mRS &gt; 2)</th>
<th>Test Value</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>66.1 ± 12.7</td>
<td>72.0 ± 12.4</td>
<td>t = -2.36</td>
<td>.02a</td>
</tr>
<tr>
<td>Female (No.) (%)</td>
<td>21 (30.9)</td>
<td>19 (46.3)</td>
<td>χ² = 2.63</td>
<td>.15</td>
</tr>
<tr>
<td>Hypertension (No.) (%)</td>
<td>46 (67.6)</td>
<td>29 (70.1)</td>
<td>χ² = 0.11</td>
<td>.83</td>
</tr>
<tr>
<td>Diabetes mellitus (No.) (%)</td>
<td>12 (17.6)</td>
<td>13 (31.7)</td>
<td>χ² = 2.86</td>
<td>.10</td>
</tr>
<tr>
<td>Atrial fibrillation (No.) (%)</td>
<td>21 (30.9)</td>
<td>21 (51.2)</td>
<td>χ² = 4.47</td>
<td>.04c</td>
</tr>
<tr>
<td>Smoking (No.) (%)</td>
<td>30 (44.1)</td>
<td>13 (31.7)</td>
<td>χ² = 1.65</td>
<td>.23</td>
</tr>
<tr>
<td>Previous stroke/TIA (No.) (%)</td>
<td>11 (16.2)</td>
<td>6 (14.6)</td>
<td>χ² = 0.05</td>
<td>1.00</td>
</tr>
<tr>
<td>Large-artery occlusion (No.) (%)</td>
<td>28 (41.2)</td>
<td>32 (78.0)</td>
<td>χ² = 14.05</td>
<td>&lt;.001f</td>
</tr>
<tr>
<td>Baseline SBP (mm Hg)</td>
<td>155.4 ± 20.5</td>
<td>151.0 ± 24.9</td>
<td>t = 0.64</td>
<td>.53</td>
</tr>
<tr>
<td>Baseline DBP (mm Hg)</td>
<td>87.1 ± 17.1</td>
<td>83.2 ± 14.8</td>
<td>t = 1.21</td>
<td>.23</td>
</tr>
<tr>
<td>ONT (min)</td>
<td>251.2 ± 74.8</td>
<td>235.1 ± 73.4</td>
<td>t = -1.11</td>
<td>.27</td>
</tr>
<tr>
<td>Baseline glucose level (mmol/L)</td>
<td>7.7 ± 2.6</td>
<td>8.1 ± 2.7</td>
<td>t = -0.77</td>
<td>.44</td>
</tr>
<tr>
<td>Baseline NIHSS score</td>
<td>6.0 (3.0–10.0)</td>
<td>14.0 (8.5–18.0)</td>
<td>Z = -5.11</td>
<td>&lt;.001f</td>
</tr>
<tr>
<td>Baseline diffusion lesion volume (mL)</td>
<td>1.0 (0–34.2)</td>
<td>8.4 (2.9–51.2)</td>
<td>Z = -3.12</td>
<td>.002f</td>
</tr>
<tr>
<td>Baseline hypoperfusion volume (mL)</td>
<td>22.7 (3.0–74.6)</td>
<td>81.8 (31.9–166.6)</td>
<td>Z = -4.16</td>
<td>&lt;.001f</td>
</tr>
<tr>
<td>SDR of TSV (%)</td>
<td>1.2 (–2.6–3.6)</td>
<td>5.8 (–11.8–9)</td>
<td>Z = -2.61</td>
<td>.009f</td>
</tr>
<tr>
<td>SDR of SV (%)</td>
<td>-0.3 (–2.4–2.6)</td>
<td>0.1 (–3.0–3.7)</td>
<td>Z = -0.45</td>
<td>.65</td>
</tr>
<tr>
<td>SDR of ICV (%)</td>
<td>-0.3 (–4.1–3.5)</td>
<td>0.9 (–5.3–5.7)</td>
<td>Z = -0.71</td>
<td>.48</td>
</tr>
<tr>
<td>Proportion of penumbral tissue loss (%)b</td>
<td>1.3 (0–19.9)</td>
<td>38.4 (4.9–95.3)</td>
<td>Z = -4.23</td>
<td>&lt;.001f</td>
</tr>
<tr>
<td>Infarct volume at 24 hr (mL)</td>
<td>4.2 (1.0–9.9)</td>
<td>38.8 (11.5–107.0)</td>
<td>Z = -5.54</td>
<td>&lt;.001f</td>
</tr>
</tbody>
</table>

Note:—SBP indicates systolic blood pressure; DBP, diastolic blood pressure; ONT, onset-to-needle time.

a Table cells express results in mean ± SD for normally distributed continuous variables, median (interquartile range) for ordinal variables and non-normally distributed continuous variables, respectively.


c P < .05.

d P < .01.

e P < .001.

TSV, and ICV. Cases with differences of >3% for SDR were reassessed by a neuroimaging physician (W.D.). The average value of the 2 SDRs was used.

**Radiologic and Clinical Assessment**

Diffusion lesion volume was calculated by applying a signal-intensity threshold of the apparent diffusion coefficient of <600 to pretreatment and 24-hour follow-up DWI.11 For patients with CT follow-up, the area of hypotenuation was manually outlined. Volumetric analysis was performed by using the commercial software MiStar (Apollo Medical Imaging Technology, Melbourne, Australia). Hypoperfusion was defined as time-to-maximum of >6 seconds. Reperfusion was assessed in patients with a baseline hypoperfusion volume of >10 mL. Successful reperfusion was defined as ≥70% reduction in hypoperfusion volume between pretreatment and 24-hour perfusion imaging.12 The proportion of penumbral tissue loss was defined as described previously: 100 × (Posttreatment Infarct Volume − Pretreatment Infarct Volume)/(Pretreatment Hypoperfusion Volume − Pretreatment Infarct Volume).13

Large-artery occlusion was defined as occlusion of the internal carotid, proximal middle cerebral (M1), or proximal posterior cerebral (P1) artery. The modified Rankin Scale score was assessed at 3 months and dichotomized into good (0–2) and poor (3–6) outcomes.

**Statistical Analysis**

All metric and normally distributed variables were reported as mean ± SD; non-normally distributed variables, as median (25th–75th percentile). Categoric variables were presented as frequency (percentage). The Fisher exact test was used to compare dichotomous variables between groups; the t test, for normally distributed continuous variables; the Mann-Whitney U test, for non-normally distributed continuous variables; and the Pearson χ² test, for categoric data. The Benjamini-Hochberg correction was used to control the false discovery rate of multiple comparisons. The interrater reliability was assessed by use of the intraclass correlation coefficient. The Spearman correlation coefficient was used to analyze the association of SDR in each vein with outcomes. The receiver operating characteristic curve analysis was performed, and the optimal threshold for deep vein asymmetry was determined by using the Youden index. The Mann-Whitney U test was used to analyze the relationship between the asymmetry of deep veins and the proportion of penumbral loss. Univariate and multivariate logistic regression analyses were conducted to investigate whether asymmetry of the deep veins was independently associated with clinical outcome. Results are reported as odds ratios with 95% confidence intervals. A P value < .05 was considered statistically significant. All statistical analyses were conducted by using SPSS, Version 19.0 (IBM, Armonk, New York).

**RESULTS**

**Patient Characteristics**

A total of 109 patients was included in the final analysis. The mean age was 68 ± 13 years with 40 (36.7%) women. The mean pretreatment NIHSS score was 7 (interquartile range, 4–14), the median time from onset to MR imaging was 193 minutes (interquartile range, 144–245 minutes), and the median time from onset to treatment was 230 minutes (interquartile range, 180–283 minutes). At 3-month follow-up, 41 (37.6%) had poor outcomes.
Deep Cerebral Veins and Outcome

Interrater agreement was excellent for the SDR of the SV (intraclass correlation coefficient = 0.82), TSV (intraclass correlation coefficient = 0.88), and ICV (intraclass correlation coefficient = 0.90). The SDR of the TSV was correlated with poor outcome (rho = 0.25, P = .008), but there was no correlation for the SV (rho = 0.04, P = .73) or ICV (rho = 0.07, P = .47). Patients with poor outcome had a higher SDR of the TSV (median, 4.5%; versus 0.7%; Z = −2.61, P = .009) (Table 1). Receiver operating characteristic analysis revealed that the SDR of the TSV was associated with poor outcome (area under the curve = 0.65; 95% CI, 0.54–0.76; P = .009), with the optimal threshold of 4.8% for identifying poor outcome (sensitivity of 53.7% and specificity of 80.9%, Youden index = 0.35).

Ipsilateral Prominent Thalamostriate Vein and Outcome

On the basis of the cutoff of 4.8%, we then defined ipsilateral prominent TSV (IPTSV) as an SDR of the TSV of ≥5% and dichotomized patients into IPTSV and non-IPTSV groups. As seen in Table 2, univariate analysis showed that patients with IPTSV had a higher rate of large-artery occlusion (74.3% versus 45.9%, P = .007), higher baseline NIHSS scores (median, 13 versus 7; P = .004), and longer onset-to-needle time (266.9 ± 68.6 minutes versus 229.0 ± 73.8 minutes, P = .01), compared with those with non-IPTSV. Baseline hypoperfusion volume was independently associated with the presence of IPTSV (OR = 1.01; 95% CI, 1.00–1.02; P = .02) after adjusting for large-artery occlusion, NIHSS score, and onset-to-needle time.

Patients with IPTSV had significantly higher rates of poor outcome than those with non-IPTSV (62.9% versus 25.7%), which remained significant in multivariate regression, which included age, NIHSS score, atrial fibrillation, diabetes, large-artery occlusion status, and baseline hypoperfusion volume (OR = 3.66; 95% CI, 1.25–10.68; P = .02) (Table 3 and Fig 2). The presence of IPTSV was associated with a high proportion of penumbral tissue loss (32.6% versus 4.5%, P = .004).

Relationship between Ipsilateral Prominent Thalamostriate Vein and Reperfusion

Reperfusion status was evaluated in 77 patients. The presence of the IPTSV was associated with a lower successful reperfusion rate (37.5% versus 60.0%; OR = 0.35; 95% CI, 0.13–0.92; P = .03) after adjusting for baseline diastolic blood pressure and onset-to-needle time.

In the subgroup of patients without reperfusion, patients with IPTSV had a significantly higher rate of poor outcome (80.0% versus 44.4%, P = .04) compared with those without IPTSV. IPTSV was not associated with poor functional outcome in patients with successful reperfusion (41.7% versus 25.9%, P = .46) (Fig 3).

DISCUSSION

Our study quantitatively described the relationship between deep cerebral veins and neurologic outcomes after IVT based on SWI.
and revealed that the presence of IPTSV (defined as an SDR of the TSV of ≥5%) was strongly associated with poor outcome after IVT. Without reperfusion, 80% of patients with IPTSV had poor outcomes, while this rate decreased to 42% if reperfused, which reinforced the importance of reperfusion treatment in patients with IPTSV.

In the current study, we found an association only between poor outcome and SDR of the TSV and not the SV or ICV. This may reflect the functional impact of the brain regions that drain to each vein. The TSV mainly drains the basal ganglia and thalamus, which play important roles in motor control and motor learning,14-16 and these functions have a strong influence on the mRS score.17 Lesions within the territory draining to the SV, such as the septum pellucidum, rostral corpus callosum, and white matter of the frontal lobe, may have less impact on the mRS score. The ICV receives tributaries from subcortical, periventricular structures and the choroid plexus within the lateral ventricle, and this extensive venous drainage territory may dilute the relationship between critical functions and changes in venous signal.

The presence of IPTSV was associated with a larger hypoperfusion volume at baseline. This finding is consistent with a previous study that demonstrated that prominent deep cerebral veins

![Figure 2](image1)

**FIG 2.** Outcome in patients with and without an ipsilaterally prominent thalamostriate vein. Two illustrative cases: a 55-year-old female patient with left-sided thromboembolic occlusion of the M1 segment (1A) and IPTSV (1A, arrow) who developed major infarct growth between pretreatment (1C) and 24-hour postthrombolysis imaging (1D). The 3-month modified Rankin Scale score was 5. A 54-year-old male patient with left-sided thromboembolic occlusion of the M1 segment (2B) without IPTSV (2A) who had similar pretreatment diffusion lesion volume (2C) but less infarct growth at 24 hours (2D). The 3-month mRS was 1.

![Figure 3](image2)

**FIG 3.** Relationship between categories of outcome with reperfusion and the presence of the ipsilaterally prominent thalamostriate vein. Note successful reperfusion is ≥70% reduction in hypoperfusion volume between pretreatment and 24 hours. Poor outcome is a 3-month modified Rankin Scale score ≥ 2. Asterisk indicates *P < .05.
on gradient-echo T2*-weighted imaging correlated with a large perfusion lesion. When the oxygen extraction fraction is elevated in hypoperfused tissue, the increased level of deoxyhemoglobin is reflected in abnormal hypointensity of the veins draining that region of the brain. Furthermore, we found that patients with IPTSV were likely to have a high proportion of penumbral tissue loss at 24-hour follow-up. The ischemic penumbra represents the potentially salvageable hypoperfused region, which retains structural integrity but has lost function after AIS. Without timely reperfusion, the penumbra becomes irreversibly damaged. Recent studies interpreted the presence of extensive prominent cortical vessels and asymmetry of deep medullary veins after AIS as evidence of poor leptomeningeal collateralization, because good collateralization should lead to less deoxygenated blood in the veins. Collateral status is one of the most important factors determining penumbral tissue loss, and good collaterals have been associated with reduced penumbral loss. Therefore, we speculated that the presence of IPTSV might indicate poor collaterals after acute ischemia, which would result in increased penumbral tissue loss. However, it is challenging to directly investigate the relationship between IPTSV and leptomeningeal collaterals in our data because we included patients with both large-artery occlusion and more distal occlusions. Further studies are required to explore this.

Our data also showed that patients with IPTSV were less likely to achieve reperfusion. Experimental studies have demonstrated venous endothelial injury and subsequent clot formation in cerebral veins due to oxidative stress and inflammation in the upstream ischemic area. Occlusion of microthrombi in venules was found within 30 minutes after acute reduction of CBF in mice, leading to decreased venous flow velocity. It is possible that the decreased venous outflow could result in an increase in intracranial pressure as blood accumulates in the capillary system. Subsequently, increased intracranial pressure may lead to decreased arterial perfusion, particularly in collateral pathways, and may contribute to the evolution of penumbral tissue into infarct.

Successful reperfusion appeared beneficial to patients with IPTSV, reducing the rate of poor outcome from 80% to 42%. This finding indicates the presence of salvageable penumbra in patients with IPTSV. Intra-arterial thrombectomy has been established as a more effective reperfusion therapy in patients with documented occlusion of the distal internal carotid or proximal middle cerebral artery. Our results may not directly affect clinical decision-making; however, further studies are required to evaluate whether patients with IPTSV are particularly likely to benefit from thrombectomy in addition to thrombolysis.

Our study has several limitations. First, it is a retrospective analysis and has a potential risk of selection bias. However, consecutive patients were included in a prospective stroke registry from our stroke center with highly homogeneous and standardized medical care. Second, both CT and MR imaging were used for follow-up imaging, which may increase the heterogeneity of volume measurement. However, the proportion of follow-up CT was <5% and was similar in patients with and without IPTSV. Third, we did not evaluate the follow-up status of veins due to the variable follow-up imaging, and this will be the subject of future study. Fourth, chronic arterial stenosis in cervicocranial arteries can result in preexisting abnormal signal in deep veins before an acute ischemic event. This may influence the assessment of asymmetry of the deep venous system after AIS. Our results should be validated in patients with stenosis of the internal carotid artery and intracranial arteries in further investigations. Finally, our conclusions were based on patients receiving IVT. Future studies are needed to validate the use of IPTSV in the context of endovascular thrombectomy.

CONCLUSIONS
Patients with IPTSV had a lower incidence of reperfusion and a higher rate of poor outcome after IVT. The risk of poor outcome was, however, markedly reduced in those patients in whom successful reperfusion was achieved. Our results emphasize the importance of reperfusion, especially for the patients with IPTSV. However, further studies are warranted to validate whether patients with IPTSV can benefit from more successful reperfusion strategies.

Disclosures: Min Lou—RELATED: Grant: Science Technology Department of Zhejiang Province [2013C3G201032] and the National Natural Science Foundation of China [81771070 & 81622017]. *Money paid to the institution.

REFERENCES


Restriction Spectrum Imaging Improves Risk Stratification in Patients with Glioblastoma


ABSTRACT

BACKGROUND AND PURPOSE: ADC as a marker of tumor cellularity has been promising for evaluating the response to therapy in patients with glioblastoma but does not successfully stratify patients according to outcomes, especially in the upfront setting. Here we investigate whether restriction spectrum imaging, an advanced diffusion imaging model, performed after an operation but before radiation therapy, could improve risk stratification in patients with newly diagnosed glioblastoma relative to ADC.

MATERIALS AND METHODS: Pre-radiation therapy diffusion-weighted and structural imaging of 40 patients with glioblastoma were examined retrospectively. Restriction spectrum imaging and ADC-based hypercellularity volume fraction (restriction spectrum imaging—FLAIR volume fraction, restriction spectrum imaging—contrast-enhanced volume fraction, ADC—FLAIR volume fraction, ADC—contrast-enhanced volume fraction) and intensities (restriction spectrum imaging—FLAIR 90th percentile, restriction spectrum imaging—contrast-enhanced 90th percentile, ADC—FLAIR 10th percentile, ADC—contrast-enhanced 10th percentile) within the contrast-enhanced and FLAIR hyperintensity VOIs were calculated. The association of diffusion imaging metrics, contrast-enhanced volume, and FLAIR hyperintensity volume with progression-free survival and overall survival was evaluated by using Cox proportional hazards models.

RESULTS: Among the diffusion metrics, restriction spectrum imaging—FLAIR volume fraction was the strongest prognostic metric of progression-free survival ($P = .036$) and overall survival ($P = .007$) in a multivariate Cox proportional hazards analysis, with higher values indicating earlier progression and shorter survival. Restriction spectrum imaging—FLAIR 90th percentile was also associated with overall survival ($P = .043$), with higher intensities, indicating shorter survival. None of the ADC metrics were associated with progression-free survival/overall survival. Contrast-enhanced volume exhibited a trend toward significance for overall survival ($P = .063$).

CONCLUSIONS: Restriction spectrum imaging—derived cellularity in FLAIR hyperintensity regions may be a more robust prognostic marker than ADC and conventional imaging for early progression and poorer survival in patients with glioblastoma. However, future studies with larger samples are needed to explore its predictive ability.

ABBRévIATIONS: CE = contrast-enhanced; CPH = Cox proportional hazards; FLAIR-HI = FLAIR hyperintensity; GBM = glioblastoma; HC = hypercellularity; HR = hazard ratio; MRSI = MR spectroscopic imaging; PFS = progression-free survival; OS = overall survival; RSI = restriction spectrum imaging; RT = radiation therapy; vf = volume fraction; vol = volume

Glioblastoma (GBM) is the most common and aggressive malignant primary brain tumor. Unfortunately, there has been only incremental improvement in the 5-year survival rate in the past decade. The standard of care for newly diagnosed GBM remains fairly uniform, with maximal permissible surgical resection followed by radiation therapy (RT) with concurrent and adjuvant temozolomide. Currently, novel molecular and cellular targeted therapies for treating GBMs are being investigated with many of

Received June 21, 2016; accepted after revision December 9.
them now in phase II clinical trials. With the advent of these new therapies and a recent study showing that radiation dose escalation to 75 Gy (above the standard dose of 60 Gy) is safe and possibly more effective in newly diagnosed GBM, stratification of patients at the highest risk for early progression is imperative because more aggressive or experimental treatments may be pursued in these individuals. These treatment decisions are usually considered within the first several weeks postsurgery once any residual tumor has been identified, making the pre-RT imaging pivotal for guiding the course of treatment.

Conventional MR imaging, including T1-postcontrast and FLAIR, is nonspecific because the former represents the breakdown of the blood-brain barrier due to tumor- and non-tumor-related causes, and the latter may represent tumor-related edema, postradiation change, or any cause of gliosis. Advanced MR imaging techniques, such as DWI, may offer more specific information related to the underlying physiology of the tissue and may complement existing measures. ADC estimates the magnitude of water diffusion in relation to the physical barriers in its environment. It is frequently used as an imaging biomarker for tumor cellularity and is inversely correlated with tumor cell density. However, at the typical b-values used clinically (b = 0, 1000 s/mm²), the diffusion signal primarily arises from the extracellular space. Therefore, in addition to estimating cell density, the ADC calculated at these b-values is also influenced by factors such as edema and necrosis, subsequently making ADC a rather nonspecific measure of tumor cellularity.

To account for the influence of edema and necrosis on ADC intensities in the tumor and peritumoral regions, histogram analysis of normalized ADC intensities (normalized with respect to mean ADC in normal-appearing white matter) and 2 Gaussian mixture modeling of the ADC intensities within the tumor have been proposed. However, these statistical methods seek only to reclassify voxels within an ROI so that voxels with presumably solid tumor are included in the analysis, while potentially problematic voxels that are confounded by partial voluming with edema and necrosis are removed. These methods have shown some promise for evaluating treatment response and predicting progression-free survival (PFS) in both the upfront and recurrent setting following treatment with antiangiogenic therapy. However, the utility of these ADC metrics for predicting response to standard chemoradiation has been less frequently explored. Some data suggest that though ADC intensities are not predictive of PFS or overall survival (OS) in the upfront setting, the volume of ADC with a large tumor burden (normalized ADC < 1.5; hypercellularity [HC] volume) within the T2 volume stratifies OS both pre- and postsurgery. However, it is unclear whether the hypercellularity volume was correlated with the underlying T2 volume and whether its predictive value merely reflects the association of the T2 volume with survival. Multiple studies have used the increase or decrease of HC volume fraction (ie, the HC volume defined with respect to the variation in a mixture of normal-appearing white and gray matter) as a predictive marker for evaluating treatment response because this metric may capture the percentage of the tumor that is highly cellular and correlated with the structural volumes. However, the utility of the HC volume fraction at individual time points for early risk stratification has not been explored, to our knowledge.

Table 1: Patient characteristics

| Resection Type: STR (n = 22): GTR (n = 18)* |
|-----------------------------|-----------------------------|
| Sex                         | 24 Male; 16 female          |
| Age (median) (range) (yr)   | 58 (31–84)                  |
| PFS (median) (range) (mo)   | 8.42 (3.5–50.53); 7 censored |
| OS (median) (range) (mo)    | 19.48 (6.37–50.93); 8 censored |
| Bevacizumab at recurrence   | 17 Patients                 |
| MGMT status                 | 12 Unmethylated, 8 methylated, 20 unknown |
| IDH status                  | 15 WT, 25 unknown           |
| EGFR amplification          | 7 Unamplified, 12 amplified, 21 unknown |
| EGRFVIII status             | 12 Positive, 9 negative, 19 unknown |

Note: —STR indicates subtotal resection; GTR, gross total resection; WT, wild type.

*Resection type was determined from the immediate postsurgical scan acquired within 48 hours of the operation for all patients, which included both T1-postcontrast and FLAIR sequences.

Multicompartment models of diffusion based on advanced multishell acquisitions can provide a more straightforward approach for mitigating the confounding effects of edema and necrosis at the voxel level. In particular, restriction spectrum imaging (RSI) is an advanced diffusion imaging model that separates the relative contributions of hindered and restricted signals originating from extracellular and intracellular water compartments, respectively, by using a multi-b-shell acquisition in conjunction with a linear mixture model. Furthermore, RSI incorporates geometric information to disambiguate isotropic-restricted diffusion in tumor cells from anisotropic-restricted diffusion in elongated neuronal processes (axons/dendrites collectively called “neurites”). Previous studies have demonstrated the increased sensitivity and specificity of RSI over ADC and DWI in both brain tumors and prostate cancer, and McDonald et al have recently demonstrated that RSI cellularity is a stronger predictor of both PFS and OS in patients following treatment with bevacizumab relative to ADC. However, its utility for predicting survival in patients newly diagnosed with GBM has not been explored, to our knowledge.

Here we investigate the application of RSI for risk stratification in newly diagnosed, resected GBMs. Our hypothesis was that RSI, due to its multi-b-shell acquisition and its inherent ability to decouple diffusion signal within tumor cells from that of extracellular pathology (eg, edema), would be a more robust marker of patient outcomes.

**MATERIALS AND METHODS**

This institutional review board–approved retrospective study included 45 patients with pathologically confirmed primary GBM who had pre-RT MRIs (median, 23 days; range, 9–113 days from the operation; median, 10 days; range, 1–29 days before start of RT) that included standardized RSI and conventional imaging sequences acquired between January 2011 and November 2015. All patients were followed for at least 6 months (May 2016). Patient characteristics are shown in Table 1. PFS and OS were defined relative to the pre-RT scan. All scans were reviewed by a neuroradiologist to ensure image quality and determine the basis for exclusion. Of the 45 eligible candidates, 17 patients underwent a second resection, with histopathology confirming tumor in 14 and showing predom-
MR Imaging Acquisition and Image Preprocessing

MR imaging was performed on a 3T Signa Excite HDx scanner (GE Healthcare, Milwaukee, Wisconsin) equipped with an 8-channel head coil. The imaging protocol included pre- and postgadolinium 3D volumetric T1-weighted inversion recovery spoiled gradient-recalled sequences (TE/TR = 2.8/6.5 ms, TI = 450 ms, flip angle = 8°, FOV = 24 cm, matrix = 0.93 × 0.93 × 1.2 mm) and a 3D T2-weighted FLAIR sequence (TE/TR = 126/6000 ms, TI = 1863 ms, FOV = 24 cm, matrix = 0.93 × 0.93 × 1.2 mm). For RSI, a single-shot pulsed-field gradient spin-echo EPI sequence was used (TE/TR = 96 ms/17 seconds, FOV = 24 cm, matrix = 96 × 96 × 48, voxel size = 2.5 mm) with 4 b-values (b = 0, 500, 1500, and 4000 s/mm²), with 6, 6, and 15 unique diffusion directions for each nonzero b-value, respectively (~8 minutes scan time).

Before analysis, raw data were corrected for geometric distortions due to susceptibility, gradient nonlinearities, and eddy currents.23 This correction was followed by correction of patient motion and rigid registration of the pre- and postcontrast 3D inversion recovery spoiled gradient-recalled images and the FLAIR images to each other by using in-house software. The diffusion maps were registered to the postcontrast images through the B0 images (b = 0 mm²/s volume), which were registered to the FLAIR images.

ADC values were calculated from a tensor fit to the b = 0, 500, and 1500 s/mm² data. Technical details of the RSI mathematic framework are described in their entirety elsewhere.15-17 and the model used has been applied in other recent publications.20 Briefly, the measured signal in each voxel was modeled as the sum of signals from 4 distinct tissue compartments: (1) the signal from water trapped within small spheric cells that is restricted in all directions, (2) the signal from water trapped in elongated neuronal processes (ie, neurites) that is restricted in the transverse direction, (3) the signal from extracellular water that is hindered by cells and neuronal processes, and (4) the signal from free water residing in CSF-filled compartments. RSI “cellularity” estimates were computed by combining the signal fraction from the intracellular compartment (1) with the isotropic restricted component of the neurite compartment (2). The RSI cellularity maps were then transformed to a standard z score by scaling each patient’s data by the population mean and SD in normal-appearing white matter of all patients in this study.

VOIs

Contrast-enhanced volumes (CEvol) and FLAIR hyperintensity (FLAIR-HI) volumes (FLAIRvol), excluding tumor and intrinsically T1 hypointense regions (ie, postsurgical blood products), were segmented semiautomatically (Amira software package; Visage Imaging, San Diego, California) on the coregistered postcontrast 3D inversion recovery spoiled gradient-recalled images and FLAIR images by a single expert image analyst with 8 years of experience. For patients whose pre-RT scan was acquired within 2 weeks after the operation, the immediate postsurgery scan was used to exclude areas of restricted diffusion caused by resection-induced cytotoxic edema.

Imaging Metrics

In this study, we chose to evaluate the utility of both the HC intensities and HC volume fractions as prognostic factors of PFS and OS. On the basis of a recent study,19 the 90th percentile of RSI cellularity values was selected for our analysis and the HC volume fraction was defined as the volume having an RSI cellularity z score of ≥1.5 within the CE and FLAIR-HI VOIs because the HC volume estimated with this pre-RT normalized ADC threshold was successful in stratifying survival.9 Imaging metrics included in the analysis were the following: FLAIR volume (FLAIRvol), CE volume (CEvol), the 90th percentile of RSI cellularity values in the CE (RSI-CE90%) and in the FLAIR-HI (RSI-FLAIR90%) VOIs and the RSI-based HC volume fraction in the CE (RSI-CEvf) and the FLAIR-HI (RSI-FLAIRvf) VOIs. The following ADC metrics were estimated for comparing against the respective RSI metrics: the 10th percentile of ADC11 in the CE (ADC-CE10%) and in the FLAIR-HI (ADC-FLAIR10%) VOIs and the ADC-based HC volume fraction (ADC z score of ≤1.5) in the CE (ADC-CEvf) and FLAIR-HI (ADC-FLAIRvf) VOIs.

Statistical Analysis

Univariate Cox proportional hazards (CPH) models were used to determine the contribution of resection type (subtotal resection, gross total resection), age, and sex to PFS and OS. Multivariate CPH models that included a single imaging metric combined with any significant clinical covariates (namely age, gender and resection type) were used to determine the relationship between each continuous pre-RT imaging metric and PFS/OS. Kaplan-Meier curves for the imaging metrics that were significant in the multivariate CPH models were obtained by dichotomizing the population on the basis of their median value and were compared by using a log-rank test. Due to the exploratory nature of the study, we did not control for type I error. P values < .05 were considered statistically significant. Statistical analysis was performed by using R Version 3.2.2 statistical and computing software (http://www.r-project.org/).23

RESULTS

Clinical

Median PFS in the final cohort (n = 40) was 8.4 months, and median OS was 19.5 months. Nine patients progressed within 6 months, 24 patients progressed within 12 months, 30 patients...
progressed within 18 months, and 33 patients progressed within 24 months. The results of univariate CPH analyses for the clinical and imaging metrics are summarized in Table 2. CPH models revealed that the resection type (subtotal resection-1 versus gross total resection-2) had a trend toward significance for PFS and was significant for OS (Table 2); hence, it was included as a covariate in multivariate CPH analyses. Age and sex (male-1, female-2) were not predictive of PFS or OS in this cohort. Boxplots of the imaging metrics are shown in Fig 1, and the boxplots split by the median PFS and OS excluding the values from the censored patients for the relevant metrics that achieved significance in Table 3 are shown in On-line Fig 1.

Table 2: Results of univariate CPH analyses of the continuous clinical imaging metrics

<table>
<thead>
<tr>
<th>Metrics</th>
<th>PFS HR (95% CI)</th>
<th>PFS P Value</th>
<th>OS HR (95% CI)</th>
<th>OS P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.562 0.992 (0.965–1.019)</td>
<td>.857 0.997 (0.965–1.030)</td>
<td>.857 0.997 (0.965–1.030)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>.063 0.512 (0.253–1.037)</td>
<td>.376 0.699 (0.316–1.546)</td>
<td>.376 0.699 (0.316–1.546)</td>
<td></td>
</tr>
<tr>
<td>Resection type</td>
<td>.066 0.511 (0.249–1.046)</td>
<td>.002* 0.235 (0.095–0.586)</td>
<td>.002* 0.235 (0.095–0.586)</td>
<td></td>
</tr>
<tr>
<td>FLAIRvol</td>
<td>.381 1.004 (0.994–1.014)</td>
<td>.404 1.005 (0.994–1.016)</td>
<td>.404 1.005 (0.994–1.016)</td>
<td></td>
</tr>
<tr>
<td>CEvol</td>
<td>.047* 1.041 (1.001–1.084)</td>
<td>.010* 1.06 (1.014–1.108)</td>
<td>.010* 1.06 (1.014–1.108)</td>
<td></td>
</tr>
<tr>
<td>ADC-FLAIRvf</td>
<td>.397 0.939 (0.881–1.087)</td>
<td>.52 0.936 (0.766–1.145)</td>
<td>.52 0.936 (0.766–1.145)</td>
<td></td>
</tr>
<tr>
<td>ADC-CEvf</td>
<td>.804 1.011 (0.925–1.105)</td>
<td>.932 0.995 (0.891–1.112)</td>
<td>.932 0.995 (0.891–1.112)</td>
<td></td>
</tr>
<tr>
<td>ADC-FLAIR10%</td>
<td>.384 1.001 (0.999–1.004)</td>
<td>.624 1.001 (0.998–1.004)</td>
<td>.624 1.001 (0.998–1.004)</td>
<td></td>
</tr>
<tr>
<td>ADC-CE10%</td>
<td>.735 0.999 (0.997–1.002)</td>
<td>.704 1.001 (0.998–1.003)</td>
<td>.704 1.001 (0.998–1.003)</td>
<td></td>
</tr>
<tr>
<td>RSI-FLAIRvf</td>
<td>.038* 1.031 (1.002–1.061)</td>
<td>.006* 1.051 (1.015–1.089)</td>
<td>.006* 1.051 (1.015–1.089)</td>
<td></td>
</tr>
<tr>
<td>RSI-CEvf</td>
<td>.12 1.03 (0.992–1.069)</td>
<td>.021* 1.048 (1.007–1.091)</td>
<td>.021* 1.048 (1.007–1.091)</td>
<td></td>
</tr>
<tr>
<td>RSI-FLAIR90%</td>
<td>.092 1.522 (0.934–2.482)</td>
<td>.028* 2.041 (1.081–3.855)</td>
<td>.028* 2.041 (1.081–3.855)</td>
<td></td>
</tr>
<tr>
<td>RSI-CE90%</td>
<td>.154 1.507 (0.858–2.649)</td>
<td>.099 1.822 (0.893–3.716)</td>
<td>.099 1.822 (0.893–3.716)</td>
<td></td>
</tr>
</tbody>
</table>

* Significant.

**CE and FLAIR-HI Volumes**

In a univariate CPH analysis, CEvol was significantly associated with PFS and OS (Table 2). When adjusted for resection type, CEvol was no longer significant for PFS and had a trend toward significance for OS (P = .063, Table 3). The Kaplan-Meier curves of the 2 groups obtained by a median split and compared by using the log-rank test were significantly different for PFS (χ²[1] = 4.7, P = .029) but not for OS (χ²[1] = 3.1, P = .078; Fig 2). Based on this stratification, the median PFS for the 2 groups was 218 and 338.5 days. FLAIRvol was not associated with PFS or OS.

FIG 1. Boxplots of the imaging metrics.
*RSI and ADC Metrics*

In a multivariate CPH analysis, RSI-FLAIR$_{90\%}$ was significantly associated with OS ($P = .043$) so that higher RSI intensities were associated with shorter survival (Table 3) and RSI-FLAIR$_{v_f}$ was significantly associated with both PFS ($P = .036$) and OS ($P = .007$, Table 3) so that higher HC volume fractions were associated with earlier progression and shorter survival.

Table 3: Results of multivariate CPH analyses of the continuous imaging metrics covaried with resection type

<table>
<thead>
<tr>
<th>Metrics</th>
<th>PFS</th>
<th>OS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$P$ Value</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td>RSI-FLAIR$_{v_f}$</td>
<td>.036*</td>
<td>1.033 (1.002–1.064)</td>
</tr>
<tr>
<td>RSI-CE$_{v_f}$</td>
<td>.258</td>
<td>1.022 (0.984–1.062)</td>
</tr>
<tr>
<td>RSI-FLAIR$_{90%}$</td>
<td>.097</td>
<td>1.535 (0.925–2.545)</td>
</tr>
<tr>
<td>RSI-CE$_{90%}$</td>
<td>.210</td>
<td>1.439 (0.814–1.543)</td>
</tr>
<tr>
<td>FLAIR$_{v_f}$</td>
<td>.845</td>
<td>1.001 (0.990–1.012)</td>
</tr>
<tr>
<td>CE$_{v_f}$</td>
<td>.165</td>
<td>1.029 (0.988–1.073)</td>
</tr>
</tbody>
</table>

* Significant.

None of the RSI metrics in the CE VOI were significant for PFS/OS.

The Kaplan-Meier curves of the 2 groups obtained by a median split of RSI-FLAIR$_{v_f}$ were significantly different for both PFS ($\chi^2 [1] = 6.1, P = .013$) and OS ($\chi^2 [1] = 7.7, P = .005$; Fig 2). The median PFS of the 2 groups was 201.5 and 367.5 days, and the median OS of the 2 groups was 451 and 750.5 days. Despite a strong trend, stratification by the median RSI-FLAIR$_{90\%}$ did not yield significant group differences in PFS ($\chi^2 [1] = 2.8, P = .095$) and OS ($\chi^2 [1] = 3.4, P = .065$).

To understand the influence of the threshold for defining HC volume fraction on the prognostic value, we repeated the analysis with HC volume fraction defined with a threshold of RSI...
and OS (P < 0.05) significantly associated with PFS (P < 0.05). Patient A had a shorter PFS and OS compared with patient B. Regionally, RSI-cellularity maps exhibited greater conspicuity in this region, reflecting higher tumor cellularity in the patient’s FLAIR-HI region. FLAIR-HI: z score maps acquired postsurgery but pre-RT for 2 patients, A and B. The VOI contours are shown in red and green for the CE and FLAIR-HI, respectively. Patient A is a 63-year-old man with a right posterior frontal GBM who underwent subtotal resection. This patient had high RSI cellularity in the FLAIR-HI region. Although there is corresponding ADC hypointensity in this region, it is subtle and inconspicuous. He had a shorter PFS and OS (PFS, 4.2 months; OS, 6.6 months) than patient B. Patient B is a 31-year-old woman with a right frontal GBM. There are no areas of high RSI cellularity or low ADC signal in the FLAIR-HI or CE region. She had a correspondingly longer PFS (PFS, 14.5 months; OS, 19.9 months).

FIG 3. Shown here are the axial T1-postcontrast, FLAIR, ADC, and RSI-cellularity z score maps acquired postsurgery but pre-RT for 2 patients, A and B. The VOI contours are shown in red and green for the CE and FLAIR-HI, respectively. Patient A is a 63-year-old man with a right posterior frontal GBM who underwent subtotal resection. This patient had high RSI cellularity in the FLAIR-HI region. Although there is corresponding ADC hypointensity in this region, it is subtle and inconspicuous. He had a shorter PFS and OS (PFS, 4.2 months; OS, 6.6 months) than patient B. Patient B is a 31-year-old woman with a right frontal GBM. There are no areas of high RSI cellularity or low ADC signal in the FLAIR-HI or CE region. She had a correspondingly longer PFS (PFS, 14.5 months; OS, 19.9 months).

cellularity z score ≥1 and ≥2. RSI-FLAIRz at z ≥ 1 was significantly associated with PFS (P = 0.03, hazard ratio [HR] = 1.023) and OS (P = 0.042, HR = 1.027). While RSI-FLAIRz at z ≥ 2 was significant for OS (P = 0.029, HR = 1.262) it was not significant for PFS (P = 0.09, HR = 1.043). Similarly, we also explored the prognostic value of the absolute HC volume with z ≥ 1.5, which were prognostic for OS (HCvol in CE: P = 0.018, HR = 1.543; FLAIR-HI: P = 0.042, HR = 1.262) in a multivariate CPH analysis. HCvol within CE was significant for PFS in a univariate analysis (P = 0.025, HR = 1.141) but was not significant in a multivariate analysis with resection type as a covariate.

In a univariate CPH analysis, none of the ADC metrics were associated with PFS (Table 2). The absolute ADC HC volumes with z ≤ 1.5 within the FLAIR-HI VOI were significantly prognostic of OS in a univariate CPH analysis (P = 0.015, HR = 58.89) but only had a trend toward significance (P = 0.079, HR = 20.06) after accounting for resection type in a multivariate CPH analysis and were not associated with PFS (univariate: P = 0.138, HR = 4.835). The ADC HC volumes within the CE VOI were not associated with outcomes.

Representative images of 2 patients in this cohort with short (patient A) and long (patient B) PFS are shown in Fig 3. Patient A had lower CEvol, FLAIRvol, and ADC-FLAIR10% but higher RSI-FLAIRz0/10%, RSI-FLAIRz1, and ADC-FLAIRz1 than patient B, likely reflecting higher tumor cellularity in the patient’s FLAIR-HI region. RSI-cellularity maps exhibited greater conspicuity in this region compared with ADC maps. Accordingly, patient A had shorter PFS and OS compared with patient B.

**DISCUSSION**

In the management of a highly aggressive tumor like GBM, the ability to stratify patient survival postoperatively is important because this information can directly impact therapeutic decision-making. In this study, we found RSI metrics to better stratify patients according to both PFS and OS compared with conventional imaging and ADC metrics. Unfortunately, CEvol and FLAIRvol may have limited prognostic value in this setting once the extent of the resection is considered. Patients with glioblastoma often have a more infiltrative tumor pattern and possibly a more hypoxic tumor biology, with some of them responding poorly to radiation therapy or conventional chemotherapy. The nonspecificity of these conventional metrics could potentially be due to the hypoxic tumors not having contrast enhancement and the presence of infiltrative tumors not having a noticeable signal in both T1 and T2 images. Hence, there is a need for better imaging metrics that perform reliably and may aid in identification of patients with GBM at high risk for early recurrence and worse survival.

Among the imaging metrics, we found that the pre-RT RSI-derived measures of cellularity within the FLAIR-HI region were associated with PFS and OS. The prognostic value of RSI appeared robust to the threshold used to determine the HC volume. Conversely, ADC metrics within the FLAIR-HI region were not associated with outcomes. Tumor progression results in areas of increased tumor cellularity (ie, decrease in diffusivity) and also areas of increased edema (ie, increase in diffusivity), both of which can occur simultaneously within an imaging voxel. Furthermore, postsurgery but pre-RT, a significant portion of the enhancing HC volume of the tumor has probably been resected and the residual HC tumor is interspersed with edema. Because ADC is a composite measure, the effects of these 2 opposing factors may cancel each other out, therefore limiting the prognostic value of ADC. RSI overcomes this limitation by separating the diffusivities associated with intracellular, restricted diffusion from the extracellular effects of edema. In contrast to a previous study, the absolute ADC HC volume was not significantly associated with OS, and this finding might be due to the inclusion of resection type as a covariate in the analysis. Even though the absolute RSI HC volume was prognostic of OS, the volume fraction was a stronger prognostic metric of PFS/OS. We did not find an association between any of our diffusion metrics in the CE region, similar to findings of previous studies. Although the reason for this outcome is not clear, it is likely that the CE region was quite limited in size in most patients, given the short interval between surgery and RT.

Other advanced imaging techniques, including perfusion, MR spectroscopic imaging (MRSI),26 and PET27 have also been shown to stratify outcomes in patients with GBM. While perfusion imaging provides information on vascular density and flow, MRSI and PET provide metabolic information. Pre-RT perfusion
imaging metrics have been shown to predict PFS, but perfusion is limited in the central core of a solid tumor with hypoxic cancer cells and with antiangiogenic therapy in which the vasculature is normalized, removing the leaky vessels in the tumor. Pre-RT MRSI metrics are associated with both PFS and OS but have lower spatial resolution compared with conventional, perfusion, and diffusion MR imaging. Although diffusion metrics should perform better than perfusion and MRSI with hypoxic tumor and antiangiogenic therapy, none of the previous studies have shown an association of pre-RT diffusion with outcomes. Here, we show that RSI-based cellularity is a prognostic metric of PFS and OS and may offer advantages over perfusion and MRSI in the pre-RT setting.

One of the main limitations of the current study is the small sample size obtained at a single institution (University of California, San Diego). The current study represents our effort to explore the clinical utility of RSI in a highly controlled study in which all the patients were scanned in the same scanner and all images were processed in a highly uniform manner, which included robust corrections for motion and geometric distortions. Given our modest size, we might be underpowered to detect smaller associations between some of our imaging metrics and survival. The generalizability and reproducibility of our results will need to be tested in prospective, multisite clinical trials across multiple vendor platforms. A second limitation is that the heterogeneity of the therapeutic approaches that each patient received following standard RT and temozolomide made it difficult to stratify patients according to treatment regimen. However, our major finding is that RSI performed better than ADC and conventional imaging in the same patient cohort, which is not confounded by between-patient treatment variance. Future studies with cohorts large enough to stratify patients according to the additional therapies received would be of great benefit for better delineating predictors of response to various therapies. Another possible limitation is that the tumors were segmented by a single imaging expert. Thus, intra- and interobserver variability of the tumor segmentations was not evaluated. In addition, although all patients had pathology-confirmed GBMs at the outset, histologic validation of tumor progression was not available for all patients. Furthermore, correlation with genomic information and molecular markers (such as MGMT and IDH status), which are known to provide prognostic information, could not be performed because these data were only available for a subset of the cohort (Table 1). Molecular information is now systematically collected for patients with GBM at our institution, facilitating future studies investigating the prognostic value of these markers. Although the results from our recent preclinical study show that the RSI cellularity metric correlates with histopathologic markers of cellularity, additional validation in patients with GBM is warranted.

CONCLUSIONS
Following the operation but before initiating RT, RSI-derived cellularity in the FLAIR-HI region performed better than ADC and conventional imaging for risk stratification in patients with GBM. Therefore, RSI could be potentially useful for identifying patients at highest risk for early progression and shorter survival. However, future studies with larger sample sizes are needed to explore its predictive ability.

ACKNOWLEDGMENTS
We would like to thank patients at the University of California, San Diego, Moores Cancer Center Neuro-Oncology Program, for their generous participation.

REFERENCES
10. Li Y, Lupo JM, Polley MY, et al. Serial analysis of imaging parame-
ters in patients with newly diagnosed glioblastoma multiforme. 
Neuro Oncol 2011;13:546–57 CrossRef Medline


A Multiparametric Model for Mapping Cellularity in Glioblastoma Using Radiographically Localized Biopsies


ABSTRACT

BACKGROUND AND PURPOSE: The complex MR imaging appearance of glioblastoma is a function of underlying histopathologic heterogeneity. A better understanding of these correlations, particularly the influence of infiltrating glioma cells and vasogenic edema on T2 and diffusivity signal in nonenhancing areas, has important implications in the management of these patients. With localized biopsies, the objective of this study was to generate a model capable of predicting cellularity at each voxel within an entire tumor volume as a function of signal intensity, thus providing a means of quantifying tumor infiltration into surrounding brain tissue.

MATERIALS AND METHODS: Ninety-one localized biopsies were obtained from 36 patients with glioblastoma. Signal intensities corresponding to these samples were derived from T1-postcontrast subtraction, T2-FLAIR, and ADC sequences by using an automated coregistration algorithm. Cell density was calculated for each specimen by using an automated cell-counting algorithm. Signal intensity was plotted against cell density for each MR image.

RESULTS: T2-FLAIR \( r = -0.61 \) and ADC \( r = -0.63 \) sequences were inversely correlated with cell density. T1-postcontrast \( r = 0.69 \) subtraction was directly correlated with cell density. Combining these relationships yielded a multiparametric model with improved correlation \( r = 0.74 \), suggesting that each sequence offers different and complementary information.

CONCLUSIONS: Using localized biopsies, we have generated a model that illustrates a quantitative and significant relationship between MR signal and cell density. Projecting this relationship over the entire tumor volume allows mapping of the intratumoral heterogeneity in both the contrast-enhancing tumor core and nonenhancing margins of glioblastoma and may be used to guide extended surgical resection, localized biopsies, and radiation field mapping.

ABBREVIATIONS: CE = contrast-enhancing; HPF = high-power field; NE = nonenhancing

Glioblastoma is a complex malignancy characterized by heterogeneous radiographic and histopathologic features. This intratumoral heterogeneity, combined with the diffuse infiltration of glioblastoma, renders correlations between imaging and underlying tissue critical for better understanding of tumor behavior. Stereotactically localized biopsies allow this comparison of local radiologic characteristics across multiple sequences with histopathologic or molecular analysis and provide a means for understanding underlying tissue characteristics that lead to macroscopic appearances on MR imaging. However, to date, histopathologic-radiographic correlations have been limited by the following: 1) the surgical challenge of obtaining stereotactically localized biopsies with reliable accuracy, 2) the maintenance of spatial fidelity across multiple MR images, and 3) the application of quantitative histopathologic analysis.

Radiographically, glioblastoma typically appears as a contrast-enhancing (CE) mass with surrounding nonenhancing (NE) tissue marked by abnormal T2-FLAIR and DWI-derived ADC signals. These heterogeneous MR signal characteristics in the NE area are known to represent a combination of vasogenic edema.
and infiltrating glioma cells, however, correlating these MR imaging abnormalities with underlying histopathology remains challenging. Nonetheless, there has been much effort with a variety of MR images to draw such radiologic-histologic correlations to better understand the tumor environment, which is critical for refining radiation treatment fields, assessing response to therapy, and evaluating progression.

Early studies have suggested that tumor cellularity as a surrogate for infiltration correlates with decreased apparent diffusivity, reduced T2-FLAIR hyperintensity, and increased gadolinium-related T1 signal. More recent effort has combined these changes on multiple MR images to help distinguish glioma cell infiltration and edema. Akbari et al used a machine-learning algorithm to identify several multiparametric features that spatially colocalized with areas of future radiographic recurrence, with the implication that these regions likely harbor underlying infiltrative tumor. However, this study did not confirm the proposed radiographic features with histologic data. Hu et al incorporated multiple MR imaging textural features with histologic data to identify areas of high tumor content for targeting diagnostic biopsies. The resulting model was able to sort tissue into 2 categories of high and low tumor content. Together, these initial studies suggested that important imaging patterns exist amid the complex signal changes within the peritumoral region and that careful analysis may yield correlates to tumor infiltration.

We hypothesized that quantitative measures of tissue cellularity from localized biopsies will correlate with voxel-level signal on multiple MR images and will provide more informative maps of tumor infiltration at the margins of glioblastomas. To this end, a fully automated cell counting and image coregistration pipeline was developed to eliminate subjective bias and improve reliability. With this approach, the objective of this study was to generate a model capable of predicting cellularity at each voxel within an entire tumor volume as a function of signal intensity, thus providing a means of quantifying tumor infiltration into surrounding brain tissue.

**MATERIALS AND METHODS**

**Patient Selection**
A retrospective review was performed on our data base of adult patients with MR imaging–localized biopsies obtained during open surgical resection for glioma at Columbia University Medical Center between January 2012 and January 2015. Patient selection was limited to those with newly diagnosed glioblastoma (World Health Organization grade IV) resected by 1 of 2 attending neurosurgeons (M.B.S., J.N.B.), who had complete preoperative MR imaging available for analysis (ADC, T2-FLAIR, and T1-weighted precontrast and postcontrast sequences). Patients were excluded if their localized samples were determined to be inadequate for any or any of the following reasons: poor neuronavigational registration, excessive MR imaging motion artifacts, errors in the automated screen capture, or problems with histologic processing pipeline.

**Image Acquisition**
All imaging was performed on a 3T MR imaging system (Signa; GE Healthcare, Milwaukee, Wisconsin) by using an 8-channel head-array coil (Signa HDxt; GE Healthcare). The following sequences were assessed for each patient: pre- and postcontrast T1-weighted, T2-FLAIR, and DWI. Pre- and postcontrast volumetric acquisition was performed with a T1-weighted 3D inversion recovery fast-spoiled gradient-recalled sequence with the following parameters: TI = 450 ms, TR = 10.2 ms, TE = 4.2 ms, flip angle $\alpha = 13^\circ$, FOV = 250 mm; matrix = 256 $\times$ 256, section thickness = 1.2 mm. The total scan time was approximately 4 minutes 15 seconds. Postcontrast images were acquired by using intravenous gadobenate dimeglumine (MultiHance; Bracco Diagnostics, Princeton, New Jersey) at a dose of 0.2 mL/kg. The time between injection and postcontrast imaging was approximately 5 minutes. Axial T2-FLAIR was performed with the following parameters: TR/TE = 9500/127 ms, TI = 2250 ms, section thickness = 5 mm with no gap, FOV = 225 mm. DWI was performed with a single-shot, spin-echo, echo-planar sequence with the following parameters: TR/TE = 7000/73 ms, FOV = 220 mm, matrix = 128 $\times$ 192, section thickness = 5 mm with no gap, bandwidth = 1953 Hz/pixel with b-values of 0 and 1000 s/mm². ADC maps were generated in FuncTool software (GE Healthcare) for the quantitative determination of mean diffusivity measurements.

**Image Preprocessing**
Raw image data from T1-precontrast, T2-FLAIR, and ADC maps were aligned to the volumetric T1-postcontrast sequence by using the FMRIB Linear Image Registration Tool (FLIRT; http://www.fmrib.ox.ac.uk/). The linear affine transformation algorithm was implemented by using 12 df, trilinear interpolation, and a mutual information cost function. Each sequence was further processed with a histogram normalization algorithm to standardize intensity values among patients. The algorithm is based on a method previously described that maximizes the cross-correlation between cumulative histogram distribution functions of each input volume with a reference standard. A mean reference template was created for each sequence by pooling data from the study subjects, each set to a mean intensity of 0 and an SD of 1. T1-postcontrast subtraction images were obtained by subtracting coregistration-normalized T1-precontrast volumes from T1-postcontrast volumes.

**Biopsy Acquisition**
All tissue sampling was performed within the normal surgical plan when it posed no additional risk to the patient. Samples were taken from the CE tumor core in all cases and the NE region around the tumor in some cases. Sampling of the CE, T2-FLAIR hyperintense region was permissible in 2 scenarios: 1) when the surgical trajectory required dissection through brain parenchyma en route to the tumor, and 2) when resection of tissue lateral or superficial to the CE region was part of the surgical goal. Biopsies were obtained before tumor debulking to maximize the spatial fidelity of localization. Stereotactic guidance was provided by a volumetric T1-weighted postcontrast sequence on a neuronavigation system (Brainlab Curve; Brainlab, Feldkirchen, Germany). The location of the biopsy was recorded by screen capture of the surgical navigation system at the time of tissue sampling.
Histopathologic Processing and Analysis

The samples were divided into 2 pieces in the operating suite. One piece was immediately flash frozen in liquid nitrogen for future genetic studies. The second piece was fixed in 10% (volume/volume) formalin and infiltrated with low-temperature paraffin for histologic analysis. Five-micrometer sections were stained with hematoxylin-eosin. The slides were then scanned and digitalized to BigTIFF files (http://bigtiff.org/) at ×40 magnification by using a Leica SCN400 system (Leica Biosystems, Wetzlar, Germany). The high-resolution image was acquired at 40,000 pixels per centimeter (2.5 × 10⁻⁵ cm per pixel).

Whole-Slide Cell Counting

A fully automated cell-counting algorithm was developed by using a convolutional neural network (Fig 1 and Online Fig 2). The algorithm was trained to count the number of nuclei present in a single high-power field (HPF), defined by a 2.25 × 2.25 mm square (900 × 900 pixels). Subsequently, the neural network was iteratively processed through tiles of HPFs until the entire H&E slide was counted (Fig 1B–D). HPFs at the border of the tissue specimen (eg, only a fraction of the field contained cells) were excluded from analysis.

The neural network was trained to classify the center pixel of a 21 × 21 × 3 RGB color input into 1 of 3 categories: cellular nuclei, densely staining artifacts mimicking cellular nuclei, and background staining. After a hyperparameter grid search, the final neural network architecture consisted of 5 convolutional layers of 5 × 5 filters (with increasing channel filter depths of 20, 40, 60, 80, and 120) followed by a final fully connected layer. Rectified linear activation functions were used after each convolutional layer, and a final softmax loss function was used to classify each input image. The network was trained on 5000 annotated examples of each class and validated on separate 1000 ground truth examples. A final validation classification error of 3.8% was achieved after training for 100 epochs.

As a final test set validation, 25 high-power fields were chosen at random. Each field was manually inspected to determine the number of nuclei present. The same field was then evaluated by the automated cell-counting algorithm. The human reviewer was blinded to the final total cell count obtained by the automated method. A high correlation was observed between the automated algorithm and the manual cell counts as determined by a Pearson coefficient ($r = 0.984$, Fig 2A).

Statistical Analysis

For each biopsy, MR signal intensity (ADC, T2-FLAIR, and T1-weighted postcontrast) was correlated against cell density per HPF. Single variable linear regression analysis and Pearson correlation coefficients were determined. In addition, a multiple linear regression model was created as defined by:

$$ F(x) = a x_1 + b x_2 + c x_3 + d, $$

where the predicted cellularity, $F(x)$, was a linear function of signal intensity on ADC, T2-FLAIR, and T1-weighted postcontrast sequences ($x_1$, $x_2$, and $x_3$, respectively).
To quantify the effects of potential biopsy misregistration, we repeated the above linear correlations iteratively by replacing the original spheric ROI with concentric shells of voxels at an increasing distance from the biopsy center. The resulting changes to the correlation of the model with increasing distance were plotted and fit with a sigmoidal regression curve. All statistical and image analyses were performed by using Matlab 2015b (MathWorks, Natick, Massachusetts).

RESULTS
Patient Population
Thirty-six patients were included in this study, of whom 15 were male and 21 were female. The mean patient age at the operation was 65.2 years (range, 24–88 years). Eight patients were excluded due to poor neuronavigational registration, excessive motion artifacts, or errors in radiographic or tissue processing. A total of 91 localized biopsies were analyzed (median, 2 samples per patient; range, 1–6 samples). Each biopsy point was co-registered individually on ADC, T2-FLAIR, and T1-weighted postcontrast subtraction sequences, yielding 273 sets of data points for analysis.

Cell Density
Ninety-one H&E slides, 30,578 HPFs (mean, 336 HPFs per sample), were analyzed.

The median cell density was 98 nuclei per HPF (range, 4–296 nuclei) among all biopsies. The median was used as a representative measure for cell density in an H&E slide, given the presence of both outlying high- and low-density HPFs in almost all tissue samples (Fig 1B, C). Other percentile cell densities were also considered, but it was shown that these represented an approximately linear translation of the median within any given H&E slide (Fig 2B, C). For example, the 98th percentile cell density of any given H&E slide was directly linearly correlated to the median cell count ($r = 0.901$), displaced by approximately 50 additional nuclei/HPFs. Thus, although the following analysis was evaluated with the median cell density, the linear nature of the relationship between signal intensity and cellularity would remain valid across most cell density percentiles.

Correlation between Signal Intensity and Cell Density
Scatterplots were used to assess the relationship between MR signal intensity and cellularity on multiple sequences (Fig 3). The correlation between ADC and cellularity was $r = 0.63$; the corre-
The correlation between T2-FLAIR and cellularity was \( r = 0.61 \); and the correlation between T1-weighted postcontrast subtraction and cellularity was \( r = 0.69 \). A multivariable linear regression was performed to assess the predictive power of the combined MR images. The Table shows the results of this analysis. The combined model correlation improves to \( r = 0.74 \) with a model variance of \( R^2 = 0.55 \).

**Effects of Possible Misregistration**

The ability to correlate voxel-level signal intensities and histopathology is contingent on accurate biopsy colocalization. To estimate the detrimental effect of possible misregistrations by our automated coregistration algorithm, we compared the relationship between MR signal intensity at varying degrees of misregistration error. The spheric ROI used for the model correlation was systematically replaced with voxels displaced at 0.5-mm increments from the center up to 5 mm of displacement. For example, the first ROI included all voxels located 0.5 mm from the center, the second ROI included all voxels 1.0 mm from the center, and so on, resulting in a series of concentric spheric shells. The correlations between signal intensity and these projected biopsy locations are shown in Fig 4. As expected, the correlation decreased as a function of distance from the original biopsy site. After 5 mm of displacement, near-zero correlation was observed for ADC and T2-FLAIR sequences, suggesting that our registration algorithm was accurate to \( \leq 5 \)-mm spatial resolution. Of note, increasing sampling distance beyond this threshold for T1-weighted postcontrast subtraction sequences inverts the correlation from positive to negative. This outcome is likely because beyond a certain distance, voxels originally within the CE tumor are now predominantly in the NE region, while voxels within the NE portion of the tumor likely begin to incorporate portions of CE tumor, thus inverting the expected relationship between enhancement and cellularity.

**FIG 3.** Cell count versus MR signal intensity. Scatterplots demonstrate median cell density as a function of signal intensity on ADC (A), T2-FLAIR (B), and T1-postcontrast subtraction sequences (C) correlated by using single-variable regression analysis. The linear regression and Pearson correlation \( r \) were significant \( (P < .05) \) for all 3 sequences. D. The scatterplot shows the actual and predicted cell counts as estimated by combining all 3 imaging modalities in a multiple-variable regression model.
Cell Density Map

Estimates for cellularity were projected at each voxel within the tumor volume by using the multivariable linear regression model and coregistered MR imaging sequences (Fig 5). Given that all biopsies were obtained within the boundaries of T2-FLAIR signal abnormality, the projected model estimates are limited to these regions.

DISCUSSION

In this study, we describe a voxel-level multiparametric MR imaging model of glioblastoma cellularity derived from the radiologic and histologic features of radiographically localized biopsies. We found that tumor cellularity is inversely correlated with ADC and T2-FLAIR signal and directly correlated with T1-weighted postcontrast signal. A multiparametric linear regression model based on these correlations captured approximately 54% of the variance in cellularity within the tumor. Projection of these estimates throughout the tumor volume provided a noninvasive map of glioblastoma cellularity, suggesting a means of visualizing infiltrative margins that may be derived from routine standard imaging. Clinically, this model may be used to guide extended surgical resection, localized biopsies, and/or radiation field mapping. Furthermore, it can be used as an outcome measure for tracking progression and response to treatment.

Cellularity was inversely correlated to ADC signal; this correlation is consistent with the notion that water diffusion is restricted in hypercellular neoplastic environments. This concept would suggest reduced ADC signal from frank tumor tissue in the area of CE region, but also from infiltrating tumor in the surrounding NE area. Our results are concordant with prior studies using localized biopsies in the CE area, which found an inverse relationship between ADC and cellularity in the CE core. However, studies that used this technique in the peritumoral region yielded contradictory results. Sadeghi et al failed to identify any significant relationship of ADC and cellularity in the peritumoral region; however, this study consisted largely of lower grade gliomas and the peritumoral region was defined histologically rather than radiologically. Conversely, the postmortem analysis of ex vivo tissue samples of LaViolette et al found a significant inverse relationship between ADC and cellularity in areas of tissue coregistered to peritumoral T2-FLAIR hyperintensity. In summary, the mixed results and differences in techniques yielded no strong conclusions but, together with our results, suggested that ADC offers some information in characterizing the peritumoral region. However, additional study is necessary to understand precisely how tumor histology is represented by ADC.

Cellularity was also inversely correlated with T2-FLAIR signal
nalization, we performed an analysis by using images obtained by intratumoral blood products or dystrophic mineralization, which are indicative of peritumoral edema, and subsequent leakage of gadolinium. A multiple linear regression model combining these individual relationships improved overall correlation from 61%–69% to 74%. This improvement confirms that these MR images contribute different and complementary information about the magnetic properties of the tumor environment. Specifically, our model is informed by tumor effects on Brownian motion of water, tissue heterogeneity in vasogenic edema, and the degree of tumor-induced blood-brain barrier breakdown.

Multiparametric modeling has been previously leveraged to investigate the imaging surrogates of peritumoral infiltration. Akbari et al.14 used imaging evidence of tumor recurrence in lieu of histologic markers for cellularity to identify MR imaging features suggestive of glioblastoma infiltration, yielding a predictive model for recurrence. Hu et al.18 expanded on this technique with histologic correlation, by using textural features extracted from multiple MR images to identify regions of tumor-rich biopsy targets (defined as ≥80% tumor content) in both the CE and NE regions. Our investigation builds on the described previous work for multiparametric analysis by implementing quantitative analysis of cellularity in the localized biopsies from both CE and NE regions.

This continuous variable provides a predictive model of underlying tissue heterogeneity.

Furthermore, this study relates radiologic and histologic characteristics at a more granular level, estimating cellularity on a voxel-by-voxel basis. This feature necessitates accurate and precise biopsy acquisition, coregistration of multiple imaging sequences, and quantitative histologic analysis of tissue samples. This effort remains nontrivial, with several potential sources of error, including the following: neuronavigational registration to the patient’s cranial features, intraoperative brain shift, initial conversion of the biopsy image capture to 3D MR imaging space, differences in acquisition matrices and patient motion, and analysis of whole-tissue specimens for full representation of their heterogeneity.2,28

Measures were taken intraoperatively, and downstream processing was fully automated to minimize these potential sources of errors. By using changes in the model correlation as a reference for deleterious effects secondary to possible misregistration, we demonstrate that gradually displacing our original biopsy coordinate by >0.5 mm causes an overall loss of signal correlation with tumor cellularity. This outcome suggests that the coordinates derived from the neuronavigation system screen capture correspond well to the true location of tissue sampling. Furthermore, the cell-counting algorithm provided unbiased measures of cell density over the entirety of these heterogeneous tissue samples.

This proposed multivariable model reflects simple and intuitive signal-intensity responses to varying tissue cellularity; however, ultimately only 55% of the variance in observed tumor cellularity. This correlation confirms earlier hypotheses of variations in T2-FLAIR hyperintensity. Specifically, areas of high cell density result in an intermediate intensity signal, while tissue with greater proportions of vasogenic edema is more uniformly hyperintense.13 This distinction was previously suggested by Akbari et al.,14 in which a mild loss in T2-FLAIR signal was found to be a significant predictor of cellular infiltration, leading to recurrent glioblastoma. Our results are the first to confirm this observation with histologic sampling, to our knowledge.

Conversely, cellularity was directly correlated with T1-shortening on gadolinium-enhanced sequences. This finding is in line with evidence of the perivascular migration of malignant cells, disruption of endothelial tight junctions, and subsequent leakage of gadolinium.16,17,26,27 Notably, instead of a bimodal distribution of T1 signal intensity corresponding to CE- versus-NE tumor, a broad range of MR imaging intensity values was observed. These observations corresponded to a similarly continuous broad distribution of cellularity. These findings suggest that even within apparently similar tumor compartments, marked histologic heterogeneity is present.

To account for the possible effects of intrinsic T1 shortening as may be seen with intratumoral blood products or dystrophic mineralization, we performed an analysis by using images obtained by subtracting T1-weighted precontrast and postcontrast data. Compared with results with native T1-weighted postcontrast images, the overall correlation in both single and multivariable models is slightly superior (On-line Fig 3). One potential explanation is the reduced confounding influence of intrinsic T1 shortening when using subtraction images, though no significant hemorrhage or dystrophic mineralization was identified on histopathologic evaluation of localized biopsy samples. A second possible source of improved correlation is that subtraction images facilitate an internal reference for intensity normalization so that local magnetic field inhomogeneities or other artifacts present on both pre- and postcontrast imaging would be reduced.
lularity is captured. The variance not accounted for by our model may stem from several sources: The first is localization error, which includes shifts in tissue position during an operation, errors in registration between the MR imaging and neuroanatomic system, and errors in registration in each patient’s multimodal MRIs. An additional source of unexplained variance comes from the MR imaging intensity of each voxel being affected by many physical and physiologic factors (eg, cell size and shape, water volume, diffusion anisotropy, and so forth). Although these factors are affected by tumor cellularity, a one-to-one relationship between cellularity and MR imaging intensity of any given sequence does not exist. While a multivariable approach, in part, reduces this phenomenon by minimizing the effects of uncorrelated noise, the proposed model is ultimately limited to linear relationships between signal intensity and cellularity, unable to account for the more spatially complex radiologic features of glioblastoma. The ability to model nonlinear behavior in signal intensity may be improved in future iterations by using machine-learning algorithms to implement both supervised and unsupervised feature detection.

While cellularity provided a simple metric for radiographic-histologic correlation, future studies would benefit from using localized biopsies to analyze additional histologic characteristics and molecular markers to capture other, more sophisticated features of biologic heterogeneity in glioblastoma. In fact, the use of machine-learning algorithms to implement both supervised and unsupervised feature detection may allow the model to account for potential complex and nonlinear radiographic-histologic relationships. The incorporation of additional physiologic imaging parameters such as dynamic susceptibility contrast perfusion, diffusion tensor imaging, and resting-state blood oxygen level-dependent imaging would likely further improve model prediction. Finally, the model may generalize to other types of gliomas (eg, low-grade gliomas); however, any model predictions beyond glioblastoma would require a validation similar to that performed in the current study.

CONCLUSIONS

This study demonstrates a significant correlation between voxel-level signal intensity and cell density in glioblastoma by single-variable regression as well as a more powerful multiparametric predictive model. Thus, with a precise and rigorous analysis pipeline, we have affirmed the feasibility of meaningful quantitative analysis at the voxel-level between signal intensity and localized biopsies. In addition, the multiparametric model proposed in this study provides a means to noninvasively map cell density at the infiltrative margins of glioblastoma. This characterization, supplemented by other histopathologic features, may be used to guide extended surgical resection, localized biopsy, and/or radiation field mapping, with significant implications for patient management.

Disclosures: Adam M. Sonabend—UNRELATED: Employment: Columbia University; Grants/Grants Pending: National Institutes of Health, Comments: I have a DNP National Institutes of Health grant from the Office of the Director, National Institutes of Health. Lawrence H. Schwartz—OTHER RELATIONSHIPS: consulting or advisory role company; Novartis, GlaxoSmithKline; research funding company; Eli Lilly, Astellas Pharma, Merck Sharp & Dohme, Pfizer, Boehringer Ingelheim; patent licensing for Varian Medical Systems. Peter Canoll—RELATED: Grant; James S. McDonnell Foundation, Irving Institute for Clinical and Translational Research: Pilot Research Award. Comments: These grants supported the analysis of the localized biopsies.

REFERENCES

18. Hu LS, Ning S, Eschbacher JM, et al. Multi-parametric MRI and
Multiparametric Evaluation in Differentiating Glioma Recurrence from Treatment-Induced Necrosis Using Simultaneous $^{18}$F-FDG-PET/MRI: A Single-Institution Retrospective Study


ABSTRACT

BACKGROUND AND PURPOSE: Differentiating glioma recurrence from treatment-induced necrosis can be a challenge on conventional imaging. This study aimed to assess the diagnostic performance of each functional MR imaging and PET parameter derived by using simultaneous FDG-PET/MR imaging individually and in combination in the evaluation of suspected glioma recurrence.

MATERIALS AND METHODS: Thirty-five treated glioma patients with 41 enhancing lesions (World Health Organization grade II = 9, III = 13, IV = 19) on MR imaging after an operation followed by radiation therapy and/or chemotherapy formed part of this study. Using PET/MR imaging, we calculated the normalized mean relative CBV, mean ADC, Cho/Cr, and maximum and mean target-to-background ratios. Statistical analysis was performed to determine the diagnostic performance of each parameter by receiver operating characteristic analysis individually and in combination with multivariate receiver operating characteristic analysis for the detection of glioma recurrence. Histopathology or clinicoradiologic follow-up was considered the criterion standard.

RESULTS: Of 35 patients, 25 (30 lesions) were classified as having a recurrence and 10 (11 lesions) patients as having treatment-induced necrosis. Parameters like rCBVmean (mean relative CBV), ADCmean, Cho/Cr, and maximum and mean target-to-background ratios were statistically significant in the detection of recurrent lesions with an accuracy of 77.5%, 78.0%, 90.9%, 87.8%, and 87.8%, respectively. On multivariate receiver operating characteristic analysis, the combination of all 3 MR imaging parameters resulted in an area under the curve of 0.913 $\pm$ 0.053. Furthermore, an area under the curve of 0.935 $\pm$ 0.046 was obtained when MR imaging parameters (ADCmean and Cho/Cr) were combined with the PET parameter (mean target-to-background ratio), demonstrating an increase in diagnostic accuracy.

CONCLUSIONS: Simultaneous PET/MR imaging with FDG offers correlative and synergistic multiparametric assessment of glioma recurrence with increased accuracy and clinical utility.

ABBREVIATIONS: AUC = area under the curve; CE = contrast-enhanced; max = maximum; rCBV = relative cerebral blood volume; ROC = receiver operating characteristic; SUV = standardized uptake value; TBR = target-to-background ratio

Glioma is currently managed by surgical resection followed by radiation therapy and/or chemotherapy, depending on the aggressiveness of the tumor. Most of these tumors, despite treatment, recur or progress during the course of the disease. Furthermore, treatment with chemoradiation therapy is associated with necrosis. Up to 30% of patients with glioblastoma on chemoradiation therapy develop treatment-related effects, which mimic tumor recurrence, and both present as a new enhancing lesion on contrast-enhanced (CE) MR imaging. This effect gets further complicated because recurrences most often occur within or adjacent to the primary tumor site. Moreover, recurrent glioma and radiation necrosis may coexist, further obscuring this differentiation. Hence, differentiation of tumor recurrence from treatment-induced necrosis, which remains a challenge on conventional CE-MR imaging, is important because their management and prognosis are completely different.

Advanced neuroimaging with MR imaging and PET has been proposed in the past for assessment of glioma recurrence. MR imaging, besides providing superior tissue contrast, also provides functional information through perfusion, diffusion, and spectroscopy. PET imaging, on the other hand, provides additional
information about tumor metabolism. FDG has been extensively used in the differentiation of glioma recurrence from radiation necrosis with variable results. High physiologic uptake of FDG in the normal brain and its uptake in inflammatory cells often result in poor tumor-to-background differentiation and make it a moderate test for characterization of tumor recurrence.\textsuperscript{10} This outcome has resulted in research into several more accurate PET tracers. Amino acid–based radio tracers such as \textsuperscript{11}C-methionine, \textsuperscript{18}F-fluoro-ethyl-tyrosine, and \textsuperscript{18}F-fluoro-l-thymidine are known to offer better imaging characteristics, however, are not easily available. \textsuperscript{11}C with a very short half-life needs an in-house cyclotron for production. \textsuperscript{18}FDG, being a commonly available tracer with low cost suitable for wider application, formed the basis of our work. The role of CT in the diagnosis of glioma recurrence is limited, with CE-MR imaging remaining the mainstay in neuroimaging.\textsuperscript{11} In an attempt to improve diagnostic accuracy, PET/CT has been used along with CE-MR imaging with limited success.\textsuperscript{12} With an aim to complement information for this purpose, multiparametric assessment with various functional MR imaging parameters such as choline/creatinine, choline/N-acetylaspartate, relative cerebral blood volume (rCBV), and apparent diffusion coefficient have been used in the past,\textsuperscript{13,14} whereas a few studies have tried coregistration of MR imaging with PET parameters such as standardized uptake value (SUV) and target-to-background ratio (TBR) performed separately on different occasions to improve diagnostic accuracy.\textsuperscript{12,15,16}

The advent of simultaneous PET/MR imaging has made it possible to assess all the parameters together in the same physical space in a single examination. We believe that simultaneous acquisition may help in overcoming some of the limitations of individual techniques and bring a synergistic effect in improved differentiation of recurrence from treatment-induced necrosis.

The present study aimed to assess the diagnostic performance of each of these parameters derived by using simultaneous FDG-PET/MR imaging individually and in combination by receiver operating characteristic (ROC) analysis for the evaluation of suspected glioma recurrence.

MATERIALS AND METHODS

Patients

In this retrospective study between March 2013 and September 2015, 41 consecutive patients who underwent simultaneous FDG-PET/MR imaging for glioma recurrence evaluation were chosen after obtaining prior approval of the institutional review board and signed informed consent. Inclusion criteria were histopathologic-confirmed glioma, previous treatment with an operation and radiation therapy with or without chemotherapy, and a high index of clinical suspicion for recurrence along with contrast enhancement on MR imaging. Exclusion criteria were nonglial primary brain tumors, proved malignancy of other sites, pregnancy, being younger than 18 years of age, standard contraindications to MR imaging, and loss of the patient to follow-up. Finally, 35 patients (6 women, 29 men; mean age, 50 ± 12 years) with 41 lesions (World Health Organization grade II = 9, III = 13, IV = 19) formed part of this study. Twenty-nine patients received chemoradiation therapy, whereas 6 patients received only radiation therapy. The interval between radiation therapy and FDG-PET/MR imaging ranged from 7 to 96 months, with a solitary case with a maximum of 233 months. The patient characteristics are available in Table 1.

Lesion Diagnosis

A combination of clinical and imaging follow-up and histopathology was considered the criterion standard. Patients with a disease-related adverse event, progressive disease on imaging, and/or a biopsy positive for viable tumor tissue were positive for recurrence. Patients who were stable or did not show any adverse events clinically and did not progress on imaging were considered positive for treatment-induced necrosis.

Instrumentation: Simultaneous PET/MR Imaging

Simultaneous FDG-PET/MR imaging was performed on a Biograph mMR scanner (Siemens, Erlangen, Germany). This system consists of a modified 3T system (Magnetom Verio; Siemens) with a fully functional PET system, equipped with avalanche photodiode technology. The MR imaging scanner features a high-performance gradient system (45 mT/m) with a slew rate of 200 T/m/s and is equipped with total imaging matrix coil technology. The PET scanner has a spatial resolution of 4.3 mm at 1 and 5.0 mm at 10 cm from the transverse FOV; its sensitivity is 1.47% at the center of the FOV and 1.38% at 10 cm.

Imaging Protocol

The patients fasted for 6 hours before intravenous FDG injection of $352.12 \pm 64.26$ MBq. PET/MR imaging started 45–60 minutes after injection for 25–30 minutes of PET acquisition, during which various MR imaging sequences were performed.

### Table 1: Patient characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (mean) (yr)</td>
<td>50 ± 12</td>
</tr>
<tr>
<td>Sex (No. of patients) (%)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>29 (82.8)</td>
</tr>
<tr>
<td>Female</td>
<td>6 (17.2)</td>
</tr>
<tr>
<td>Primary histopathology (No. of lesions) (%)</td>
<td></td>
</tr>
<tr>
<td>Glioblastoma</td>
<td>17 (41.5)</td>
</tr>
<tr>
<td>Anaplastic astrocytoma</td>
<td>9 (21.9)</td>
</tr>
<tr>
<td>Anaplastic oligodendroglioma</td>
<td>4 (9.8)</td>
</tr>
<tr>
<td>Oligodendroglioma</td>
<td>7 (17.1)</td>
</tr>
<tr>
<td>Others</td>
<td>4 (9.7)</td>
</tr>
<tr>
<td>WHO classification (No. of lesions) (%)</td>
<td></td>
</tr>
<tr>
<td>Grade II</td>
<td>9 (22.0)</td>
</tr>
<tr>
<td>Grade III</td>
<td>13 (31.7)</td>
</tr>
<tr>
<td>Grade IV</td>
<td>19 (46.3)</td>
</tr>
<tr>
<td>Primary site of glioma (No. of lesions) (%)</td>
<td></td>
</tr>
<tr>
<td>Frontal</td>
<td>15 (36.6)</td>
</tr>
<tr>
<td>Parietal</td>
<td>2 (4.9)</td>
</tr>
<tr>
<td>Temporal</td>
<td>10 (24.4)</td>
</tr>
<tr>
<td>Multilobar</td>
<td>14 (34.1)</td>
</tr>
<tr>
<td>Primary treatment (No. of patients) (%)</td>
<td></td>
</tr>
<tr>
<td>Operation + radiation therapy</td>
<td>6 (17.2)</td>
</tr>
<tr>
<td>Operation + radiation therapy + chemotherapy</td>
<td>29 (82.8)</td>
</tr>
<tr>
<td>Final diagnosis (No. of lesions) (%)</td>
<td></td>
</tr>
<tr>
<td>Recurrence</td>
<td>30 (73.2)</td>
</tr>
<tr>
<td>Histopathology</td>
<td>21 (51.2)</td>
</tr>
<tr>
<td>Clinicoradiologic follow-up</td>
<td>9 (22.0)</td>
</tr>
<tr>
<td>Treatment-induced necrosis</td>
<td>11 (26.8)</td>
</tr>
<tr>
<td>Histopathology</td>
<td>2 (4.9)</td>
</tr>
<tr>
<td>Clinicoradiologic follow-up</td>
<td>9 (21.9)</td>
</tr>
</tbody>
</table>

Note:—WHO indicates World Health Organization.
A simultaneous brain PET/MR imaging protocol was composed of a transversal T1WI ultrashort TE sequence (TR/TE, 11.94/0.07/22.46 ms) for attenuation correction and other MR imaging sequences for complete diagnostic evaluation of the brain, which included an axial FLAIR sequence (TR/TE, 7000/94 ms; TI, 2215.2 seconds; section thickness, 5 mm); a T2-weighted turbo spin-echo sequence (TR/TE, 4300/100 ms; section thickness, 5 mm); DWI (TR/TE, 4600/101 ms; b = 0, 400, 1000 s/mm²); PWI/perfusion EPI (TR/TE, 2550/31 ms); and 3D-encoded MPRAGE (TR/TE/TI, 1500/2.33 ms/900 seconds; spatial resolution, 1.2 × 1.1 × 1 mm) in the sagittal plane. In each case, 3D multivoxel 1H-MR spectroscopy (TR/TE, 1510/135 ms; FOV, 120 mm; matrix, 8 × 8 × 8 cm; acquisition, 1 average; scanning time, 5 minutes 41 seconds) was performed.

After the scan, all coincident data were sorted into a 2D-PET sinogram, which was subsequently reconstructed into transaxial sections, with an iterative 3D-ordered-subset expectation maximization algorithm with 3 iterations and 21 subsets, Gaussian smoothing of 4 mm in full width at half maximum, and a zoom of 1. The voxel size of brain PET images was 1.39 × 1.39 × 2.03 mm. MR imaging, PET, and PET/MR imaging scans were reviewed at a syngo.via platform (Siemens) by using the mMR general workflow.

**PET and MR Image Quality Control**

All PET/MR imaging studies were performed by using standard imaging protocol for both PET and MR imaging, adhering to the principle of the Quantitative Imaging Biomarkers Alliance protocol and standards (https://www.rsa.org/qbba/); in case of PET, tracer dose, time delay after injection, acquisition time; and, in case of MR imaging, consistency of imaging sequences. Distortion-correction-enabled diffusion-weighted images were acquired to minimize error in the ADC calculation. For PET images, emission data were corrected for randoms, dead time, and scatter, and attenuation correction was performed by using an MR imaging-based ultrashort TE sequence.

**Image Analysis**

Quantitative analysis of FDG-PET/MR imaging scans was performed by a radiologist and a nuclear medicine physician in consensus, with >10 years’ experience in diagnostic radiology or nuclear medicine, who were blinded to the clinical information, histopathologic data, clinicoradiologic follow-up information, and final diagnosis for each lesion.

**ROI Selection**

First, by subtraction of precontrast from postgadolinium T1WIs, we achieved a software-based postcontrast subtraction series wherever possible to accurately isolate areas of enhancing tumor. We calculated ADC on-line by using system software, applying a monoexponential model. The most representative spectra observed within/around the lesion showing maximum PET uptake was taken for calculation of Cho/Cr.

A freehand 2D ROI was drawn manually over the most representative section with maximum enhancement and was stored. This freehand ROI was duplicated by using the copy-paste function over the FDG-PET image, software generated ADC and CBV maps to find the respective maximum standardized uptake value (SUVMAX, mean standardized uptake value (SUVMean), mean ADC (ADCmean), and mean CBV (CBVmean) (Figs 1 and 2). Maximum target-to-background ratio (TBRmax), mean target-to-background ratio (TBRmean), and rCBVmean were calculated by dividing SUVMAX, SUVMean, and CBVmean obtained from the lesion ROI by the SUVMean, SUVMean, and CBVmean, respectively, from a separate ROI drawn for normalization on normal contralateral white matter.

**Statistical Analysis**

Data were checked for normality by using the Shapiro-Wilk test before statistical analysis. Individual PET/MR imaging parameters were evaluated for the difference between glioma recurrence and treatment-induced necrosis by using a 2-tailed independent Student t test or Mann-Whitney U test. Optimal-threshold value, area under the curve (AUC), sensitivity, specificity, positive predictive value, negative predictive value, and accuracy were calculated by using receiver operating characteristic analysis for each of the PET/MR imaging parameters for their ability to detect recurrence. Multivariate analysis of variance was performed to check whether a combination of parameters was better than individual parameters. A multivariate ROC analysis was performed to analyze the increment offered by a combination of various parameters by conducting a logistical regression analysis to generate a combined ROC curve of different combinations of parameters. All statistical analyses were performed with the SPSS software package (Version 17.0; IBM, Armonk, New York). For all statistical tests, a P value < .05 was considered a significant difference.

**RESULTS**

**Patient Information**

Thirty-five patients enrolled in this retrospective study had 41 lesions for evaluation. Five patients had multicentric lesions. Histopathology after repeat operation showed viable tumor in 16 patients (21 lesions) and evidence of treatment-induced necrosis in 2 patients (2 lesions).

Eight patients (9 lesions) without histopathologic evaluation were classified as having treatment-induced necrosis based on a stable clinical state on extended follow-up periods, absence of any new neurologic symptoms, and no remarkable increase of the lesion size and/or metabolic activity observed on follow-up PET/MR imaging up to a mean period of 11.8 ± 4.5 months (range, 7–24 months). Nine patients (9 lesions) were finally classified as having recurrent brain tumors due to the development of neurologic symptoms and a progressive increase in size on CE-MR imaging (1 patient) or PET/MR imaging (8 patients) during follow-up.

**Scan Information**

All 35 patients underwent simultaneous FDG-PET/MR imaging. One patient did not have perfusion MR imaging; in 4 patients (5 lesions), the MR spectroscopy study was not contributory because of noisy spectra, and it was not acquired in 2 patients (3 multicentric lesions) and hence was not included in the analysis.

Among considered parameters, rCBVmean, ADCmean, Cho/Cr, TBRmax, and TBRmean were found significant (P < .05) and used
for further combined analysis (Table 2 and Fig 3). The tumor size ranged from 0.92 to 9.5 cm².

Comparison of Advanced MR Imaging Parameters

PWI: rCBVmean. In this study, rCBVmean was not significantly higher in patients with recurrence than in patients with treatment-induced necrosis (2.41 ± 1.02 versus 1.82 ± 1.12, \( P = .082 \); Fig 3A). However, the diagnostic accuracy of rCBVmean for correct identification of glioma recurrence reached 77.5% with rCBVmean ≥ 1.71 (AUC = 0.680 ± 0.111, \( P = .008 \)), which was found to be significant (Table 2 and Fig 4A).

DWI: ADCmean. The ADCmean was significantly lower in patients with recurrence than in those with treatment-induced necrosis (1283.13 ± 210.96 versus 1558.55 ± 313.32 × 10⁻⁶ mm²/s, \( P = .015 \); Fig 3B). The diagnostic accuracy of ADCmean for correct identification of glioma recurrence reached 78% with ADCmean ≤ 1507 × 10⁻⁶ mm²/s (AUC = 0.752 ± 0.015, \( P = .013 \); Table 2 and Fig 4A).

MR Spectroscopy: Cho/Cr. Cho/Cr was significantly higher in patients with glioma recurrence than in those with treatment-induced necrosis (3.47 ± 2.20 versus 1.63 ± 0.64, \( P = .002 \); Fig 3C). The diagnostic accuracy of Cho/Cr for correct identification of glioma recurrence reached 90.9% by using Cho/Cr ≥ 1.405 (AUC = 0.861 ± 0.08, \( P < .001 \); Table 2 and Fig 4A).

Comparison of FDG Uptake Indices (TBRmax and TBRmean)
The TBRmax and TBRmean were significantly higher in patients with glioma recurrence than in those with treatment-induced necrosis (TBRmax, 2.50 ± 0.97 versus 1.53 ± 0.59, \( P = .001 \) and TBRmean, 1.70 ± 0.61 versus 0.98 ± 0.36, \( P < .001 \); Fig 3D). The diagnostic accuracy of TBR values for the correct identification of
recurrence of brain gliomas reached 87.8% with TBR\textsubscript{max} ≥ 1.58 (AUC = 0.827 ± 0.078, \( P < .001\); Table 2 and Fig. 4C) and also with TBR\textsubscript{mean} ≥ 1.18 (AUC = 0.888 ± 0.059, \( P < .001\); Table 2 and Fig 4C), though with a superior AUC.

Comparison of Combined Parameters
On multivariate analysis of variance, a statistically significant difference between glioma recurrence and treatment-induced necrosis (\( F_{5,26} = 5.871, P = .001\); Wilks \( \Lambda = 0.470\), partial \( \eta^2 = 0.530\) was observed.

Table 2: Diagnostic performance of individual parameters in the detection of glioma recurrence

<table>
<thead>
<tr>
<th>Parameter</th>
<th>rCBV\textsubscript{mean} ≥ 1.709</th>
<th>ADC\textsubscript{mean} ≤ 1507</th>
<th>Cho/ Cr ≥ 1.405</th>
<th>TBR\textsubscript{max} ≥ 1.579</th>
<th>TBR\textsubscript{mean} ≥ 1.179</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>82.8%</td>
<td>86.7%</td>
<td>100.0%</td>
<td>93.3%</td>
<td>90.0%</td>
</tr>
<tr>
<td>Specificity</td>
<td>63.6%</td>
<td>54.5%</td>
<td>66.7%</td>
<td>72.7%</td>
<td>81.8%</td>
</tr>
<tr>
<td>PPV</td>
<td>85.7%</td>
<td>83.9%</td>
<td>88.9%</td>
<td>90.3%</td>
<td>93.1%</td>
</tr>
<tr>
<td>NPV</td>
<td>58.3%</td>
<td>60.0%</td>
<td>100.0%</td>
<td>80.0%</td>
<td>75.0%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>77.5%</td>
<td>78.0%</td>
<td>90.9%</td>
<td>87.8%</td>
<td>87.8%</td>
</tr>
<tr>
<td>AUC±SE</td>
<td>0.68 ± 0.111</td>
<td>0.752 ± 0.015</td>
<td>0.861 ± 0.08</td>
<td>0.827 ± 0.078</td>
<td>0.888 ± 0.059</td>
</tr>
<tr>
<td>( P ) value</td>
<td>( .008^b )</td>
<td>( .013^b )</td>
<td>( &lt;.001^b )</td>
<td>( &lt;.001^b )</td>
<td>( &lt;.001^b )</td>
</tr>
</tbody>
</table>

Note: — PPV indicates positive predictive value; NPV, negative predictive value; SE, standard error.

*The ADC\textsubscript{mean} value is expressed as \( \times 10^{-6}\) mm\(^2\)/s and rCBV\textsubscript{mean}, Cho/Cr, TBR\textsubscript{max}, and TBR\textsubscript{mean} are ratios and hence unitless. \( P \) values mentioned are generated while calculating the AUC from the ROC analysis.

\( b \) \( P \) values of rCBV\textsubscript{mean}, ADC\textsubscript{mean}, Cho/Cr, TBR\textsubscript{max}, and TBR\textsubscript{mean} are \( <.05 \) and hence are statistically significant. \( P \) value reflects the significance of the ROC analysis of individual parameters represented in the table.
Combinational Analysis

Various combinations of several parameters were performed, and multivariate ROC analysis was performed as depicted in Table 3 (Fig 4B - D).

Combined Analysis of MR Imaging Parameters: Cho/Cr, rCBVmean, and ADCmean

Individually, Cho/Cr achieved a maximum AUC of 0.861 ± 0.080 in the detection of glioma recurrence. On multivariate ROC analysis, the addition of ADCmean to either rCBVmean or Cho/Cr resulted in an improved diagnostic capability of both rCBVmean and Cho/Cr as noted by an increment in the AUC of the combinations. The combination of all 3 MR imaging parameters resulted in a combined AUC of 0.913 ± 0.053 (Table 3 and Fig 4B).

Combined Analysis of MR Imaging and FDG Parameters: TBRmax, TBRmean, Cho/Cr, rCBVmean, and ADCmean

Individually, TBRmax has the maximum AUC (0.888 ± 0.059) of all the MR imaging and PET parameters. Among intervariable evaluation with multivariate analysis of variance for the predictability of diagnosis, ADCmean (P = .001) and TBRmean (P = .001) were found to be the most significant variables in predicting glioma recurrence. On multivariate ROC analysis, the maximum AUC of 0.935 ± 0.046 was achieved with a combination of ADCmean, Cho/Cr, and TBRmean (Fig 4D). Moreover, a combination of either TBRmax or TBRmean with ADCmean and/or Cho/Cr improved the AUC value significantly beyond that of individual combining parameters (Table 3).

Summary of Results

Among all individual parameters, Cho/Cr in MR imaging and TBRmean in PET are the most significant discriminators for the prediction of recurrence. Among MR imaging parameters alone, rCBVmean should be used in conjunction with Cho/Cr and ADCmean for differentiating recurrence from treatment-induced necrosis. The maximum AUC is achieved by combining TBRmean, ADCmean, and Cho/Cr.

DISCUSSION

Treatment-induced necrosis is a common treatment-related morbidity in the management of gliomas, and the rate of radiation necrosis increases with incorporation of temozolomide into high-grade glioma management. The differentiation of glioma
recurrence or progression from radiation injury can be a radiologic dilemma, irrespective of the imaging technique used.20

It is known that functional imaging methods complement anatomic information and yield different aspects of pathophysiology. The combination of PET and CE-MR imaging information with retrospective fusion has been reported to increase diagnostic accuracy over any single technique.12,15,16 In an integrated PET/MR system, PET and MR imaging information are not only combined but an additional value is expected owing to co-interpretation of both the signals.21

In our study, the results derived from individual PET and functional MR imaging parameters from simultaneously ac-

FIG 4. Receiver operating characteristic curves with their respective AUC values of MR imaging parameters (A) showing the high diagnostic performance of Cho/Cr in the detection of glioma recurrence. With multivariate ROC analysis, the ROC curve and AUC of all 3 MR imaging parameters combined show a significant increment in AUC over the individual MR imaging parameters (B). ROC curves with their respective AUC values of FDG parameters (C) show the high diagnostic performance of TBR mean in the detection of glioma recurrence. Multivariate ROC analysis, ROC curve, and AUC of the best performing FDG-PET/MR imaging combination of ADCmean, Cho/Cr, and TBRmean show a significant increment over individual MR imaging or PET parameters (D).

Table 3: Multivariate ROC analysis showing AUC±SE values for various combinations of FDG-PET and MRI parameters in the detection of glioma recurrence

<table>
<thead>
<tr>
<th></th>
<th>rCBVmean</th>
<th>ADCmean</th>
<th>Cho/Cr</th>
<th>ADCmean + rCBVmean</th>
<th>ADCmean * Cho/Cr</th>
<th>rCBVmean + Cho/Cr</th>
<th>ADCmean * rCBVmean</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADCmean (0.752 ± 0.085)</td>
<td>0.781 ± 0.079</td>
<td>0.800 ± 0.083</td>
<td>0.951 ± 0.053</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cho/Cr (0.861 ± 0.080)</td>
<td>0.860 ± 0.083</td>
<td>0.912 ± 0.051</td>
<td>0.931 ± 0.053</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBRmax (0.827 ± 0.078)</td>
<td>0.831 ± 0.081</td>
<td>0.864 ± 0.072</td>
<td>0.945 ± 0.059</td>
<td>0.820 ± 0.073</td>
<td>0.935 ± 0.044</td>
<td>0.889 ± 0.061</td>
<td>0.932 ± 0.046</td>
</tr>
<tr>
<td>TBRmean (0.888 ± 0.059)</td>
<td>0.884 ± 0.063</td>
<td>0.888 ± 0.058</td>
<td>0.935 ± 0.044</td>
<td>0.897 ± 0.057</td>
<td>0.935 ± 0.046°</td>
<td>0.928 ± 0.047</td>
<td>0.932 ± 0.048°</td>
</tr>
</tbody>
</table>

* Each entry represents the AUC ± SE of a combination of parameters mentioned in the respective rows and columns. Values in parentheses represent AUC ± SE of that particular parameter or combination. Blank cells are left to avoid repetition of values.

° A combination of FDG-PET and MR imaging parameters has a better AUC ± SE than a combination of MR imaging parameters.

° Note that the maximum AUC was achieved with a combination of ADCmean, Cho/Cr, and TBRmean.
quired FDG-PET/MR imaging are in agreement with the published literature.

With an optimized threshold of $1507 \times 10^{-6}\text{mm}^2/\text{s}$ for ADCmean, an AUC of 0.752 ± 0.015 was obtained, which is close to the reported threshold of $1490 \times 10^{-6}\text{mm}^2/\text{s}$ and AUC of 0.779.23 We achieved an accuracy of 77.5% and an AUC of 0.68 ± 0.111 with an optimized threshold of 1.709 for rCBVmean. Our calculated sensitivity and specificity of 82.8% and 63.6% are similar to 86% and 70%, respectively, reported with a comparatively lower threshold of 1.3.24 However, wide variation has been reported in rCBV thresholds to detect recurrence,24 and reasons ascribed include rapid extravasation of gadolinium-based contrast agent and vascular leak in the irradiated bed, especially with chemoradiation therapy.25 With an optimized threshold of 1.405 for Cho/Cr, we achieved an accuracy and AUC of 90.9% and 0.861, we achieved an accuracy and AUC of 90.9% and 0.861, respectively, which are in concordance with the reported values of 93.1% and 0.913, respectively, with 1.54 as a threshold.26 FDG-PET has been used to detect glioma recurrence with a varied range of reported sensitivities from 43% to 95% and specificities from 50% to 100%.9 Despite known limitations of FDG in the detection of glioma recurrence of any histology, a recent meta-analysis of FDG found a respectable pooled AUC of 0.866 ± 0.034 for FDG-PET.9 We achieved an AUC of 0.888 ± 0.059 by using an optimized threshold of 1.18 for TBRmean, the detection of glioma recurrence that is in agreement with the reported AUC of 0.898 by Nozawa et al16 in the differentiation of recurrent/residual high-grade glioma from posttreatment changes or low-grade glioma, however, by using TBRmax with a threshold of 1.8.

Because there is no single technique that can solve this clinical dilemma, researchers have attempted to combine functional MR imaging parameters to achieve an improved diagnostic accuracy beyond any single parameter.15,27

Moreover, in the past, retrospective fusion of FDG-PET with CE-MR imaging has aided in improved differentiation of glioma recurrence qualitatively, compared with FDG-PET or MR imaging alone.12,15,16 In a group of 30 patients with high-grade gliomas, Estrada et al12 reported an accuracy of 73% for FDG-PET, which improved to 93% with visual qualitative analysis of fused FDG-PET with CE-MR imaging.

Matsusue et al.,27 with a semi-quantitative multiparametric scoring system for ADC ratio, rCBV, Cho/Cr, and Cho/NAA in 15 patients, reported an accuracy of 93.3% in the detection of glioma recurrence. We achieved an AUC of 0.913 ± 0.053 by combining ADCmean, rCBVmean, and Cho/Cr by multivariate ROC analysis.

Our quantitative analysis of combined parameters also supports this body of evidence. We observed that combinations of individual parameters demonstrate superior results in comparison with individual parameters. We achieved an AUC of 0.913 ± 0.053 when all MR imaging parameters were combined, which further improved to 0.932 ± 0.046 when TBRmax or TBRmean was added, clearly highlighting the utility of combinational analysis. However, the highest AUC of 0.935 ± 0.046 was achieved by combining ADCmean, Cho/Cr, and TBRmean.

We found that a simultaneous PET/MR imaging study offers advantages such as being logically simple, implying, thereby, that it is performed as a single examination to obtain PET and MR imaging datasets; reduced examination time, thereby improving patient compliance; and correlative image reading and interpretation, which could help in localizing the most suspicious area on the PET image in voxel selection for MR spectroscopy. Furthermore, as observed by Mong et al.,28 we also found that PET uptake could help in accurate classification of suspicious ADC lesions. In the present study, we used an ultrashort TE sequence for MR imaging–based attenuation correction for brain PET, which is reported to result in accurate quantification of PET parameters,29 more importantly offering comparable accuracy of CT-based attenuation correction.

Our study has several limitations. It is a retrospective study with a small group and a relatively heterogeneous class of patients to perform robust statistical analysis. The present study only provides within-sample prediction characteristics and is not adequately sized to provide a decent out-of-sample prediction. Although desirable, histologic diagnosis in all the lesions could not be obtained due to ethical concerns associated with high morbidity and sampling error.30

Hence, further large-scale prospective multicentric trials are required with simultaneous PET/MR imaging and FDG, a commonly available radiotracer, to validate our results.

CONCLUSIONS

Simultaneous PET/MR imaging with FDG offers correlative and synergistic multiparametric assessment of glioma recurrence with increased accuracy and clinical utility.

ACKNOWLEDGMENTS

We thank Dr. Aru Singh for technical help in the statistical analysis.

REFERENCES


ABSTRACT

BACKGROUND AND PURPOSE: Current imaging assessment of high-grade brain tumors relies on the Response Assessment in Neuro-Oncology criteria, which measure gross volume of enhancing and nonenhancing lesions from conventional MRI sequences. These assessments may fail to reliably distinguish tumor and nontumor. This study aimed to classify enhancing and nonenhancing lesion areas into tumor-versus-nontumor components.

MATERIALS AND METHODS: A total of 140 MRI scans obtained from 32 patients with high-grade gliomas and 6 patients with brain metastases were included. Classification of lesion areas was performed using a support vector machine classifier trained on 4 components: enhancing and nonenhancing, tumor and nontumor, based on T1-weighted, FLAIR, and dynamic-contrast-enhancing MRI parameters. Classification results were evaluated by 2-fold cross-validation analysis of the training set and MR spectroscopy. Longitudinal changes of the component volumes were compared with Response Assessment in Neuro-Oncology criteria.

RESULTS: Normalized T1-weighted values, FLAIR, plasma volume, volume transfer constant, and bolus-arrival-time parameters differentiated components. High sensitivity and specificity (100%) were obtained within the enhancing and nonenhancing areas. Longitudinal changes in component volumes correlated with the Response Assessment in Neuro-Oncology criteria in 27 patients; 5 patients (16%) demonstrated an increase in tumor component volumes indicating tumor progression. These changes preceded Response Assessment in Neuro-Oncology assessments by several months. Seven patients treated with bevacizumab showed a shift to an infiltrative pattern of progression.

CONCLUSIONS: This study proposes an automatic classification method: segmented Response Assessment in Neuro-Oncology criteria based on advanced imaging that reliably differentiates tumor and nontumor components in high-grade gliomas. The segmented Response Assessment in Neuro-Oncology criteria may improve therapy-response assessment and provide earlier indication of progression.

ABBREVIATIONS: BAT = bolus arrival time; DCE = dynamic contrast-enhanced; GB = glioblastoma; HGG = high-grade gliomas; k = interstitium-to-plasma rate constant; v = interstitial volume; v = plasma volume; RANO = Response Assessment in Neuro-Oncology; SPGR = spoiled gradient-recalled; sRANO = segmented RANO; SVM = support vector machine

High-grade gliomas (HGG), specifically glioblastoma (GB), remain the most common and aggressive brain tumors in adults. Despite recent advances in treatment, long-term survival remains low. Treatment of newly diagnosed and recurrent HGG consists of a combination of surgery, radiation therapy, and chemotherapy and, more recently, biologic and immunotherapies, some of which are under clinical investigation. Unique imaging phenomena characterize therapeutic responses to these therapies, challenging conventional radiologic interpretation.

MRI is the method of choice for the initial diagnosis, follow-up, and therapy response assessment of brain tumors. HGG are typically characterized by contrast-enhanced areas on T1-weighted imaging and hyperintensity on FLAIR and T2WI. The current standard for the radiologist's assessment in
patients with HGG relies on the Response Assessment in Neuro-Oncology (RANO) criteria, which expand upon the earlier Macdonald criteria, to incorporate the nonenhancing component of the tumor, as this component may indicate infiltrative or diffuse tumor growth. However, the current RANO criteria, which rely on conventional imaging, may fail to distinguish different tissue components that have similar imaging patterns (ie, contrast-enhanced T1WI) yet different prognostic implications (tumor-versus-treatment effects), the identification of which is crucial for appropriate clinical management.

Glioblastoma shows high heterogeneity within the lesion area, including areas of active tumor, necrosis, infiltrative tumor cells, and vasogenic edema. Enhancing lesion area can be measured on conventional T1WI; however, it may represent a variety of both tumor and non tumor processes, including tumor progression, pseudoprogression, postsurgical changes, and radiation effects. Differentiation of the nonenhancing area between infiltrative tumor and nontumor area (edema and/or gliosis) is challenging, as both are characterized by hyperintense FLAIR/T2WI signal. Hyperintense FLAIR/T2WI signal can be expected to represent pure vasogenic edema in metastatic tumors (characteristically with well-defined borders between malignant cells and normal brain), while in HGG, the hyperintensity may reflect areas of infiltrating tumor cells and/or vasogenic edema.

Glioblastoma is characterized by a highly complex neovascularization process. The increased vascularity of GB allows analysis and characterization of the tumor using advanced vascular MRI methods, including perfusion and permeability. Dynamic contrast-enhanced (DCE) imaging has been widely used for characterization of tumor biology and therapy-response assessment. DCE relies on dynamic T1WI sequences during bolus injection of contrast agent. Hemodynamic parameters, including plasma volume \( (v_p) \) and tissue transfer constants (specifically, volume transfer constant \( [K_{trans}] \) and interstitium-to-plasma rate constant \( [k_{ep}] \)), can be extracted from DCE data.

Several studies have suggested the utility of DCE parameters as important imaging markers for brain tumor diagnosis, grading, and therapy-response monitoring; to differentiate between active and nonactive tumor components or infiltrative and noninfiltrative tumor; and for early-stage assessment of antiangiogenic therapy response in HGG. However, none of these studies have aimed to classify the entire lesion area, both enhancing and nonenhancing, into tumor and nontumor areas at the voxel level.

Various techniques have been proposed for segmentation and classification of brain lesions. Support vector machine (SVM) is a supervised binary classifier, originally aimed at classifying data into 2 classes. Based on a training dataset, the algorithm finds the hyperplane that maximally separates between points. New data are then mapped onto the same space and assigned to a particular class. SVM has been widely used as a classifier of medical images and has proved advantageous over other algorithms of its kind.

The aim of this study was to classify lesion areas in patients with HGG using SVM into 4 distinct components: 1) enhancing tumor, 2) enhancing nontumor, 3) nonenhancing tumor, and 4) nonenhancing nontumor. Segmentation of lesion areas into tumor and nontumor components refines the RANO criteria and may improve therapy response MRI assessment.

**MATERIALS AND METHODS**

**Patients**
A total of 140 MRI scans were obtained from 38 patients, who were scanned every 2–3 months. Ninety-five scans were from 26 patients with GB; 31 scans (5 patients) were of anaplastic astrocytomas; 3 scans (1 patient) were of anaplastic oligodendrogliomas; and 11 scans (6 patients) were of brain metastases (4 breast and 2 lung cancers) (Table 1). Inclusion criteria required a normal glomerular filtration rate and no contraindication to MRI. All patients with HGG who were scanned longitudinally \((n = 26)\) had recurrent HGG and had been treated initially with standard chemoradiation. Additional therapies are detailed in the On-line Appendix. Nine patients received bevacizumab at recurrence and had pre- and post bevacizumab MRIs.

The study was approved by the Tel Aviv Sourasky Medical Center review board, and written informed consent was obtained from all patients.

**MR Imaging Protocol**

Scans were performed on 3T MRI scanners: 91 scans were performed on a Signa Excite scanner (GE Healthcare) using an 8-channel head coil, and 49 scans were performed on a Magnetom Prisma scanner (Siemens) using a 20-channel head coil. The protocol included conventional imaging: spoiled gradient-recalled (SPGR)/FLASH T1WI performed before and after contrast agent injection and FLAIR. DCE data were acquired using multiphase 3D T1WI SPGR/FLASH before, during, and after contrast agent injection (FOV 250 mm; matrix 256X256/256X184; section thickness 5 mm; TR/TE = 5/2.2 ms; and flip angle = 20°). For the T1 maps, variable flip angle SPGR/FLASH data were acquired with nominal flip angles of 5°, 10°, 15°, 20°, and 30°. Dynamic data were acquired with a temporal resolution of 6 seconds and scan duration of 6 minutes. A power injector was used to infuse a single dose \((0.2\ mL/kg)\) of contrast agent (gadoterate meglumine, 0.5 mol/L, Dotarem; Guerbet, Aulnay-sous-Bois, France) followed by 20 mL saline flush, both at a constant rate of 5 mL/s second. Fourteen-to-twenty sections were centered on the tumor area as identified in the conventional images, providing brain coverage of 70–100 mm. MR spectroscopy data were acquired with a single-voxel point-resolved sequence by using a cubic volume of \(\approx 8\ mL\) located at the lesion area and normal-appearing white matter (NAWM) in the contralateral hemisphere (TR/TE = 1500/35 ms).

<table>
<thead>
<tr>
<th>Table 1: Patient characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosis</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Glioblastoma</td>
</tr>
<tr>
<td>Anaplastic astrocytoma</td>
</tr>
<tr>
<td>Anaplastic oligodendroglioma</td>
</tr>
<tr>
<td>Brain metastasis</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

**Lesion Classification**

Lesion classification involved several steps as detailed in the Online Appendix and shown in On-line Fig 1. Following preprocessing, the enhancing and nonenhancing lesion areas, designated as targets for classification, were automatically segmented from the T1WI and FLAIR (respectively) images by using an adaptive threshold as previously described. Manual correction was performed when necessary. Normalized images of FLAIR (nFLAIR) and post-contrast T1WI (nT1WI+Gd) were calculated relative to NAWM. The DCE pharmacokinetic parameters, $v_p$, $K_{trans}$, interstitial volume ($v_i$), $k_{ep}$, and bolus arrival time (BAT), were estimated by using DCE-Up-Sampled-Temporal-Resolution (DUSTER).

Classification of the lesion area into 4 components was performed automatically by using the SVM. For the training data, volumes of interest of 5–10 voxels each (30–60 cc) were manually defined with the FMRIB Software Library (FSL; http://www.fmrib.ox.ac.uk/fsl) in areas with known pathology for the 4 tissue types: 1) enhancing tumor, defined retrospectively in patients with progressive HGG based on the RANO assessment on follow-up scans; 2) enhancing nontumor, defined retrospectively in patients with nonprogressive HGG, with treatment-related changes based on stable longitudinal (>6-month) RANO assessments; 3) nonenhancing tumor (infiltrative), defined as a peritumoral area (<2 cm from the enhanced tumor margins) in a nonenhancing FLAIR hyperintense region in patients with HGG with progressive disease; and 4) nonenhancing, nontumor (edema), defined in patients with brain metastasis, in whom the hyperintense FLAIR signal represented pure vasogenic edema, without tumor cell infiltration.

Seven MRI parameters were measured within and compared between tissue types: nT1WI+Gd, nFLAIR, $v_p$, $K_{trans}$, $v_i$, $k_{ep}$, and BAT. SVM classifier with a linear kernel was trained only on the MRI parameters that were found to be significantly different ($P < .05$) among the tissue types. Voxelwise classification using SVM was performed separately for the enhancing and nonenhancing lesion areas in all patients.

**Evaluation**

Two-Fold Cross-Validation Analysis. To test the sensitivity and specificity of the method with validating confidence intervals, we performed 300 iterations of the training and testing data based on different random selections of the even and odd partitions. The group was divided into 2 subgroups (randomized, even-versus-odd) with 1 subgroup as training data, the other as the test data, and vice versa.

**MR Spectroscopy.** Classification results were validated based on MR spectroscopy results, performed off-line using LCModel (http://www.lcmodel.com). Metabolite ratios of glycerophosphocholine + phosphocholine (GPC+PCho) and of N-acetyl aspartate + N-acetyl aspartylglutamate (NAA+NAAG) relative to creatine + phosphocreatine (Cr+Pcr) were calculated. While many MR spectra were acquired for clinical purposes, MR spectroscopy data were used in this study when voxel location included only 1 tissue component or were defined within the NAWM.

**Longitudinal Assessment.** Longitudinal assessment of the classification results was performed in patients with primary brain tumors (32 patients). Longitudinal changes in the volume of each component were evaluated relative to RANO criteria assessed by a senior neuroradiologist defining 4 categories: complete response, partial response, stable disease, or progression.

**Statistical Analysis**

To identify differences between tumor and nontumor components for the various MRI parameters (separately for the enhancing and nonenhancing area), we used the Wilcoxon test (because parameters were not normally distributed).

**RESULTS**

**Classification Results**

**Training Data.** We defined VOIs for the 4 tissue types: the enhancing tumor component in 28 scans from 5 patients with progressive HGG (4 GB and 1 anaplastic astrocytoma); the enhancing nontumor component in 10 scans from 3 patients with high-grade tumors (2 GB and 1 oligodendroglioma); the nonenhancing tumor component in 24 scans from 5 patients with GB; and the nonenhancing, nontumor component in 11 scans from 6 patients with brain metastasis.

Mean values of the 7 MRI parameters, nT1WI+Gd, nFLAIR, $v_p$, $K_{trans}$, $v_i$, $k_{ep}$, and BAT, measured in the 4 components (tissue types) defined for the training data and NAWM, are provided in Table 2. For the enhancing components, significant differences were detected for the nT1WI+Gd, $v_p$, $K_{trans}$, $k_{ep}$, and $v_i$ parameters, demonstrating higher vascularity ($v_p$) and impaired permeability (nT1WI+Gd, $K_{trans}$) in the tumor component relative to...
the nontumor component. For the nonenhancing components, significant differences between components were detected for all parameters, demonstrating higher vascularity (v_p) and slightly impaired permeability (nT1WI+Gd, K^trans) in the tumor component relative to the nontumor component. The v_e and k_ep parameters showed significant group differences yet a substantially high SD (>150%) and were therefore not used in the training set.

**Classification of Lesion Area**

The SVM classifier was trained on the nT1WI+Gd, v_p, and K^trans parameters for enhancing components and on the nT1WI+Gd, nFLAIR, v_p, K^trans, and BAT parameters for nonenhancing components. Voxelwise classification was performed on all scans of all patients (n = 140), separately for enhancing and nonenhancing lesion areas. Figure 1 shows classification results obtained in 3 patients. In the first patient (Fig 1, case 1) with breast cancer metastasis, the non-enhancing lesion area was classified as nonenhancing, nontumor (ie, edema) and the enhancing lesion area was classified as enhancing tumor. The second patient (Fig 1, case 2) with GB was an example in whom most of the nonenhancing lesion was classified as tumor component. This patient was diagnosed with progressive disease on a subsequent follow-up scan (2 months later). In the third patient (Fig 1, case 3), also with GB, the enhancing lesion and most of the non-enhancing area were classified as tumor component, consistent with a diagnosis of progressive disease as diagnosed on subsequent follow-up.

**Longitudinal Assessment**

Longitudinal assessment of the volume of each component and the correlating radiologic assessment based on RANO, for 26 patients with high-grade lesions scanned longitudinally >3 times (a total of 118 scans) are shown in On-line Fig 2. Consistent results were obtained between changes in the volume of the lesion components and the radiologist’s assessment in 27 patients (of 32, 84%). However, in 5 (16%; patients 1, 17, 18, 20, and 26), increased volume of the nonenhancing tumor component indicating tumor progression preceded the radiologic diagnosis based on RANO criteria by several months.

Figure 2 shows longitudinal data obtained from a 54-year-old patient with anaplastic astrocytoma (patient 18). While the conventional imaging showed a pattern of stability at scans 3 and 5, with reduction/no substantial changes in the enhancing and non-enhancing lesion areas, classification results revealed a different clinical scenario—progressive tumor growth. A continual increase in the nonenhancing tumor component indicated a pattern of infiltrative disease progression through scans 2–5 (particularly from scan 2 to 3). This patient was diagnosed with progressive disease at scan 4 and died 71 days following scan 5.

In 7 of 9 patients who received bevacizumab, substantial reductions (mean, 56%) were detected mainly in the volume of the nonenhancing, nontumor component (interpreted as edema). Three patients of 7 (patients 12,13, and 23), though showing a major reduction in the volume of the nonenhancing, nontumor component, had an increase in the nonenhancing tumor component, suggesting a shift to an infiltrative pattern of tumor progression during bevacizumab treatment.

**Evaluation**

**Sensitivity and Specificity of the Training Data.** Analysis demonstrated 100% sensitivity (85%/100% for the 5th/90th percentiles) and 100% specificity (100%/100%) for the identification of enhancing tumor and enhancing nontumor components, and 100% sensitivity (100%/100%) and 100% specificity (100%/100%) for identification of the nonenhancing (infiltrative) tumor and nonenhancing nontumor components.

**MR Spectroscopy Results.** Forty-one spectra were obtained with good quality,
with voxel location consisting of 1 of the 4 components or within NAWM: 14 from NAWM; 6 from nonenhancing, nontumor component; 1 from nonenhancing tumor; and 19 from enhancing tumor components. MR spectroscopy results are provided in Table 2. Spectra from the enhancing, nontumor components showed primarily lipid peaks, without a detectable metabolite ratio. MR spectroscopy supported the classification results, demonstrating a >1.5-fold increase in glycerolphosphocholine + phosphocholine/creatine + phosphocholine (GPC+Pch/Cr+PCr) ratio in the tumor components relative to the nontumor component and NAWM.

**DISCUSSION**

In this study, longitudinal classification of lesion area into tumor and nontumor components was performed in patients with HGG based on multi-MRI parameters extracted from conventional imaging and DCE MRI. RANO criteria improved the earlier MacDonald criteria by incorporating the nonenhancing component of the lesion. Our proposed method, referred to as segmented RANO (sRANO), provides a logical next step in the evolution of MR clinical imaging, to better assess tumor growth and therapy response. Both enhancing and nonenhancing lesion areas may include tumor and nontumor components. The proposed sRANO classifies the lesion areas and defines each component separately. In 16% of our cases followed longitudinally, we were able to identify tumor progression several months in advance of RANO criteria.

Conventional MRI lacks the precision to reliably differentiate tumor and nontumor areas with similar imaging patterns. However, the implications of such differentiation regarding treatment are substantial. Within the enhancing lesion area, recognition of dynamic, active tumor typically requires a change in therapy. Likewise, the identification of radiation necrosis or treatment-related changes can prevent unnecessary interventions. While salvage re-irradiation of brain tumors can typically be performed with a minimum of resultant treatment necrosis, re-irradiation to an area of preexisting treatment-related necrosis could exacerbate existing tissue damage and lead to potentially life-threatening neurologic toxicity.

Classification between tumor and nontumor areas within the nonenhancing lesion may provide new guidelines for radiation. Standards of treatment-planning techniques differ geographically. Accurate delineation of tumor area may either limit the size of the radiation field and decrease radiation exposure to normal brain tissue or define more accurately the optimal radiation field to include the entire tumor area and enable more efficient treatment.

In addition, as the use of immune-mediating therapies in glioma increases, the complexity of interpreting inflammatory responses to treatment (“pseudoprogression”) in enhancing and nonenhancing lesion areas will also increase. Misinterpretation of the response to a potentially helpful therapy could lead to premature cessation of a useful therapy and compromise patient outcome.

Within the enhancing lesion area, all measured vascular parameters (except the BAT) clearly differentiated tumor and nontumor components. While both components demonstrated increased tissue permeability, the tumor component was characterized by much higher permeability and \( v_p \) compared with the nontumor component and relative to NAWM. GB is characterized by a complex neovascularization process that results in formation of new, abnormal blood vessels and an impaired blood-brain barrier, and thus manifests as a hyperperfusion area with increased blood flow, volume, and permeability, as was seen in our cohort study. The nontumor component demonstrated reduced \( v_p \) compared with NAWM, consistent with data that have shown areas of treatment-related changes characterized by impaired blood brain barrier yet with reduced blood flow and volume, due to treatment-induced vascular endothelial damage and coagulative necrosis.

Within the nonenhancing lesion area, despite a similar appearance between tumor and nontumor components on conventional imaging (FLAIR and T2WI), significant differences were found in all measured MRI parameters. The nonenhancing, nontumor component was characterized by reduced \( v_p \), in comparison with the NAWM and the nonenhancing tumor component. This pattern is consistent with previous reports of tissues proximal to brain tumors in patients and animal models and can be explained by compression of regional capillaries caused by vasogenic edema. The tumor component in our cohort was characterized by increased tissue permeability and increased perfusion as demonstrated by \( K_{trans} \) and \( v_p \) parameters, consistent with findings in previous studies.

In this work, classification of lesion area was performed with a linear SVM classifier, a simple form of SVM, highly suitable for a limited training set size, for which the separating hyperplane is simply a plane in the feature space and the relative contribution of the features can be easily obtained. Various machine learning– and computational intelligence–based methods have been proposed for segmentation of brain tumors. In this study, SVM was found to be a robust, rapid method, tailored to the clinical data, easily implemented, and most important, it differentiated lesion components with high sensitivity and specificity.

The success of a supervised algorithm is directly determined by the selected training set. Thus, only clear-cut cases were used for this study while training the classifier, based on retrospective radiologic assessment for several months. Classification into tumor and nontumor components was performed based on conventional parameters (nT1WI+Gd, and nFLAIR) and parameters extracted from DCE (\( v_p \), \( K_{trans} \), and BAT). While perfusion parameters can be obtained using dynamic susceptibility contrast imaging and have been shown to differentiate between tumor and nontumor components, DCE imaging is preferable due to its higher spatial resolution, less sensitivity to susceptibility artifacts, and provision of quantitative parameters, including permeability (\( K_{trans} \)).

Several issues are important when interpreting the clinical relevance of classification results. The relatively small sample size and absence of histology-proved diagnosis in our cohort call for additional studies. Tumor components were found to be highly vascular while the nontumor components were found to have reduced vascularity. However, the enhancing, nontumor component may represent a more complex situation. It may reflect treatment-related changes (such as radiation necrosis) but may also represent tumor-associated hy-
voxic necrosis, which is one of the histologic hallmarks of GB.37 Thus, in such cases, it is important to consider the non-tumor component in the context of active tumor surrounding the necrotic area.

CONCLUSIONS
The current study proposes a model of segmented RANO criteria, sRANO, that classifies tumor and non-tumor components within a lesion area, with high sensitivity and specificity. Longitudinal assessment in patients demonstrated consistency between the classification results and radiologist’s assessment in most cases; in 16% of patients, the segmentation results identified growth of highly vascular components and preceded the conventional radiologic diagnosis of tumor progression by several months. The proposed sRANO method and results presented in this study demonstrate the importance and contribution of segmentation of the enhancing and nonenhancing lesion areas into tumor and nontumor components, to improve therapy-response assessment of patients with malignant brain tumors.

ACKNOWLEDGMENTS
We are grateful to Vicki Myers for editorial assistance, to Faina Vitenstein and Tuvia Ganot for their assistance in patient recruitment and performing MRI scans, and to Dr Ricardo Tarrasch for statistical consulting.

Disclosures: Deborah T. Blumenthal—UNRELATED: Board Membership: Medical Advisory board for Vascular Biogenics. Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed: Roche, Bristol-Myers Squibb. Comments: travel grants to international meetings.* Money paid to the institution.

REFERENCES
27. Cloughesy TF, Cavenee WK, Mischel PS. Glioblastoma: from molec-
33. Senturk S, Oğuz KK, Cila A. Dynamic contrast-enhanced suscepti-

ABSTRACT

BACKGROUND AND PURPOSE: Among patients with vertebrobasilar dolichoectasia is a subset of patients with disease affecting the anterior circulation as well. We hypothesized that multivessel intracranial dolichoectasia may represent a distinct phenotype from single-territory vertebrobasilar dolichoectasia. The purpose of this study was to characterize clinical characteristics and angiographic features of this proposed distinct phenotype termed “diffuse intracranial dolichoectasia” and compare them with those in patients with isolated vertebrobasilar dolichoectasia.

MATERIALS AND METHODS: We retrospectively reviewed a consecutive series of patients with diffuse intracranial dolichoectasia and compared their demographics, vascular risk factors, additional aneurysm prevalence, and clinical outcomes with a group of patients with vertebrobasilar dolichoectasia. “Diffuse intracranial dolichoectasia” was defined as aneurysmal dilation of entire vascular segments involving ≥2 intracranial vascular beds. Categoric and continuous variables were compared by using χ² and Student t tests, respectively.

RESULTS: Twenty-five patients had diffuse intracranial dolichoectasia, and 139 had vertebrobasilar dolichoectasia. Patients with diffuse intracranial dolichoectasia were older than those with vertebrobasilar dolichoectasia (70.9 ± 14.2 years versus 60.4 ± 12.5 years, P = .0002) and had a higher prevalence of abdominal aortic aneurysms (62.5% versus 14.3%, P = .01), other visceral aneurysms (25.0% versus 0%, P < .0001), and smoking (68.0% versus 15.9%, P < .0001). Patients with diffuse intracranial dolichoectasia were more likely to have aneurysm growth (46.2% versus 21.5%, P = .09) and rupture (20% versus 3.5%, P = .007) at follow-up. Patients with diffuse intracranial dolichoectasia were less likely to have good neurologic function at follow-up (24.0% versus 57.6%, P = .004) and were more likely to have aneurysm-related death (24.0% versus 7.2%, P = .02).

CONCLUSIONS: The natural history of patients with diffuse intracranial dolichoectasia is significantly worse than that in those with isolated vertebrobasilar dolichoectasia. Many patients with diffuse intracranial dolichoectasia had additional saccular and abdominal aortic aneurysms. These findings suggest that diffuse intracranial dolichoectasia may be a distinct vascular phenotype secondary to a systemic arteriopathy affecting multiple vascular beds.

Fusiform and dolichoectatic intracranial aneurysms are generally associated with a poor natural history with high rates of rupture, mass effect, and ischemic stroke.¹⁻⁴ These aneurysms often form secondary to circumferential loss or weakening of the internal elastic lamina of a vessel segment and can involve an entire vessel or may be limited to a single vessel segment (eg, basilar artery, cavernous or supraclinoid ICA, proximal MCA, and so forth).⁵⁻⁸ Notably, most prior reports of intracranial dolichoectasia have focused on involvement of single-vessel territories, most commonly the vertebrobasilar system. These dolichoectatic aneurysms are associated with a wide range of proposed, underlying etiologies, including healed, focal arterial dissections; chronic hypertension; prior radiation therapy; connective tissue diseases; arterial tortuosity syndrome; glycogen storage diseases; infection; and myxomatous emboli.⁵⁻⁸

During the past several years, we have noticed in our clinical practice a subset of patients who have segmental aneurysmal dolichoectasia of multiple intracranial vascular territories in the absence of a known, underlying insult such as radiation, infection, genetic connective tissue disease, or dissection. We hypothesized that multivessel presentation of intracranial dolichoectasia may represent a process distinct from the frequently described single-territory dolichoectasia, potentially with distinct presentation, associations, and natural history. For this report, we have defined “diffuse intracranial dolichoectasia” as aneurysmal dilation of entire segments that involve ≥2 in-
tracranial vascular beds. These criteria were chosen due to the notion that fusiform aneurysmal dilation of multiple intracranial vascular beds indicates a systemic vascular insult, while fusiform aneurysmal dilation of a single bed suggests a more localized pathologic process.

The primary purpose of this article was to fully characterize the clinical risk factors and anatomic and angiographic characteristics of this proposed distinct entity, which we term “diffuse intracranial dolichoectasia.” The secondary purpose was to compare clinical risk factors and anatomic and angiographic characteristics of patients with diffuse intracranial dolichoectasia with those in a group of patients with dolichoectasia isolated to the vertebrobasilar system.

**MATERIALS AND METHODS**

**Patient Population**

Following institutional review board approval, we identified all patients with diffuse intracranial dolichoectasia evaluated at our institution from January 1, 2001, to December 31, 2014. We also identified a control group of patients with isolated vertebrobasilar dolichoectasia. To identify patients, we searched our imaging database for reports of head MR imaging, MRA, CTA, and DSA examinations in which any one of the following terms was used in the body or conclusions of the report: “fusiform,” “aneurysmal dilation,” “dilation,” “fusiform aneurysm,” “dolichoectasia,” “dolichoectatic,” and “dolicho.” All identified imaging examinations were then reviewed by 2 radiologists in consensus. Cases with fusiform aneurysmal dilation of multiple vascular segments were then identified for additional review to determine whether they met the diagnostic criteria for diffuse intracranial dolichoectasia. Cases in which there was fusiform aneurysmal dilation with or without tortuosity isolated to the vertebrobasilar system were defined as “isolated vertebrobasilar dolichoectasia.” Of the cases that were identified for additional review as potentially representing diffuse intracranial dolichoectasia, maximum measurements were obtained of the basilar artery, supraclinoid ICAs, and bilateral M1 segments.

**Diagnostic Criteria**

To be included in this diffuse intracranial dolichoectasia group, patients had to meet the following diagnostic inclusion criteria: adult patients with fusiform aneurysmal dilation of entire vascular segments (ie, supraclinoid ICA, basilar artery, M1 segment of the MCA) that involved ≥2 intracranial vascular beds (ie, vertebrobasilar system, left anterior circulation, or right anterior circulation). Aneurysmal dilation of a vascular segment was defined as a diameter of at least 2 times the normal diameter of a given vascular segment as defined by an angiographic atlas. Isolated vertebrobasilar dolichoectasia was defined as an aneurysmal dilation 2 times the normal diameter without a definable neck involving a portion of an arterial segment (either vertebral or basilar) with any degree of tortuosity. The criteria for aneurysmal dilation of a vessel segment are summarized in Table 1. Exclusion criteria included the following: 1) pediatric patients, 2) patients with underlying connective tissue diseases (ie, Ehlers-Danlos and Marfan syndromes, Loeys-Dietz syndrome, autosomal dominant polycystic kidney disease, neurofibromatosis type 1, and so forth), 3) patients with a history of cranial radiation, 4) patients with an underlying infectious etiology (ie, varicella zoster vasculopathy, HIV, and so forth), and 5) patients with poststenotic dilations secondary to atherosclerosis or dissection. All patients in the vertebrobasilar dolichoectasia control group and 11 patients in the diffuse intracranial dolichoectasia group were also included in a prior study on the imaging natural history of patients with vertebrobasilar, nonsaccular dolichoectatic aneurysms.

**Baseline Demographics and Clinical Risk Factors**

We collected the following baseline patient demographic characteristics and comorbidities: age; sex; the presence of an ectatic abdominal aorta (maximum diameter, 2.6–2.9 cm) or an abdominal aortic aneurysm (maximum diameter, ≥3 cm); and the presence of other vascular ectasias/aneurysms, fibromuscular dysplasia, coronary artery disease, peripheral artery disease, hypertension, diabetes mellitus, hyperlipidemia, tobacco use (current, former, never), pack-years of tobacco use (if available), alcohol abuse, sympathomimetic abuse, and family history of abdominal aortic aneurysm, stroke, or aneurysm. In addition to baseline demographic and clinical risk factors, we collected data on presenting symptoms at the time of the discovery of the intracranial dolichoectasia (ie, incidental, stroke, subarachnoid hemorrhage, neuralgia, mass effect, headache, hemifacial spasm, other). Baseline demographic and clinical risk factors were compared between the diffuse intracranial dolichoectasia group and the isolated vertebrobasilar dolichoectasia group.

**Imaging Characteristics**

Among patients who met the diagnostic criteria for diffuse intracranial dolichoectasia and those in the isolated vertebrobasilar dolichoectasia control group, we collected the following information from their imaging examinations: the presence of a saccular aneurysm; the presence of other intracranial vascular dilations, subarachnoid hemorrhage, or intraparenchymal hemorrhage on presentation; white matter ischemic disease; and thrombus within the aneurysm. All measurements on CTA and MRA were obtained by using multiplanar reformatting software (Terarecon, San Mateo, California). The maximum cross-sectional diameter of the vessel perpendicular to the longitudinal axis of the vessel was obtained. For DSA images, the maximum diameter of the vessel perpendicular to the longitudinal axis of the vessel was obtained. Imaging characteristics were compared between the diffuse intracranial dolichoectasia group and the isolated vertebrobasilar dolichoectasia group.

<table>
<thead>
<tr>
<th>Vascular Segment</th>
<th>Definition of Vascular Segment</th>
<th>Definition of Aneurysmal Dilation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cavernous ICA</td>
<td>Entry point of ICA into the cavernous segment to the ophthalmic artery origin</td>
<td>≥8.5 mm</td>
</tr>
<tr>
<td>Supraclinoid ICA</td>
<td>ICA from the ophthalmic artery origin to the ICA bifurcation</td>
<td>≥8.0 mm</td>
</tr>
<tr>
<td>M1 segment of MCA</td>
<td>Origin of MCA to the MCA bifurcation</td>
<td>≥5.0 mm</td>
</tr>
<tr>
<td>Basilar artery</td>
<td>Confluence of vertebral arteries to the basilar bifurcation</td>
<td>≥6.0 mm</td>
</tr>
</tbody>
</table>

Table 1: Definition of vascular segments and aneurysmal dilation
Outcomes Studied
In cases in which there was serial imaging, we collected data on presence of dolichoectatic aneurysm growth or rupture and new ischemic stroke. We also collected data on the clinical status of the patient at last follow-up. Clinical status was defined as no or minimal disability (mRS ≤ 2), moderate or severe disability (mRS 3–5), and death. Among patients who died, the cause of death was noted, when available. Outcomes were compared between the diffuse intracranial dolichoectasia group and the isolated vertebrobasilar dolichoectasia group.

Statistical Analysis
All categoric variables are reported as number (percentage), while continuous variables are reported as mean ± SD. Categoric variables were compared by using χ² tests, and continuous variables were compared by using a Student t test. To determine whether a diagnosis of diffuse intracranial dolichoectasia was independently associated with aneurysm growth, rupture, and new infarct, we performed a multivariate logistic regression analysis, adjusting for baseline factors that were different between groups (ie, age and smoking history). All analysis was performed by using the SAS-based statistical software package JMP 12.0 (SAS Institute, Cary, North Carolina).

RESULTS
Baseline Patient Characteristics for Diffuse Intracranial Dolichoectasia
The search of radiology reports (ie, MR imaging, MRA, CTA, and DSA) yielded 890 unique patients with the terms “fusiform,” “aneurysmal dilation,” “dilation,” “fusiform aneurysm,” “dolichoectasia,” “dolichoectatic,” and “dolicho” in their reports. After further review of imaging reports and radiologic images, 25 patients met the diagnostic criteria for diffuse intracranial dolichoectasia.

The mean patient age was 70.9 ± 14.2 years. Two patients were younger than 50 years of age (8.0%), and 21 patients (84.0%) were 60 years of age or older. Eighty-four percent (21/25) of patients were men. The most common clinical presentation was acute ischemic stroke or transient ischemic attack (13/25, 52%). Of the ischemic strokes, 6 affected the posterior circulation and 7 affected the anterior circulation. Five patients (20%) had symptoms due to mass effect from the aneurysmal dilations, 3 patients (12%) had dizziness, and in the remaining 4 patients (16%), the aneurysms were discovered incidentally during evaluation for loss of consciousness, syncope, or visual blurring.

Among patients with abdominal imaging, 50% (8/16) had an abdominal aortic aneurysm and an additional 13% (2/16) had aortic ectasia. Hypertension was present in 88% (22/25) of patients, and 68% (17/25) were current or former smokers. Hyperlipidemia was present in 52% (13/25) of patients, and intracranial atherosclerosis was present in 48% (12/25). These data are summarized in Table 2.

Angiographic and Anatomic Characteristics for Diffuse Intracranial Dolichoectasia
The most common distribution of disease was dilation of the basilar artery and 1 internal carotid artery (10 patients, 40%). Five patients (20%) had diffuse dilation of the basilar artery and bilateral ICAs, and 4 patients (16%) had dilation of the bilateral ICAs without dilation of the basilar artery. Two patients (8.0%) had dilation of the basilar artery and bilateral ICAs and MCAs.

The mean maximum diameter of the dilated basilar arteries was 15.2 ± 8.3 mm. The mean maximum diameter of the dilated

Table 2: Summary of patient characteristics and outcomes

<table>
<thead>
<tr>
<th>Variable</th>
<th>DID</th>
<th>VBD</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>25</td>
<td>139</td>
<td>~</td>
</tr>
<tr>
<td>Mean age (SD) (yr)</td>
<td>70.9 (14.2)</td>
<td>60.4 (12.5)</td>
<td>.0002</td>
</tr>
<tr>
<td>Age group (yr)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20–29</td>
<td>1 (4.0)</td>
<td>5 (3.6)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>30–39</td>
<td>0 (0.0)</td>
<td>6 (4.3)</td>
<td></td>
</tr>
<tr>
<td>40–49</td>
<td>1 (4.0)</td>
<td>6 (4.3)</td>
<td></td>
</tr>
<tr>
<td>50–59</td>
<td>2 (8.0)</td>
<td>44 (31.6)</td>
<td></td>
</tr>
<tr>
<td>60–69</td>
<td>7 (28.0)</td>
<td>43 (30.9)</td>
<td></td>
</tr>
<tr>
<td>70–79</td>
<td>6 (24.0)</td>
<td>30 (21.5)</td>
<td></td>
</tr>
<tr>
<td>≥80</td>
<td>8 (32.0)</td>
<td>5 (3.6)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>21 (84.0)</td>
<td>105 (75.5)</td>
<td>.51</td>
</tr>
<tr>
<td>Female</td>
<td>4 (16.0)</td>
<td>34 (24.5)</td>
<td></td>
</tr>
<tr>
<td>Indication for imaging</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroke</td>
<td>13 (52.0)</td>
<td>45 (32.7)</td>
<td>.10</td>
</tr>
<tr>
<td>Mass effect</td>
<td>5 (20.0)</td>
<td>25 (18.0)</td>
<td>.81</td>
</tr>
<tr>
<td>Other/incidental</td>
<td>7 (28.0)</td>
<td>69 (49.6)</td>
<td>.08</td>
</tr>
<tr>
<td>Comorbidities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominal aortic aneurysm or ectasia</td>
<td>10 (62.5)</td>
<td>12 (43.3)</td>
<td>.01</td>
</tr>
<tr>
<td>Other visceral aneurysms</td>
<td>4 (25.0)</td>
<td>0 (0.0)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Fibromuscular dysplasia</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>1.0</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>10 (40.0)</td>
<td>40 (29.0)</td>
<td>.38</td>
</tr>
<tr>
<td>Peripheral artery disease</td>
<td>4 (16.0)</td>
<td>6 (4.3)</td>
<td>.07</td>
</tr>
<tr>
<td>Hypertension</td>
<td>22 (88.0)</td>
<td>97 (70.3)</td>
<td>.11</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>3 (12.0)</td>
<td>22 (15.9)</td>
<td>.85</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>13 (52.0)</td>
<td>73 (52.9)</td>
<td>.96</td>
</tr>
<tr>
<td>Smoking</td>
<td>17 (68.0)</td>
<td>22 (15.9)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Alcohol abuse</td>
<td>1 (4.0)</td>
<td>7 (5.0)</td>
<td>.82</td>
</tr>
<tr>
<td>Sympathomimetic abuse</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>1.0</td>
</tr>
<tr>
<td>Family history</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominal aortic aneurysm</td>
<td>3 (12.0)</td>
<td>9 (6.4)</td>
<td>.58</td>
</tr>
<tr>
<td>Ischemic stroke</td>
<td>7 (28.0)</td>
<td>32 (23.0)</td>
<td>.78</td>
</tr>
<tr>
<td>Intracranial aneurysm</td>
<td>1 (4.0)</td>
<td>9 (6.4)</td>
<td>.98</td>
</tr>
<tr>
<td>Other saccular aneurysms present</td>
<td>7 (28.0)</td>
<td>22 (15.8)</td>
<td>.24</td>
</tr>
<tr>
<td>Growth on follow-up</td>
<td>6 (46.2)</td>
<td>30 (21.6)</td>
<td>.09</td>
</tr>
<tr>
<td>SAH on follow-up</td>
<td>5 (20.0)</td>
<td>5 (3.5)</td>
<td>.007</td>
</tr>
<tr>
<td>Infarct on follow-up</td>
<td>7 (28.0)</td>
<td>18 (13.0)</td>
<td>.10</td>
</tr>
<tr>
<td>Clinical status at last follow-up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good neurologic function</td>
<td>6 (24.0)</td>
<td>80 (57.6)</td>
<td>.004</td>
</tr>
<tr>
<td>Poor neurologic function</td>
<td>6 (24.0)</td>
<td>22 (15.8)</td>
<td>.48</td>
</tr>
<tr>
<td>Death</td>
<td>13 (52.0)</td>
<td>37 (26.6)</td>
<td>.02</td>
</tr>
<tr>
<td>Aneurysm-related death</td>
<td>6 (24.0)</td>
<td>10 (7.2)</td>
<td>.02</td>
</tr>
</tbody>
</table>

Note: —VBD indicates vertebrobasilar dolichoectasia; DID, diffuse intracranial dolichoectasia.
These data are summarized in Table 3.

Outcomes for Diffuse Intracranial Dolichoectasia

Thirteen patients had imaging follow-up (mean, 54.1 ± 78.4 months). Of these patients, 6 (46%) had growth of at least 1 aneurysmal dolichoectasia, which occurred at a mean of 56.0 ± 54.3 months. Given a total of 63.1 patient-years of imaging follow-up, the overall growth rate was 9.5%/patient-year. Seven patients (56%) had imaging evidence of a new infarct after a mean of 60.9 ± 66.5 months of follow-up, including 2 patients with brainstem perforator infarcts, 5 patients with cerebellar or thalamic infarcts, and 4 patients with anterior circulation infarcts.

Five patients (5/25, 20.0%) had imaging evidence of subarachnoid hemorrhage at a mean of 59.0 ± 73.1 months of follow-up. Of these 5 patients, 3 had rupture of 2 basilar dolichoectatic/fusiform aneurysms, 1 had rupture of a posterior inferior cerebellar artery dissecting aneurysm, and in 1 patient, the source of the subarachnoid hemorrhage was unclear (On-line Table). Including patients without imaging follow-up, the overall rate of subarachnoid hemorrhage was 5.9%/year (5 cases over 84.9 patient-years).

The mean clinical follow-up for all 25 patients was 40.8 ± 62.3 months. At last follow-up, 13 patients (52%) had died, 6 patients (24%) had poor neurologic outcome, and 6 patients (24%) had good neurologic outcome. Among the patients who died, 4 died from aneurysm rupture, 2 died from hydrocephalus secondary to mass effect from a growing basilar artery fusiform aneurysm, and 1 patient died from intraparenchymal hemorrhage, which was not aneurysm-related. The remaining deaths were noncerebrovascular in nature, including 4 patients who died from cancer and 2 patients who died from unknown causes. In 4 of the 6 patients who had poor neurologic outcome at last follow-up, the cause of the morbidity was related to diffuse intracranial dolichoectasia. Representative case examples are provided in Figs 1–4.

**Table 3: Angiographic characteristics and distribution of DID**

<table>
<thead>
<tr>
<th>Vessels involved</th>
<th>No. (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anterior circulation only</td>
<td>4 (16.0)</td>
</tr>
<tr>
<td>Bilateral ICAs</td>
<td>4 (16.0)</td>
</tr>
<tr>
<td>Anterior and posterior circulation</td>
<td>21 (84.0)</td>
</tr>
<tr>
<td>Basilar artery + 1ICA</td>
<td>10 (40.0)</td>
</tr>
<tr>
<td>Basilar artery + bilateral ICAs</td>
<td>5 (20.0)</td>
</tr>
<tr>
<td>Basilar artery + bilateral ICAs + bilateral MCAs</td>
<td>2 (8.0)</td>
</tr>
<tr>
<td>Basilar artery + 1ICA/ipsilateral MI</td>
<td>2 (8.0)</td>
</tr>
<tr>
<td>Basilar artery + 1ICA + bilateral MCAs</td>
<td>1 (4.0)</td>
</tr>
<tr>
<td>Basilar artery + 1 MCA</td>
<td>1 (4.0)</td>
</tr>
<tr>
<td>Mean (SD) maximum diameter of dilated (mm)</td>
<td></td>
</tr>
<tr>
<td>Basilar artery</td>
<td>15.3 (8.3)</td>
</tr>
<tr>
<td>Internal carotid artery</td>
<td>12.7 (7.0)</td>
</tr>
<tr>
<td>MCA</td>
<td>10.1 (4.3)</td>
</tr>
<tr>
<td>Thrombus in fusiform/dolichoectatic aneurysm</td>
<td>7 (28.0)</td>
</tr>
</tbody>
</table>

**Note:**—DID indicates diffuse intracranial dolichoectasia.

**FIG 1.** A 67-year-old man who is a former smoker with a history of a penetrating atheromatous ulcer of the aortic arch (white arrow), a 5.6-cm abdominal aortic aneurysm (circle), and a celiac artery aneurysm (curved arrow, A). The patient had an episode of dizziness and headache and underwent a noncontrast CT of the head, which demonstrated enlarged intracranial arteries. An MRA demonstrated fusiform aneurysmal dilation of the entire right M1 segment measuring 10 mm in maximum diameter and a largely thrombosed fusiform aneurysm of the basilar artery, which measured 18 mm in maximum diameter (B and C). Approximately 9 months later, the aneurysm grew to 25 mm in diameter and started causing obstructive hydrocephalus (D). The patient also had a new perforator pontine infarct at the time (not shown). A programmable ventriculoperitoneal shunt was placed; however, the patient died due to complications of hydrocephalus 3 months later.

Comparison of Diffuse Intracranial Dolichoectasia and Isolated Verteobasilar Dolichoectasia Groups

A total of 139 patients had vertebrobasilar dolichoectasia with a mean follow-up of 44 ± 42.9 months. A case of vertebrobasilar dolichoectasia is demonstrated in Fig 5. None of the patients with vertebrobasilar dolichoectasia were older (70.9 ± 14.2 years versus 60.4 ± 12.5 years, \( P = .0002 \)), and the age distribution of patients with diffuse intracranial dolichoectasia was older (\( P < .0001 \)). There was a predominance of males in both patient groups (84.0% versus 75.5%, \( P = .51 \)). Patients with diffuse intracranial dolichoectasia were significantly more likely to have abdominal aortic aneurysms or ectasias (62.5% versus 14.3%, \( P = .01 \)) as well as visceral aneurysms (25.0% versus 0.0%, \( P < .0001 \)). Patients with diffuse intracranial dolichoectasia trended toward higher rates of peripheral artery disease (16.0% versus 4.3%, \( P = .17 \)) and hypertension (88.0% versus 70.3%, \( P = .11 \)). Patients with dif-
fuse intracranial dolichoectasia were significantly more likely to smoke (68.0% versus 15.9%, $P < .0001$).

Patients with diffuse intracranial dolichoectasia trended toward higher aneurysm growth rates (46.2% versus 21.6%, $P = .09$) and cerebral infarction rates (28.0% versus 13.0%, $P = .10$) and were more likely to have SAH on follow-up (20.0% versus 3.5%, $P = .007$). Patients with diffuse intracranial dolichoectasia were less likely to have good neurologic function at follow-up (24.0% versus 57.6%, $P = .004$), were more likely to die (52.0% versus 26.6%, $P = .02$), and were more likely to die due to aneurysm-related factors (24.0% versus 7.2%, $P = .02$). These data are summarized in Table 2.

On multivariate analysis, adjusting for age and smoking history, a diagnosis of diffuse intracranial dolichoectasia was independently associated with higher odds of aneurysm rupture (OR = 6.33; 95% CI, 1.23–35.31; $P = .02$) and higher odds of new ischemic stroke (OR = 18.62; 95% CI, 3.63–146.57; $P = .0003$). There was no significant difference in the odds of aneurysm growth on multivariate analysis (OR = 2.23; 95% CI, 0.57–8.68; $P = .24$).

**DISCUSSION**

Our study of 25 patients with diffuse intracranial dolichoectasia demonstrated a number of notable findings. First, this condition is a relatively rare vascular phenotype, predominantly affecting elderly male smokers with hypertension. The natural history of this disease is dismal, with only a minority surviving with good outcome at a mean follow-up of 3.5 years. Furthermore, among patients with abdominal imaging, >60% had abdominal aortic aneurysms or ectasias and nearly 30% of patients had an additional intracranial saccular aneurysm. Compared with a control group of patients with isolated vertebrobasilar dolichoectasia, patients with diffuse intracranial dolichoectasia were more likely to have aortic and visceral aneurysms, were more likely to smoke, and were more likely to experience aneurysm growth, rupture, and morbidity and mortality on both univariate and multivariate analyses. Overall, these findings suggest that diffuse intracranial dolichoectasia may represent a distinct vascular phenotype from isolated vertebrobasilar dolichoectasia and is likely the manifestation of a diffuse, systemic arteriopathy affecting multiple vascular beds. A summary of our proposed diagnostic criteria for diffuse intracranial dolichoectasia and how they compare with current definitions of vertebrobasilar dolichoectasia is provided in Table 4.

**FIG 2.** The patient is a 47-year-old man. He presented in his 20s with 2 large fusiform aneurysms of his cavernous carotid arteries. He later developed an aneurysm of his basilar artery and had a subarachnoid hemorrhage from a dissecting aneurysm of his PICA (not shown). A, CTA image demonstrates a large fusiform aneurysm of the right cavernous carotid artery (long white arrow) and a thrombosed/calcified aneurysm of the left cavernous carotid artery (short white arrows). There is also an aneurysm of the basilar tip (curved black arrow). B, Right ICA cerebral angiogram shows a large fusiform aneurysm of the right petroccavernous carotid artery (straight black arrow) and a basilar tip aneurysm (curved black arrow). C. There was suspicion for underlying connective tissue disease. The patient underwent a skin biopsy. Electron microscopy of the skin biopsy shows multiple abnormally enlarged collagen fibers (black circles) consistent with collagen flowers. These are typically seen in Ehlers-Danlos syndrome. The patient later underwent genetic testing for Loeys-Dietz, Ehlers-Danlos, and Marfan syndromes. The findings of all tests were negative.

**FIG 3.** A 51-year-old man with a long history of headaches with associated nausea and vomiting. He had acute-onset left-sided weakness with a prominent left facial droop along with left face, arm, and leg numbness and slurred speech. The patient had no family history of cerebral aneurysms, though his father had an abdominal aortic aneurysm. A, MR imaging at the time of the initial evaluation showed a medial left pontine infarct (black arrow). There was evidence of a large dolichoectatic aneurysm of the basilar artery on MR imaging, and the patient underwent cerebral angiography for further evaluation. B. Cerebral angiography demonstrated a fusiform-type aneurysm of the basilar artery with a filling defect that was consistent with thrombus (black arrows). The patient also had diffuse arteriomegaly with dilation of the right supraclinoid ICA to 6 mm and dilation of the left supraclinoid ICA to 10 mm (white arrow C). D. The day following the angiography, the patient had a 10/10 headache. Noncontrast CT at the time showed diffuse subarachnoid hemorrhage with most of the blood products surrounding the basilar artery aneurysm. He died the next day.
We fully acknowledge that patients with multivessel intracranial dolichoectasia have been reported previously in the literature. However, prior reports were generally quite small and did not specifically identify multivessel involvement as a distinct entity and did not provide detailed baseline demographic, imaging, and outcome data to fully characterize our proposed subgroup. In a review of 9 studies on vertebrobasilar dolichoectasia, Gutierrez et al\textsuperscript{11} noted that approximately 45% of patients with vertebrobasilar dolichoectasia had some degree of ectasia or dilation of the anterior circulation as well. However, on closer inspection of these studies, it is apparent that the degree of dilation or tortuosity of the anterior circulation and the outcomes of patients with diffuse intracranial dolichoectasia were not reported.\textsuperscript{3,10,12-14} For example, in a case series that our group previously published, 63 of 152 patients (41.5%) with vertebrobasilar nonsaccular dolichoectatic aneurysms had concomitant dolichoectasias in the anterior circulation by visual inspection. However, only 11 of those patients (7.2%) are included in the current study because the other 52 did not meet the size criteria set forth in Table 1. In fact, even in the large surgical series reported by Anson et al\textsuperscript{15} and Drake et al\textsuperscript{16} of patients with surgically treated large fusiform aneurysms, <5 percent of cases would meet the diagnostic criteria of diffuse intracranial dolichoectasia. Other small case series and case reports have reported outcomes of patients with diffuse intracranial dolichoectasia with a well-defined underlying etiology such as prior radiation, autosomal dominant polycystic kidney disease, connective tissue disease, or infection. However, findings from these studies do not necessarily apply to the patient population included in our series due to substantial differences in risk factors and etiology.\textsuperscript{17-23}

Our comparison of the baseline clini-

**FIG 4.** An 87-year-old man with a history of a right third-nerve palsy. A, Right ICA cerebral angiogram demonstrates a 20-mm cavernous carotid fusiform aneurysm with associated dilation of the supraclinoid ICA as well. B, Left ICA cerebral angiogram shows dilation of the left supraclinoid ICA to approximately 10 mm. C, Left vertebral artery cerebral angiogram shows diffuse dilation and tortuosity of the basilar artery measuring 9 mm in maximum diameter. The cause of the third-nerve palsy was thought to be the right cavernous aneurysm.

**FIG 5.** Vertebrobasilar dolichoectasia in a 67-year-old man. A, Right and left ICA cerebral angiograms demonstrate normal-caliber internal carotid arteries, MCAs, and anterior cerebral arteries bilaterally. B, Left vertebral artery cerebral angiogram demonstrates an irregular dolichoectatic and fusiform aneurysm involving the entirety of the basilar artery.

### Table 4: Comparison of diagnostic criteria and outcomes

<table>
<thead>
<tr>
<th>Diffuse Intracranial Dolichoectasia</th>
<th>Vertebrobasilar Dolichoectasia\textsuperscript{a}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imaging appearance</td>
<td>Fusiform aneurysmal dilation of an entire vascular segment (ie, supraclinoid ICA, basilar artery, M1 segment of the MCA)</td>
</tr>
<tr>
<td></td>
<td>Fusiform: aneurysmal dilation without definable neck involving a portion of an arterial segment with any degree of tortuosity Dolichoectatic: uniform aneurysmal dilation of an artery involving the entire basilar or vertebral or both with any degree of tortuosity Transitional: uniform aneurysmal dilation of an artery with superimposed dilation of a portion of the involved arterial segment Verteobasilar system only</td>
</tr>
<tr>
<td>Distribution</td>
<td>≥2 Intracranial vascular beds (ie, vertebrobasilar system, left anterior circulation, or right anterior circulation) Basilar artery diameter of &gt;5.0 mm</td>
</tr>
<tr>
<td>Size criteria</td>
<td>Cavernous ICA: ≥8.5 mm Supraclinoid ICA: ≥8.0 mm MCA: ≥5.0 mm Basilar artery: ≥6.0 mm</td>
</tr>
<tr>
<td></td>
<td>7%/year\textsuperscript{2-10} 3%/year\textsuperscript{2-10} 2%/year\textsuperscript{2-10}</td>
</tr>
<tr>
<td>Growth rate</td>
<td>10%/year 11%/year 6%/year</td>
</tr>
<tr>
<td>Ischemic stroke risk</td>
<td>7%/year\textsuperscript{2-10} 3%/year\textsuperscript{2-10} 2%/year\textsuperscript{2-10}</td>
</tr>
<tr>
<td>Aneurysm rupture risk</td>
<td>7%/year\textsuperscript{2-10} 3%/year\textsuperscript{2-10} 2%/year\textsuperscript{2-10}</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Definitions of vertebrobasilar dolichoectasia proposed by Flemming et al.\textsuperscript{2}
cal characteristics and outcomes of patients with diffuse intracranial dolichoectasia with patients with isolated vertebrobasilar dolichoectasia suggests that at the very least, this entity is a distinctive phenotype or a more diffuse and severe form of vertebrobasilar dolichoectasia. In general, patients with diffuse intracranial dolichoectasia were older, were more likely to have hypertension and other atherosclerotic risk factors, and were more likely to smoke. Furthermore, patients in the diffuse intracranial dolichoectasia group were significantly more likely to have intracranial, abdominal aortic, and visceral arterial aneurysms than those with isolated vertebrobasilar dolichoectasia. These differences are not only isolated to patients studied at our institution. Compared with the systematic review of the literature on vertebrobasilar dolichoectasia published by Gutierrez et al, patients in our current series were generally older (71 versus 63 years), were more often male (84% versus 64%), and were more likely to have hypertension (88% versus 66%), smoke (68% versus 42%), and have other atherosclerotic risk factors. Furthermore, considering all patients in our series with abdominal imaging, 50% of patients had abdominal aortic aneurysms compared with an average of 11% in the vertebrobasilar dolichoectasia literature.11 The rate of concomitant saccular aneurysm formation is twice as high as that reported in the literature as well (28% versus 15%).

The natural history of patients with diffuse intracranial dolichoectasia was also substantially more aggressive than that of the isolated vertebrobasilar dolichoectasia group as well as that reported in the vertebrobasilar dolichoectasia literature. Twenty percent of patients in our series had subarachnoid hemorrhage on follow-up, for a subarachnoid hemorrhage rate of 6.9%/patient-year. This compares with rupture rates of just 1%–2% per year reported for dolichoectatic and fusiform aneurysms of both the anterior and posterior circulations.3,4,24 The growth rate for patients with diffuse intracranial dolichoectasia was also substantially higher than that reported in the literature among patients with single-vessel dolichoectatic aneurysms (9.5%/year versus 6.5%/year).10 More than 50% of the patients in the current series had ischemic stroke after a mean of approximately 5 years of follow-up. Meanwhile, the rates of ischemic stroke among previously published large series of vertebrobasilar dolichoectasia are 3%, 11%, 16%, and 38% at 1, 5, 10, and 12 years, respectively.11

Limitations
Our study has limitations. First, this is a retrospective study spanning >15 years at a single institution. Therefore, there is wide variation in the type of imaging used, clinical management of these patients, and degree of follow-up. In addition, our study is prone to substantial selection bias. Patients who presented with diffuse intracranial dolichoectasia at our tertiary referral institution were more likely to have complications related to this disease, which could explain the poor natural history seen in our study. We do not have access to pathologic specimens or gene samples from the patients included in this study. Thus, we are unable to determine whether the pathologic basis of this disease is similar to that in other fusiform aneurysms and ectasias or whether the degree of vessel wall weakening is, in fact, more severe. While none of the patients in our series had a clinical diagnosis of a genetic connective tissue disease such as Marfan syndrome, Ehlers-Danlos syndrome, or Loeys-Dietz syndrome, most of the patients included in this series did not receive any genetic testing, so we cannot rule out the presence of an undiagnosed connective tissue disease. Therefore, the total number of patients or the association to secondary diseases may not be as high as that reported in our study. In our comparison of the natural history of diffuse intracranial dolichoectasia with that of vertebrobasilar dolichoectasia, we did not account for the fact that the morphology of the vertebrobasilar dolichoectatic lesions has a significant impact on their natural history. As defined in Table 4, dolichoectatic vertebrobasilar aneurysms have a much more favorable natural history than fusiform or transitional lesions.10 Last, because this entity has not been extensively studied outside small case series and case reports, there are no well-defined diagnostic criteria. Our choice of size criteria in diagnosing diffuse intracranial dolichoectasia are somewhat arbitrary (doubling of mean values reported in the literature). Using such strict criteria could have resulted in exclusion of patients with milder or earlier forms of this disease.

CONCLUSIONS
Our study of 25 patients with diffuse intracranial dolichoectasia demonstrated that this rare disease is associated with a very poor natural history, with about 20% of patients having subarachnoid hemorrhage on follow-up, 30% of patients having ischemic stroke, and 25% of patients with aneurysm-related mass effect. Nearly two-thirds of patients with this disease have abdominal aortic aneurysms, and 25% have additional saccular intracranial aneurysms. Given the high morbidity and mortality rates associated with this disease and the high rate of concomitant intracranial saccular aneurysm and abdominal aortic aneurysms in this population, these findings suggest that diffuse intracranial dolichoectasia may represent a distinct vascular phenotype. Further research is needed to better characterize the histopathologic and genetic features of diffuse intracranial dolichoectasia and determine ideal treatment options.

Disclosures: Giuseppe Lanzino—UNRELATED: Consultancy: Coviden/Medtronic.* David F. Kallmes—UNRELATED: Consultancy: Minnetronix.* Comments: Data and Safety Monitoring Board. Grants/Grants Pending: Medtronic. MicroVention, Sentre Medical, Sumorodix, Codman, NeuroSigma, Comments: preclinical research and clinical trials; Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed: Medtronic, Comments: FDA panel meeting.* Money paid to the institution.
Detection of Focal Longitudinal Changes in the Brain by Subtraction of MR Images


ABSTRACT

BACKGROUND AND PURPOSE: The detection of new subtle brain pathology on MR imaging is a time-consuming and error-prone task for the radiologist. This article introduces and evaluates an image-registration and subtraction method for highlighting small changes in the brain with a view to minimizing the risk of missed pathology and reducing fatigue.

MATERIALS AND METHODS: We present a fully automated algorithm for highlighting subtle changes between multiple serially acquired brain MR images with a novel approach to registration and MR imaging bias field correction. The method was evaluated for the detection of new lesions in 77 patients undergoing cardiac surgery, by using pairs of fluid-attenuated inversion recovery MR images acquired 1–2 weeks before the operation and 6–8 weeks postoperatively. Three radiologists reviewed the images.

RESULTS: On the basis of qualitative comparison of pre- and postsurgery FLAIR images, radiologists identified 37 new ischemic lesions in 22 patients. When these images were accompanied by a subtraction image, 46 new ischemic lesions were identified in 26 patients. After we accounted for interpatient and interradiologist variability using a multilevel statistical model, the likelihood of detecting a lesion was 2.59 (95% CI, 1.18–5.67) times greater when aided by the subtraction algorithm (P = .017). Radiologists also reviewed the images significantly faster (P < .001) by using the subtraction image (mean, 42 seconds; 95% CI, 29–60 seconds) than through qualitative assessment alone (mean, 66 seconds; 95% CI, 46–96 seconds).

CONCLUSIONS: Use of this new subtraction algorithm would result in considerable savings in the time required to review images and in improved sensitivity to subtle focal pathology.

Increasingly, diagnostic radiologic assessment does not involve a single examination at 1 time point but makes use of follow-up imaging for monitoring responses during the course of treatment or to assess the progression of a condition with a view to intervention.1–3 Manual interpretation of this type of data is burdensome and prone to errors. A challenge for medical image analysts is the development of reliable techniques to assist the radiologist.4 Easier and faster detection of pathologic changes facilitates diagnosis and makes it more cost-effective to monitor disease evolution in the long term.

Manual analysis of medical images acquired at different time points is generally limited to a qualitative comparison. The repositioning of the patient is never identical, the scanning equipment used may be different or may have been upgraded, the acquisition parameters may have altered between scans, and local shape deformations of nonrigid anatomic structures occur. The expert radiologist compensates intuitively for some of these unwanted differences with knowledge of anatomy and the expected signal differences to identify and reject certain artifactual changes.

One of the major artifacts associated with MR imaging is caused by the spatially nonuniform transmission and reception properties of the radiofrequency coils.5 Commonly, a large transmitter coil built into the bore of the magnet is used for radiofrequency transmission, resulting in relatively uniform transmission properties. However, to improve the signal-to-noise ratio, an array of small coils is used for signal reception, and these have very spatially nonuniform reception properties.6 It is common to correct for this by mapping the reception properties of the coils and
rescaling the image intensities pixel by pixel.2 This mapping, however, is based on a number of assumptions and does not remove the image-intensity shading, known as the bias field, completely. It is also possible to perform a bias field correction as a postprocessing step, but again this is based on a number of assumptions about the nature of the field and its stability.9,10 The algorithm described in the current work circumvents these issues by correcting the bias field in subtraction images rather than in the source images themselves.

This article describes an image registration and subtraction technique that can highlight local changes in MR images of the brain acquired at ≥2 time points. It is robust to changes in the patient position and in the bias field. The technique was evaluated by using a set of 77 patients who were scanned before and after cardiac surgery, with the aim of assessing whether this algorithm could aid in the detection of small ischemic brain lesions resulting from emboli shed from the vascular system as a result of an operation.10 The value of the system was assessed by examining the lesion detection rate and time required for the assessment of scans for radiologists examining pairs of images with and without the aid of our subtraction method.

**MATERIALS AND METHODS**

**Registration and Subtraction**

The method is based on registration of images acquired at different points in time, the correction of any differences due solely to radiofrequency reception nonuniformity, and subsequent subtraction to reveal any differences due to focal pathology. The method is described in detail elsewhere11 and is illustrated in Fig 1. Briefly, input images are registered to a common frame of reference at the average position of all input images. Then, radiofrequency bias field variations are modeled as a smoothly varying polynomial function; then, individual registered images are corrected by pixel-wise division by the modeled field. Finally, individual pairs of registered and corrected images are subtracted (Fig 2).

**Patients**

We examined data from patients enrolled in a clinical study of neurologic deficits following cardiac surgery, as detailed elsewhere.10 All clinical data were collected in accordance with local research ethics procedures. All patients provided written informed consent for their anonymized data to be used in this research.

**MR Imaging**

MR imaging was performed by using a 3T system (Magnetom Skyra; Siemens, Erlangen, Germany) with a 20-channel phased array head and neck coil for signal reception. MR imaging studies were performed 1–2 weeks before cardiac surgery and repeated 6–8 weeks after surgery to identify any new postoperative ischemic lesions. The same scanning protocol was used on both occasions, except that the angiogram was not obtained postsurgery. After a localizer image, the scanning protocol consisted of a diffusion-weighted image, time-of-flight angiography, a susceptibility-weighted image, and a FLAIR image.12 The total imaging time was approximately 30 minutes.

For this study, only the FLAIR images were analyzed. The FLAIR images consisted of 48, 3-mm-thick contiguous axial sections positioned parallel to the anterior/posterior commissure line. The k-space matrix size was 320 × 224, zero-filled to give an image matrix of 320 × 320 and an FOV of 240 × 240 mm (TR/TE/TI, 6770/108/2170 ms).

**Image Review and Analysis**

Images were analyzed independently by 3 experienced neuroradiologists in 2 stages. First, for each patient, the pre- and postsurgery FLAIR images were presented side by side on a computer display and the radiologist paged through both the images section by section to identify any new ischemic lesions. The anatomic regions of any new lesions were also noted. Signal hyperintensity clearly extending outside the brain parenchyma or minor variations in ependymal signal hyperintensity (commonly encountered on FLAIR sequences) were considered artifactual. A digital timer was used by each neuroradiologist to note the time taken to review images from each patient.

Second, for each patient, the coregistered pre- and postsurgery FLAIR images along with the difference image were presented. The neuroradiologist paged through all 3 images simultaneously section by section to identify any new ischemic lesions on the difference image, with the pre- and postsurgery images used to confirm or refute the findings. Thus, it is unlikely that there were any false-positive findings. A subsequent review of the original
FLAIR images confirmed that lesions originally overlooked were visible in retrospect.

The time required to perform each type of analysis was recorded. There was a gap of at least 4 weeks between the first and second analysis for any given patient, and patient scans were presented in an arbitrary order to avoid familiarity with the images.

Statistical Analysis
The total evaluation time and the number of lesions detected in 77 patients by 3 radiologists were compared for both the qualitative and subtraction-aided review methods. We assumed that the total number of lesions follows a Poisson distribution and modeled our results with a generalized linear mixed model. The model included the method of review (2 levels: qualitative- and subtraction-aided review) as a fixed effect, individual patient and radiologist as random effects, and the natural logarithm function as a link function. To assess sensitivity to detect the presence of a lesion, we created a binary variable to represent the presence or absence of a lesion with a value of 1 when the total number of lesions was greater than zero, and zero otherwise. We assumed that this binary variable follows a Bernoulli distribution and modeled the presence of lesion data by using a generalized linear mixed model approach (logistic generalized linear mixed model). The logistic generalized linear mixed model included the logit function as a link function and the same fixed and random effects structure as described for the Poisson generalized linear mixed model.

To account for increased variability with the mean review time, we transformed the data on total processing time by using the natural logarithm function, and the transformed data were analyzed by using a linear mixed model incorporating the method of review as a fixed effect and individual patient and radiologist as random effects. For all models, a predictor was considered statistically significant if the 2-sided type I error rate was <5% (ie, \( P < .05 \)). All statistical analyses were performed by using R statistical and computing software, Version 3.3 (http://www.r-project.org/) with appropriate packages (lme4, ggplot2).

RESULTS
Pairs of images from 77 patients (72 men; mean age, 62.8 ± 10.4 years; range, 32–80 years) were evaluated independently by 3 radiologists. These 77 patients had postoperative MR imaging performed at a mean of 7 ± 1.25 weeks after surgery. Five of the 77 (7%) patients had perioperative strokes: Patients 12 and 61 had a lacunar infarct in the right corona radiata, patient 46 had a lacunar infarct in the left corona radiata, patient 18 had a lacunar infarct in the right frontal lobe, and patient 62 had 2 small lacunar infarcts in the right superior parietal lobule and the left medial percentual gyrus.

With just the pre- and postsurgery FLAIR images, on average, 37 new ischemic lesions were identified in 22 patients (median, 1 lesion per patient; range, 1–5 lesions). The anatomic locations of these new lesions were the following: left frontal WM (18 lesions); right frontal WM (9 lesions); right parieto-occipital subcortical WM (2 lesions); left corona radiata (2 lesions); left centrum semiovale (1 lesion); left anterior striatum (2 lesions); and left (1 lesion) and right (2 lesions) cerebellum. Lesions missed during qualitative assessment were frequently subsequently identified among preexisting white matter hyperintensities.

Using the difference image in combination with the pre- and postsurgery FLAIR images, we identified an average of 46 new ischemic lesions in 26 patients (median, 1 lesion per patient; range, 1–8 lesions). Therefore, the number of new ischemic lesions identified increased by 24%, and the number of patients who had new lesions increased by 18% when review was aided by a subtraction image. The anatomic locations of these new lesions were the following: left frontal WM (20 lesions); right frontal WM (9 lesions); left (1 lesion) and right (1 lesion) parieto-occipital subcortical WM; right parietal WM (1 lesion); left temporal lobe (2 lesions); left (2 lesions) and right (1 lesion) corona radiata; left centrum semiovale (2 lesions); left parietal cortex (1 lesion); left centrum semiovale (2 lesions); left parietal cortex (1 lesion); left

counts were 67 and 80 for the qualitative and subtractive methods, respectively. The cerebellar lesions not appreciated on the subtraction image were larger and were definitely present on repeat viewing. Of the 3 radiologists, only 1 confirmed these lesions by using the image-subtraction algorithm. The exceptions were 1 lesion in the left and 2 in the right frontal WM and 1 lesion in the left and 1 in the right cerebellum. On further examination of the original FLAIR images, the lesions in the frontal white matter were very small (<1 mm) and of a level similar to that of the background “mottle.” These were confirmed by only 1 of the 3 radiologists. The cerebellar lesions not appreciated on the subtraction image were larger and were definitely present on repeat viewing.

When we pooled the data over all radiologists, the mean lesion counts were 67 and 80 for the qualitative and subtractive methods, respectively, though there was no evidence that the mean lesion counts were significantly different between methods ($P = .101$). When we accounted for the variabilities between patients and radiologists, the logistic generalized regression model suggested that the likelihood of detecting the presence of a lesion was 2.59 (95% CI, 1.18–5.67) times greater with the subtraction compared with the qualitative method ($P = .017$).

When we accounted for the variability in the data between patients (interpatient variability) and between radiologists (inter-radiologist variability), the mean time to review the images was significantly ($P < .001$) lower with the subtraction algorithm (mean, 42 seconds; 95% CI, 29–60 seconds) compared with qualitative assessment (mean, 66 seconds; 95% CI, 46–96 seconds), reducing the processing time by 37%. Estimates of between-patient variability were slightly higher (0.27; 95% CI, 0.22–0.34) than between-radiologist variability (0.17; 95% CI, 0.07–0.42).

**Figure 3.** The logarithm of time (seconds) taken to process the data by qualitative comparison and aided by the image-subtraction algorithm. The scatterplots show the triplicate log-transformed time data (of 3 radiologists) for each patient along with boxplots presenting the median, lower, and upper quartiles for each algorithm. Points with the same color represent data from the same patient. The plot also presents the predicted mean time (square marker) along with corresponding 95% confidence intervals (error bar) for both methods obtained from a linear mixed-model fit to the data.

DISCUSSION
We have demonstrated a method for the registration, subtraction, and bias field correction of serial MR images, which can be applied in longitudinal patient studies. The subtraction image highlights the presence of new lesions, which can be confirmed or refuted on closer inspection of the source images. The method is tailored toward robust detection of focal differences in pathology that develop or resolve between scans.

Axial FLAIR imaging frequently has CSF flow-related artifacts in the posterior fossa, which necessarily result in a more heterogeneous parenchymal background on the subtraction image. On the subtraction image, the preexisting lesions were removed, but new lesions were highlighted as hyperintense, thus aiding detection. Similarly, posterior fossa lesions on FLAIR are sometimes difficult to separate from adjacent flow-related signal hyperintensity and can be easily overlooked, particularly at the periphery of the cerebellum.

Although several other studies have previously used subtraction imaging, the procedure outlined here has several advantages. First, it can be used to analyze MR imaging data acquired serially over many time points in an unbiased fashion. Most previous studies have either selected one of the image time points as the registration target or have been restricted to showing the differences between pairs of images by realigning them to a position half-way between the 2 scans. By transforming all of the images to their average position, the current method allows the alignment of any number of image time points. This is achieved by first registering them to a generic atlas target to facilitate relative positioning.

Second, previous studies have incorporated bias field correction by using, for example, the N3 method, in which the input images are individually bias-corrected as the first processing step. However, the N3 method is prone to errors in the bias field correction scheme, leading to residual shading artifacts in the subtracted images. The current method provides an alternative correction scheme that relies on only the assumption that in the absence of bias field differences, tissues that are not involved in focal pathologic changes should remain substantially unchanged in intensity between image acquisitions. Thus, the differences in the bias field between scans can be estimated and used to correct the subtracted images.

We believe that this method could provide a useful tool for radiologists when reviewing serially acquired MR images. First, the time required to review the images was reduced by approximately 37% when using the registered and subtracted images. All
processing was fully automatic and therefore could be performed automatically as an adjunct to the acquisition of new scans and retrieval of previous scans from a PACS system.

Second, the number of new ischemic lesions identified increased by 24%, and the number of patients who had new lesions increased by 18% when using the subtraction images compared with simply viewing the 2 scans. The likelihood of observing a lesion was also 2.6 times greater when using the new algorithm. These represent substantial increases in sensitivity in detecting new ischemic tissue damage. Although this pathology consisted of relatively small ischemic lesions and only new pathology was identified, subtle changes in tissue in other diseases, such as MS, might be equally identified and could prove important in monitoring disease progression or stability and guiding treatment decisions.

On the other hand, when using the subtraction images, 2 lesions in the cerebellum were missed. On subsequent review, these lesions were clearly present in the view of the radiologists. The 2 scans for each patient were acquired approximately 8–10 weeks apart, so there was little chance for more substantial tissue changes to develop or for cerebral atrophy to advance. We used a rigid-body (translation and rotation) registration to align the 2 scans, which proved adequate in all cases. More sophisticated non-linear registration would be required for longer follow-up periods or to account for more profound structural changes in the brain, for example, in the follow-up of patients with brain tumor and after an operation or in cases of severe atrophy.

To provide an unbiased image registration, the method uses a generic atlas in the first processing step. Because the atlas is used only to find the approximate head position for the serially acquired images, it is likely that other generic atlases could also be used without compromising the quality of the final registration. The scan parameters were kept constant across both scans because the method is unlikely to work well when there are gross differences in contrast between the 2 images. Constraining the scan parameters was easy to do in the context of our study but might be more problematic when patients are rescanned in a routine setting, possibly with different scanner types from different manufacturers. On the other hand, particularly if a 3D acquisition is used, the method should be relatively robust to gross changes in the patient’s head position because of the novel bias field correction used. However, the effects of head position and scanner settings would need to be assessed in future studies.

CONCLUSIONS
We have developed a fully automated image registration and subtraction scheme capable of highlighting small changes in serial MR imaging scans of the brain. The method should now be used in larger studies to investigate its clinical potential in monitoring progression in different diseases and for longer time periods.


REFERENCES

Acute Cytotoxic and Vasogenic Edema after Subarachnoid Hemorrhage: A Quantitative MRI Study

J.M. Weimer, S.E. Jones, and J.A. Frontera

ABSTRACT

BACKGROUND AND PURPOSE: The mechanism of early brain injury following subarachnoid hemorrhage is not well understood. We aimed to evaluate if cytotoxic and vasogenic edema are contributing factors.

MATERIALS AND METHODS: A retrospective analysis was conducted in patients with SAH undergoing diffusion-weighted MR imaging within 72 hours of onset. Apparent diffusion coefficient values derived from DWI were evaluated by using whole-brain histograms and 19 prespecified ROIs in patients with SAH and controls with normal findings on MRI. Cytotoxic edema observed outside the ROIs was assessed in patients with SAH. The average median ADC values were compared between patients with SAH and controls and patients with SAH with mild (Hunt and Hess 1–3) versus severe early brain injury (Hunt and Hess 4–5).

RESULTS: We enrolled 33 patients with SAH and 66 controls. The overall average median whole-brain ADC was greater for patients with SAH (808 × 10⁻⁶ mm²/s) compared with controls (788 × 10⁻⁶ mm²/s, P < .001) and was higher in patients with SAH across ROIs after adjusting for age: cerebral gray matter (826 versus 803 × 10⁻⁶ mm²/s, P = .059), cerebral white matter (793 versus 758 × 10⁻⁶ mm²/s, P = .023), white matter tracts (797 versus 739 × 10⁻⁶ mm²/s, P < .001), and deep gray matter (754 versus 713 × 10⁻⁶ mm²/s, P = .016). ADC values trended higher in patients with Hunt and Hess 4–5 versus those with Hunt and Hess 1–3. Early cytotoxic edema was observed in 13 (39%) patients with SAH and was more prevalent in those with severe early brain injury (87.5% of patients with Hunt and Hess 4–5 versus 24.0% of those with Hunt and Hess 1–3, P = .001).

CONCLUSIONS: Age-adjusted ADC values were globally increased in patients with SAH compared with controls, even in normal-appearing brain regions, suggesting diffuse vasogenic edema. Cytotoxic edema was also present in patients with SAH and correlated with more severe early brain injury.

ABBREVIATIONS: EBI = early brain injury; HH = Hunt and Hess

E arly brain injury (EBI) incurred during aneurysm rupture in spontaneous subarachnoid hemorrhage is a major predictor of poor functional outcome, yet the mechanism for EBI is not well-understood. In both animal and human models, SAH leads to transiently elevated intracranial pressure with concomitant inadequate cerebral blood flow and, in severe cases, intracranial circulatory arrest. This transient global hypoperfusion is associated with endothelial activation, microthrombosis, ischemia, and vasogenic edema in animal models. As a part of routine clinical MR imaging at many institutions, diffusion-weighted imaging presents a unique opportunity for the study of patients with SAH in the acute period. Apparent diffusion coefficient values may serve as a practical and useful biomarker for the severity of EBI following SAH. In humans, we and others have demonstrated that MR imaging–detected infarctions on DWI occur acutely after SAH and before the onset of delayed cerebral ischemia/vasospasm.

These infarctions are more common in patients with more severe EBI (Hunt and Hess [HH] 4–5); occur in unusual, nonvascular patterns (eg, corpus callosum, bilateral medial frontal lobes); and are associated with an increased risk of delayed cerebral ischemia and worse 3-month functional outcomes. The volume of infarction associated with worse functional outcomes is small, however. Thus, this finding led us to hypothesize that conventional, nonquantitative MR imaging techniques are not sensitive enough to detect the full extent of brain injury.
We hypothesized that patients with SAH (compared with controls) would demonstrate global reductions in apparent diffusion coefficient values, indicating diffuse cytotoxic edema presumably due to ischemia/hypoperfusion, when whole-brain ADC mapping and ROI quantitative analyses were applied. We further hypothesized that patients with SAH with evidence of severe EBI would have a greater burden of cytotoxic edema than those with mild EBI.

MATERIALS AND METHODS

Subject Enrollment and Data Collection
A retrospective review of prospectively enrolled patients with spontaneous SAH was conducted between February 2013 and March 2015 as part of the Intracranial Hemorrhage ICU Project of Cleveland Clinic and Early Status of Coagulation, Platelet Activation and Outcome after Subarachnoid Hemorrhage (Principal Investigator, J.A.F.). As part of these studies, all eligible patients admitted to the neurologic intensive care unit within 72 hours of the onset of SAH symptoms were approached for enrollment. Inclusion criteria were being 18 years of age or older; SAH diagnosed by CT, MR imaging, or lumbar puncture; consent to participate in 3- and 12-month follow-up interviews; and DWI performed within 72 hours of ictus, before the clinical or radiographic onset of delayed cerebral ischemia/cerebral vasospasm as performed elsewhere (Hunt and Hess 4–5) as performed elsewhere.11

The control group was identified from a population of consecutive patients undergoing brain DWI on the same 1.5T inpatient scanners and were matched to patients with SAH in a 2:1 ratio. Records were queried from the hospital electronic record on in-patients 40–60 years of age (because this is the typical age range of patients with spontaneous SAH) undergoing brain MR imaging between January 2010 and December 2013 (the most recent time frame in which control data were available). Radiologists’ image reports were reviewed to identify patients with normal findings on MRI and no evidence of intracranial pathology. Images were then independently reviewed by 2 reviewers (J.A.F., J.M.W.) to confirm the absence of intracranial pathology. Basic and demographic data were collected for both patients with SAH and controls. Admission clinical and radiographic data were also recorded for patients with SAH. This study was approved by our institutional review board.

MR Imaging Postprocessing and Analysis
All MRIs were performed on the same 1.5T whole-body scanners (Magnetom; Siemens, Erlangen, Germany), and diffusion-weighted images were obtained by using a standard single-shot, spin-echo, echo-planar sequence with b-values of 0 mm²/s and 1000 mm²/s. ADC images were automatically generated from diffusion-weighted sequences (b=0 and b=1000 series) following acquisition. To confirm their accuracy, ADC images were also reconstructed off-line by using Matlab R2013 software (MathWorks, Natick, Massachusetts). Because differences between the scanner-generated and off-line-reconstructed images were negligible, representing <1% of the maximal ADC signal, automatically generated ADC images were used for subsequent analyses.

Brain extractions were next performed on scans of all controls and patients with SAH with the FSL Brain Extraction Tool (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET), available through FSL (http://www.fmrib.ox.ac.uk/fsl; 2015) as part of the Mango image viewing software package (http://ric.uthscsa.edu/mango; 2015). Floor and ceiling threshold filters were applied to limit the images to voxel intensities between 250 and 1250 × 10⁻⁶ mm²/s to restrict contributions from any remaining nonparenchymal signal, such as from the extra-axial or subarachnoid space (thereby excluding blood products in the subarachnoid space and intraventricular blood). The median whole-brain ADC was calculated for each subject. Whole-brain histograms with bin widths of 10 × 10⁻⁶ mm²/s were generated. Histograms were normalized to account for variations in brain volume among patients, and histogram peaks were calculated. Smooth curves were fit to model the histograms by applying a moving average low-pass filter. Median ADC values were compared for both patients with SAH versus control subjects and those with mild EBI (Hunt and Hess 1–3) versus severe EBI (Hunt and Hess 4–5) as performed elsewhere.11

Traditional regional spatial profile analysis was also performed by using standardized volumetric software in a manner similar to that previously described.14 Images were loaded into the Medical Image Processing, Analysis, and Visualization software package (National Institutes of Health, Bethesda, Maryland; 2014), and symmetric ROIs were placed bilaterally over 19 pre-defined normal-appearing regions: gray matter of the frontal, parietal, occipital, and temporal lobes; white matter of the frontal, parietal, occipital, and temporal lobes; internal capsule, genu, and splenium of the corpus callosum; caudate head, lentiform nucleus, and thalamus; midbrain; hippocampus; pons; cerebellar nuclei; and white matter (On-line Figure).12 Median ADC values were reported for each subject ROI due to the possibility of individual voxel outliers within an individual ROI. Conclusions were unchanged if mean ADC values were used during ROI analysis and therefore were not reported. Average median ADC intensity was calculated for the following ROI combinations: cerebral gray matter (frontal, parietal, occipital, and temporal); cerebral white matter (frontal, parietal, occipital, and temporal); white matter tracts (genu, splenium, and internal capsule); deep gray matter (thalamus, caudate, and lentiform nucleus); and summed ROIs (all 19 regions). Hyperintense FLAIR signal related to remote lesions (eg, old stroke) or transependymal periventricular edema was excluded from ROI analysis. Ischemic lesions were not included in summed ROI comparisons and are reported separately as their own ROI group.

For patients with SAH, DWI and ADC images were also reviewed for evidence of cytotoxic edema outside prespecified ROIs, defined qualitatively as DWI hyperintense/ADC hypointense lesions. Cytotoxic edema lesions were manually traced to calculate total volume. Procedural-related lesions, such as those along known external ventricular drain tracts or craniotomy tra-
EBI were performed by using the t-test. Comparisons between patients with SAH and controls and SAH admission data are summarized in Table 1. The median age of patients with SAH (59 years; interquartile range, 44.5–66 years) was significantly older than that in the control cohort (49 years; interquartile range, 43.4–53.2; P < .001). There were no other significant baseline differences, including sex, race, history of smoking, history of alcohol abuse, hypertension, diabetes, hyperlipidemia, or aspirin use. Eight patients with SAH (24.2%) had severe EBI (HH 4–5) on admission. Most patients with SAH had either modified Fisher grade 3 (n = 18, 54.5%) or modified Fisher grade 4 bleeds (n = 9, 27.3%).

The average median whole-brain ADC value was significantly higher for the SAH cohort (808 × 10⁻⁶ mm²/s) compared with the control cohort (788 × 10⁻⁶ mm²/s, P < .001). Similarly, whole-brain histograms were right-shifted toward higher ADC values in the SAH cohort compared with controls (Fig 1). The peak of the averaged SAH histogram was located at the bin centered at 780 × 10⁻⁶ mm²/s, while the peak of the averaged control histogram was located at the bin centered at 750 × 10⁻⁶ mm²/s.

Average median ADC values for ROIs in patients with SAH and controls are reported in Table 2. Across all ROIs, the average median ADC was higher in the SAH cohort compared with the control cohort (788 × 10⁻⁶ mm²/s, P < .001). The greatest differences were observed in the white matter tracts and the cerebral white matter of the brain, along with deep gray matter structures, while differences among gray matter ROIs were smaller.

When we compared patients with SAH with mild EBI (HH 1–3) versus severe EBI (HH 4–5), the average median whole-brain ADC values were similar between groups (808 × 10⁻⁶ mm²/s for HH 1–3 versus 811 × 10⁻⁶ mm²/s for HH 4–5, P = .514). However, the histogram peak for those with mild EBI

Table 1: Summary of admission, demographic, clinical, and radiographic characteristics of patients with SAH4

<table>
<thead>
<tr>
<th></th>
<th>Patients with SAH (n = 33)</th>
<th>Controls (n = 66)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic and demographic data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (yr)</td>
<td>59 (44.5–66)</td>
<td>49 (43.4–53.2)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Female sex</td>
<td>16 (48.5%)</td>
<td>41 (62.1%)</td>
<td>.196</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>25 (75.8%)</td>
<td>47 (72.3%)</td>
<td>.343</td>
</tr>
<tr>
<td>Black</td>
<td>8 (24.3%)</td>
<td>14 (21.5%)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>4 (6.2%)</td>
<td></td>
</tr>
<tr>
<td>History of smoking</td>
<td>21 (63.6%)</td>
<td>34 (51.5%)</td>
<td>.253</td>
</tr>
<tr>
<td>History of alcohol abuse</td>
<td>7 (22.6%)</td>
<td>10 (15.2%)</td>
<td>.370</td>
</tr>
<tr>
<td>Hypertension</td>
<td>14 (45.2%)</td>
<td>33 (50.0%)</td>
<td>.657</td>
</tr>
<tr>
<td>Diabetes</td>
<td>4 (12.9%)</td>
<td>14 (21.2%)</td>
<td>.326</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>12 (38.7%)</td>
<td>24 (36.4%)</td>
<td>.824</td>
</tr>
<tr>
<td>Aspirin use</td>
<td>10 (31.3%)</td>
<td>22 (33.3%)</td>
<td>.837</td>
</tr>
<tr>
<td>History of ischemic stroke</td>
<td>1 (3.0%)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>History of intracranial hemorrhage</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>History of head trauma</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>History of brain tumor</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Family history of aneurysm</td>
<td>4 (13.3%)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Cocaine use within 24 hr of ictus</td>
<td>1 (3.0%)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Admission clinical and radiographic data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Premorbid modified Rankin Scale score</td>
<td>0 (0–0)</td>
<td>5 (15.2%)</td>
<td></td>
</tr>
<tr>
<td>HH grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5 (15.2%)</td>
<td>14 (42.4%)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>6 (18.2%)</td>
<td>5 (15.2%)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3 (9.1%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2 (6.1%)</td>
<td>18 (54.5%)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>9 (27.3%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glasgow Coma Score</td>
<td>15 (11.5–15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>APACHE II score</td>
<td>7 (3–10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modified Fisher grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4 (12.1%)</td>
<td>2 (6.1%)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2 (6.1%)</td>
<td>18 (54.5%)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>9 (27.3%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>13 (5–21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hija drug score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebral edema on admission CT</td>
<td>8 (25.0%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>External ventricular drain placed</td>
<td>14 (42.4%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: —APACHE II indicates Acute Physiology and Chronic Health Evaluation II. 4 Values are reported as No. (%) or median (interquartile range). Comparative basic and demographic data are provided for controls. b Statistical significance.

Statistical Plan
Statistical analyses were performed by using the JMP 9.0 software package (SAS Institute, Cary, North Carolina; 2010). Univariate comparisons between patients with SAH and controls and between patients with SAH with severe (HH 4–5) versus mild (HH 1–3) EBI were performed by using the χ² test for categoric variables and the Wilcoxon rank sum test for continuous variables. Multisample comparisons were performed by using the Kruskal-Wallis test. For average median ADC comparisons between patients with SAH and controls with whole-brain and combined ROI measurements (cerebral white matter, cerebral gray matter, white matter tracts, deep gray matter, and summed ROIs), multivariable logistic regression models adjusted for patient age were also performed because there were significant age differences between patients with SAH and controls. Correlations between ADC values and age, the modified Fisher score, the presence of elevated intracranial pressure, and external ventricular drainage were evaluated by using 2-sided Spearman correlation coefficients.

RESULTS
During the study period, 33 patients with SAH underwent MR imaging within 72 hours of ictus and 66 controls 40–60 years of age with normal-appearing MR imaging were evaluated. The most common indications for MR imaging in the control group were headache, altered mental status, weakness, and vertigo/dizziness/syncope. Basic and demographic data for patients with SAH and controls and SAH admission data are summarized in Table 1. The median age of patients with SAH (59 years; interquartile range, 44.5–66 years) was significantly older than that in the control cohort (49 years; interquartile range, 43.4–53.2; P < .001). There were no other significant baseline differences, including sex, race, history of smoking, history of alcohol abuse, hypertension, diabetes, hyperlipidemia, or aspirin use. Eight patients with SAH (24.2%) had severe EBI (HH 4–5) on admission. Most patients with SAH had either modified Fisher grade 3 (n = 18, 54.5%) or modified Fisher grade 4 bleeds (n = 9, 27.3%).

The average median whole-brain ADC value was significantly higher for the SAH cohort (808 × 10⁻⁶ mm²/s) compared with the control cohort (788 × 10⁻⁶ mm²/s, P < .001). Similarly, whole-brain histograms were right-shifted toward higher ADC values in the SAH cohort compared with controls (Fig 1). The peak of the averaged SAH histogram was located at the bin centered at 780 × 10⁻⁶ mm²/s, while the peak of the averaged control histogram was located at the bin centered at 750 × 10⁻⁶ mm²/s.

Average median ADC values for ROIs in patients with SAH and controls are reported in Table 2. Across all ROIs, the average median ADC was higher in the SAH cohort compared with the control cohort (789 × 10⁻⁶ mm²/s, P < .001). The greatest differences were observed in the white matter tracts and the cerebral white matter of the brain, along with deep gray matter structures, while differences among gray matter ROIs were smaller.

When we compared patients with SAH with mild EBI (HH 1–3) versus severe EBI (HH 4–5), the average median whole-brain ADC values were similar between groups (808 × 10⁻⁶ mm²/s for HH 1–3 versus 811 × 10⁻⁶ mm²/s for HH 4–5, P = .514). However, the histogram peak for those with mild EBI

jectors, and lesions due to the presence of blood products were manually excluded from analysis.

Statistical Plan
Statistical analyses were performed by using the JMP 9.0 software package (SAS Institute, Cary, North Carolina; 2010). Univariate comparisons between patients with SAH and controls and between patients with SAH with severe (HH 4–5) versus mild (HH 1–3) EBI were performed by using the χ² test for categoric variables and the Wilcoxon rank sum test for continuous variables. Multisample comparisons were performed by using the Kruskal-Wallis test. For average median ADC comparisons between patients with SAH and controls with whole-brain and combined ROI measurements (cerebral white matter, cerebral gray matter, white matter tracts, deep gray matter, and summed ROIs), multivariable logistic regression models adjusted for patient age were also performed because there were significant age differences between patients with SAH and controls. Correlations between ADC values and age, the modified Fisher score, the presence of elevated intracranial pressure, and external ventricular drainage were evaluated by using 2-sided Spearman correlation coefficients.
was centered at $780 \times 10^{-6}$ mm$^2$/s, while the histogram peak for those with severe EBI was centered at $800 \times 10^{-6}$ mm$^2$/s, resulting in the distributions shown in Fig 2. Across most individual ROIs, the average median ADC intensity was greater in patients with SAH with severe EBI compared with those with mild EBI (Table 3), though the results were not statistically significant. Similarly, there was no significant difference in average median ADC values for summed areas of cerebral gray matter ($P = .159$), cerebral white matter ($P = .147$), deep gray matter ($P = .074$), white matter tracts ($P = .916$), and total summed ROIs ($P = .200$) between patients with SAH with mild-versus-severe EBI.

ADC values increased significantly with age in the combined cohort of controls and patients with SAH (Pearson correlation coefficient for median summed ROI ADC and age = 0.514, $P < .001$), though this effect was primarily driven by patients with SAH (Pearson correlation coefficient = 0.602, $P < .001$), while there was no correlation between age and ADC values among controls alone (Pearson correlation coefficient = 0.012, $P = .924$). Among patients with SAH, ADC values increased as the modified Fisher score increased (median summed ROI modified Fisher score 1: 766 $\times 10^{-6}$ mm$^2$/s; modified Fisher score 2: 754 $\times 10^{-6}$ mm$^2$/s; modified Fisher score 3: 797 $\times 10^{-6}$ mm$^2$/s; modified Fisher score 4: 813 $\times 10^{-6}$ mm$^2$/s; Pearson correlation coefficient = 0.440, $P = .010$). There was no correlation between ADC values and sex, elevated intracranial pressure, or external ventricular drainage.

In multivariable analysis, after we adjusted for age, whole-brain ADC values were still significantly higher in patients with SAH compared with controls ($P = .028$). Similarly, after we adjusted for age, the average median ADC values remained higher in grouped ROIs for cerebral white matter ($P = .023$), white matter tracts ($P < .001$), deep gray matter ($P = .016$), and summed ROIs ($P < .001$, Table 2).

Evidence of cytotoxic edema was observed in 13 (39%) patients with SAH. Cytotoxic edema was more prevalent in patients with SAH with more severe EBI (87.5% of patients with HH 4–5 versus 24.0% of patients with HH 1–3, $P = .001$). The median volume of ischemia was also higher in patients with HH 4–5 than in those with HH 1–3 (0.8 mL; range, 0–88.8 versus 0 mL; range, 0–2.4; $P = .001$). The average median ADC value of ischemic lesions across all patients with SAH was $605 \times 10^{-6}$ mm$^2$/s. There was no significant difference in ischemic ADC values between patients with mild and severe EBI ($602 \times 10^{-6}$ mm$^2$/s for HH 1–3 versus $619 \times 10^{-6}$ mm$^2$/s for HH 4–5, $P = .721$).

**DISCUSSION**

This study is one of the first to provide quantitative evidence of both early cytotoxic and vasogenic edema occurring within 72 hours of SAH compared with controls. On the basis of our previous work, we anticipated globally decreased ADC values in patients with SAH compared with controls because we hypothesized that the abrupt increase in intracranial pressure that follows aneurysm rupture would result in inadequate cerebral blood flow and subsequent ischemia, similar to cardiac arrest models. Indeed, animal studies demonstrate increased intracranial pressure followed by cerebral blood flow values below the infarction threshold immediately after SAH onset. Other mechanisms may concomitantly reduce cerebral blood flow, such as primary SAH-induced vasoconstriction, as shown in preclinical

---

**Table 2:** Average median ADC values in ROIs in patients with SAH and controls

<table>
<thead>
<tr>
<th>ROI</th>
<th>ROI</th>
<th>Control (n = 66)</th>
<th>SAH (n = 33)</th>
<th>$P$ Value</th>
<th>Age-Adjusted $P$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebral gray matter</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>803 (36.2)</td>
<td>826 (37.7)</td>
<td>.007</td>
<td>.059</td>
</tr>
<tr>
<td>Cerebral white matter</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>802 (65.2)</td>
<td>823 (66.4)</td>
<td>.137</td>
<td></td>
</tr>
<tr>
<td>Parietal</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>820 (48.0)</td>
<td>843 (53.5)</td>
<td>.025b</td>
<td></td>
</tr>
<tr>
<td>Occipital</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>798 (67.7)</td>
<td>814 (51.7)</td>
<td>.042b</td>
<td></td>
</tr>
<tr>
<td>Temporal</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>791 (43.6)</td>
<td>822 (60.6)</td>
<td>.006b</td>
<td></td>
</tr>
<tr>
<td>Cerebral white matter</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>758 (33.9)</td>
<td>793 (44.9)</td>
<td>.003b</td>
<td>.023b</td>
</tr>
<tr>
<td>Frontal</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>735 (53.0)</td>
<td>768 (55.8)</td>
<td>.007b</td>
<td></td>
</tr>
<tr>
<td>Parietal</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>767 (50.9)</td>
<td>799 (51.3)</td>
<td>.007b</td>
<td></td>
</tr>
<tr>
<td>Occipital</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>764 (36.8)</td>
<td>795 (47.7)</td>
<td>.003b</td>
<td></td>
</tr>
<tr>
<td>Temporal</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>767 (45.4)</td>
<td>809 (61.8)</td>
<td>.003b</td>
<td></td>
</tr>
<tr>
<td>White matter tracts</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>739 (33.5)</td>
<td>797 (59.3)</td>
<td>&lt;.000b</td>
<td>&lt;.000b</td>
</tr>
<tr>
<td>Genu</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>761 (45.2)</td>
<td>816 (69.1)</td>
<td>&lt;.000b</td>
<td></td>
</tr>
<tr>
<td>Splenium</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>736 (45.6)</td>
<td>789 (91.4)</td>
<td>&lt;.000b</td>
<td></td>
</tr>
<tr>
<td>Internal capsule</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>719 (53.2)</td>
<td>785 (96.4)</td>
<td>&lt;.000b</td>
<td></td>
</tr>
<tr>
<td>Deep gray matter</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>713 (27.4)</td>
<td>754 (59.6)</td>
<td>&lt;.000b</td>
<td>.016b</td>
</tr>
<tr>
<td>Thalamus</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>730 (45.1)</td>
<td>777 (86.8)</td>
<td>.020b</td>
<td></td>
</tr>
<tr>
<td>Caudate</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>705 (37.2)</td>
<td>739 (65.7)</td>
<td>.002b</td>
<td></td>
</tr>
<tr>
<td>Lentiform</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>705 (45.5)</td>
<td>747 (67.5)</td>
<td>&lt;.000b</td>
<td></td>
</tr>
<tr>
<td>Hippocampus</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>854 (67.7)</td>
<td>878 (64.8)</td>
<td>.072</td>
<td></td>
</tr>
<tr>
<td>Midbrain</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>795 (93.5)</td>
<td>813 (47.4)</td>
<td>.006b</td>
<td></td>
</tr>
<tr>
<td>Pons</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>706 (50.6)</td>
<td>756 (54.6)</td>
<td>&lt;.000b</td>
<td></td>
</tr>
<tr>
<td>Cerebellar gray matter</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>756 (48.9)</td>
<td>796 (67.7)</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Cerebellar white matter</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>720 (50.3)</td>
<td>731 (43.7)</td>
<td>.214</td>
<td></td>
</tr>
<tr>
<td>Summed ROIs</td>
<td>ADC ($\times 10^{-6}$ mm$^2$/s)</td>
<td>759 (23.3)</td>
<td>795 (36.9)</td>
<td>&lt;.000b</td>
<td>&lt;.000b</td>
</tr>
</tbody>
</table>

*Values are reported as average median with SD.

Statistical significance.
While we did observe evidence of cytotoxic edema (reduced ADC values, which may represent ischemia) in 87.5% of patients with SAH with severe EBI, the volume of cytotoxic edema was too small to account for the clinical differences between those with severe-versus-mild EBI.

Instead, we noted diffusely increased ADC values in normal-appearing brain parenchyma in patients with SAH compared with controls and relatively higher ADC values in patients with SAH with more severe EBI. These data suggest the early presence of diffuse vasogenic edema, which is a direct consequence of abnormal blood-brain barrier permeability. Animal models of SAH demonstrate cerebral microvessel endothelial basal lamina disruption, acute loss of collagen IV, and increased microvessel permeability within minutes of SAH onset. Resultant blood-brain barrier breakdown and inflammation are evidenced by the escape of platelet aggregates and neutrophils into the brain parenchyma.

Several other clinical studies have identified an increase in immunologic and inflammatory markers in patients with both unruptured intracranial aneurysms and subarachnoid hemorrhage. The association of higher ADC values with higher modified Fisher scores (and hence more intracranial hemorrhage) in our study suggests that the presence of blood products may play a role in the development of vasogenic edema.

Other MR imaging studies performed in patients with SAH corroborate our findings. In one study, the volume of DWI and FLAIR hyperintensities (assessed through manual ROI segmentation and documented within 48 hours of SAH ictus) was found to correlate with worse admission neurologic status and poor 3-month functional outcomes. However, much of the FLAIR signal changes in this study were related to chronic injury in addition to the evolution of DWI positive for ischemic lesions and transependymal edema, possibly due to hydrocephalus. Our data do not demonstrate a significant association of intracranial pressure or external ventricular drain use and ADC values, suggesting that hydrocephalus and transependymal edema are not the causes of higher ADC measures. Additionally, elevated ADC values occurred in ROIs that are not immediately juxtaposed to the ventricles, where transependymal edema occurs.

In another analysis, 100 patients with aneurysmal SAH underwent DWI evaluation an average of 9 days after bleed onset. In this study, normal-appearing brain parenchyma had significantly higher ADC in the cerebral white matter and deep gray matter compared with healthy age-matched volunteers, while cerebral gray matter did not significantly differ between groups. Although evidence of radiographic vasospasm was not described, symptomatic vasospasm was observed in 14% of patients in this cohort. The authors also found that age was significantly associated with higher ADC values, but admission neurologic status (Hunt and Hess grades 1–2 versus 3–4, no subjects with Hunt and Hess 5 scores), treatment (surgical clipping versus endovascular coiling), and the severity of bleeding (modified Fisher grades 1–2 versus 3–4) were not significantly associated with ADC values. Limitations of this study include MRIs being performed during the vasospasm period (typically days 3–14 following aneurysm rupture) and signal change possibly representing pathology related to vasospasm rather than EBI due to aneurysm rupture.

In contrast to the above study, in which increases in ADC values were observed predominantly in the white matter and deep gray matter, we observed increases diffusely, including in the cortical gray matter, brain stem, and cerebellum. The timing of MR imaging in our study may account, in part, for this difference. Elevated ADC values have also been observed in other types of neurologic injury. In a study of traumatic brain injury, higher

---

**Table 3: Average median ADC values in ROIs among patients with SAH with HH 1–3 versus 4–5**

<table>
<thead>
<tr>
<th>ROI</th>
<th>ADC (×10⁻⁶ mm²/s)</th>
<th>ADC (×10⁻⁶ mm²/s)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebral gray matter</td>
<td>823 (25.9)</td>
<td>833 (64.1)</td>
<td>.159</td>
</tr>
<tr>
<td>Frontal</td>
<td>821 (60.9)</td>
<td>831 (85.9)</td>
<td>.785</td>
</tr>
<tr>
<td>Parietal</td>
<td>844 (55.5)</td>
<td>843 (50.6)</td>
<td>.556</td>
</tr>
<tr>
<td>Occipital</td>
<td>813 (41.0)</td>
<td>819 (80.3)</td>
<td>.172</td>
</tr>
<tr>
<td>Temporal</td>
<td>816 (52.1)</td>
<td>840 (83.5)</td>
<td>.147</td>
</tr>
<tr>
<td>Cerebral white matter</td>
<td>787 (43.1)</td>
<td>811 (48.7)</td>
<td>.147</td>
</tr>
<tr>
<td>Frontal</td>
<td>764 (57.2)</td>
<td>781 (52.6)</td>
<td>.425</td>
</tr>
<tr>
<td>Parietal</td>
<td>801 (49.8)</td>
<td>792 (58.7)</td>
<td>1.000</td>
</tr>
<tr>
<td>Occipital</td>
<td>789 (45.8)</td>
<td>814 (51.6)</td>
<td>.115</td>
</tr>
<tr>
<td>Temporal</td>
<td>793.5 (55.4)</td>
<td>856 (59.7)</td>
<td>.014</td>
</tr>
<tr>
<td>White matter tracts</td>
<td>796 (58.4)</td>
<td>798 (66.1)</td>
<td>.916</td>
</tr>
<tr>
<td>Genu</td>
<td>810 (63.5)</td>
<td>834 (86.6)</td>
<td>.401</td>
</tr>
<tr>
<td>Splenium</td>
<td>799 (79.1)</td>
<td>758 (123.8)</td>
<td>.950</td>
</tr>
<tr>
<td>Internal capsule</td>
<td>780 (103.7)</td>
<td>801 (71.9)</td>
<td>.401</td>
</tr>
<tr>
<td>Deep gray matter</td>
<td>748 (52.1)</td>
<td>773 (80.0)</td>
<td>.074</td>
</tr>
<tr>
<td>Thalamus</td>
<td>763 (81.4)</td>
<td>820 (94.7)</td>
<td>.097</td>
</tr>
<tr>
<td>Caudate</td>
<td>737 (57.8)</td>
<td>743 (90.8)</td>
<td>.130</td>
</tr>
<tr>
<td>Lentiform</td>
<td>744 (64.4)</td>
<td>756 (80.7)</td>
<td>.239</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>881 (62.9)</td>
<td>870 (74.4)</td>
<td>1.000</td>
</tr>
<tr>
<td>Midbrain</td>
<td>811 (43.2)</td>
<td>818 (61.9)</td>
<td>.378</td>
</tr>
<tr>
<td>Pons</td>
<td>756 (50.3)</td>
<td>756 (70.4)</td>
<td>.834</td>
</tr>
<tr>
<td>Cerebellar gray matter</td>
<td>794 (75.5)</td>
<td>800 (36.8)</td>
<td>.324</td>
</tr>
<tr>
<td>Cerebellar white matter</td>
<td>730 (45.6)</td>
<td>732 (39.5)</td>
<td>.883</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ROI</th>
<th>ADC (×10⁻⁶ mm²/s)</th>
<th>ADC (×10⁻⁶ mm²/s)</th>
<th>P Value</th>
</tr>
</thead>
</table>

Values are reported as average median with SDs.

Statistical significance.
whole-brain peak ADC in normal-appearing parenchyma was significantly associated with worse Glasgow Coma Scores.26 While we found a trend toward higher ADC values in patients with SAH with worse admission neurologic status, our study was underpowered to detect a significant difference.

Although elevated ADC is typically due to increased extracellular water content, breakdown of the blood-brain barrier, and vasogenic edema, a 1 additional possibility includes the pseudonormalization of ADC values during the subacute stages of ischemia. While there is some evidence of ADC heterogeneity early after acute ischemia,27 the overall time course of ischemic ADC changes involves 2 well-defined phases: an early reduction phase lasting between 96 and 144 hours, followed by a late elevation phase in which ADC values return to normal and even exceed normative values.28 Our study investigated an earlier time period of SAH, notably the acute phase lasting the first 72 hours. While we observed evidence of early brain ischemia in patients with SAH, ischemic lesions were often punctate and randomly distributed. The low volumes of ischemia, the presence of elevated ADC in patients with SAH without evidence of acute ischemia, and a short timeframe from ictus to imaging are inconsistent with pseudonormalization and instead are strongly supportive of the occurrence of global vasogenic edema early after SAH.

Some limitations of our study should be mentioned. Because this was a retrospective study, there is a patient selection bias. MR imaging is not routinely performed at our institution following SAH, and patients who underwent MR imaging may differ from those who did not. Furthermore, because MR imaging was performed as part of routine clinical practice, we were unable to account for possible subtle differences in patient movement or the effects of magnetic susceptibility, which may influence ADC values. Another limitation was that though the control group had normal MR imaging findings, these patients were not healthy controls but rather patients with a clinical symptom that prompted an MR imaging evaluation. It was necessary to select a control group that underwent imaging on the same inpatient MR imaging scanner to limit technical confounders. Hence, it was more expedient to use a patient group with normal-appearing MR imaging findings (based on neuroradiologist conventional assessment) rather than healthy controls who would be imaged on a different outpatient MR imaging scanner. We performed a 2:1 control to patients with SAH analysis to maximize our statistical power to detect a difference between control and SAH ADC values. Due to the low number of patients with SAH, we were unable to detect a significant association of ADC values with evidence of EBI or worse admission neurologic status. Although we did observe trends for higher ADC values in patients with HH 4–5, a larger study would be necessary to confirm this effect. Finally, our patients with SAH were older than the controls. Previous studies have demonstrated increases in ADC values with age. However, even after controlling for age, differences in ADC values remained significant.

CONCLUSIONS

We observed evidence of early cytotoxic and vasogenic edema occurring within 72 hours of SAH ictus. Patients with SAH had significantly higher whole-brain and ROI ADC values than controls. Those with more EBI had significantly more cytotoxic edema and trended toward higher ADC values (more vasogenic edema) than those with better admission neurologic status. Both cytotoxic and vasogenic edema may be related to the mechanism of EBI following SAH or may represent biomarkers for the severity of EBI. Further study is necessary to determine whether either MR imaging marker is a significant predictor of functional outcome.

REFERENCES

17. Veelken JA, Laing RJ, Jakubowski J. The Sheffield model of sub-
arachnoid hemorrhage in rats. Stroke 1995;26:1279–84; discussion 1284 CrossRef Medline


Progressive Cortical Neuronal Damage and Extracranial-Intracranial Bypass Surgery in Patients with Misery Perfusion

H. Yamauchi, S. Kagawa, Y. Kishibe, M. Takahashi, and T. Higashi

ABSTRACT

BACKGROUND AND PURPOSE: Misery perfusion may cause selective neuronal damage in atherosclerotic ICA or MCA disease. Bypass surgery can improve misery perfusion and may prevent neuronal damage. The purpose of this retrospective study was to determine whether progression of cortical neuronal damage in surgically treated patients with misery perfusion is larger than that in surgically treated patients without misery perfusion or medically treated patients with misery perfusion.

MATERIALS AND METHODS: We evaluated the distribution of benzodiazepine receptors twice by using PET and 11C-labeled flumazenil in 18 surgically treated patients with atherosclerotic ICA or MCA disease (9 with misery perfusion and 9 without) and no perioperative stroke before and after bypass surgery; in 8 medically treated patients with misery perfusion and no intervening ischemic event; and in 7 healthy controls. We quantified abnormal decreases in the benzodiazepine receptors of the cerebral cortex within the MCA distribution and compared changes in the benzodiazepine receptor index among the 3 groups.

RESULTS: The change in the benzodiazepine receptor index in surgically treated patients with misery perfusion (27.5 ± 15.6) during 7 ± 5 months was significantly larger than that in surgically treated patients without misery perfusion (5.2 ± 9.4) during 6 ± 4 months (P < .001) and in medically treated patients with misery perfusion (3.2 ± 15.4) during 16 ± 6 months (P < .01).

CONCLUSIONS: Progression of cortical neuronal damage in surgically treated patients with misery perfusion and no perioperative stroke may occur and may be larger than that in medically treated patients with misery perfusion and no intervening ischemic event.

ABBREVIATIONS: BZR = benzodiazepine receptor; CMRO2 = cerebral metabolic rate of oxygen; FMZ = flumazenil; FMZ-BP = flumazenil-binding potential; MP = misery perfusion; OEF = oxygen extraction fraction

Chronic hemodynamic impairment, as indicated by an increased oxygen extraction fraction (OEF; misery perfusion [MP]) on PET, is a risk factor for subsequent ischemic stroke in patients with atherosclerotic ICA or MCA occlusive diseases. Furthermore, MP is thought to cause cognitive impairment independent of infarction. Extracranial-intracranial bypass surgery has been demonstrated to improve MP. However, whether bypass surgery improves stroke risk or cognitive impairment in patients with MP is controversial.

The Carotid Occlusion Surgery Study of patients with recently symptomatic ICA occlusion and MP failed to show that bypass surgery compared with medical therapy reduced the risk of ipsilateral ischemic stroke at 2 years. Furthermore, the Randomized Evaluation of Carotid Occlusion and Neurocognition trial, an ancillary study of the Carotid Occlusion Surgery Study, showed that bypass compared with no bypass did not lead to improved cognitive function after 2 years in patients with no recurrent stroke. However, in the Carotid Occlusion Surgery Study, the perioperative stroke rate was sufficiently high to nullify any benefit. Therefore, the study was criticized for the possibility that patients with MP in the bypass group may also have experienced deleterious effects on cognition (ie, silent stroke), which overshadowed any beneficial effects of improved cerebral perfusion. However, no study has investigated silent stroke after bypass surgery in patients with MP, to our knowledge.
MP may cause not only cerebral infarction but also selective neuronal damage in the cerebral cortex that is not detectable as an infarction on CT or MR imaging. Because most cortical neurons express central-type benzodiazepine receptors (BZRs), specific imaging of these receptors has made possible the in vivo visualization of neuronal alterations induced by ischemia. Selective neuronal damage can be detected in humans by using PET and C-labeled flumazenil, a neuronal tracer. Bypass surgery can improve MP and may prevent neuronal damage. Conversely, surgery conveys a risk of selective neuronal damage. However, no study has investigated changes of BZRs after bypass surgery in patients with or without MP. Carotid endarterectomy for patients with chronic hemodynamic impairment was found to result in a decrease in BZRs that was correlated with postoperative cognitive impairment associated with cerebral hyperperfusion or ischemia. If this result is also the case in bypass surgery, the progression of neuronal damage in surgically treated patients with MP should be compared with that in medically treated patients with MP, in order for the procedure to prevent neuronal damage.

The purpose of this retrospective study was to determine whether the progression of cortical neuronal damage, evaluated as a decrease in BZRs, in surgically treated patients with MP is larger than that in surgically treated patients without MP or medically treated patients with MP.

MATERIALS AND METHODS

Patients

We retrospectively analyzed data collected from 18 surgically treated patients and 8 medically treated patients enrolled in an observational study that investigated the relationship between hemodynamic compromise and selective neuronal damage in patients with atherosclerotic occlusive disease of the ICA or MCA. Patients were referred to our PET unit from 2002 to 2013 for evaluation of the hemodynamic effects of ICA or MCA disease as part of a comprehensive clinical evaluation to determine the necessity of vascular reconstruction surgery.

Using PET before and after bypass surgery (anastomosis of superficial temporal artery branch to a MCA cortical branch), we studied 18 patients (12 men and 6 women; mean age, 59 ± 10 years) with atherosclerotic occlusion or stenosis of the ICA or MCA (Table 1). The interval between baseline and follow-up PET studies ranged from 2 to 16 months (mean, 6 ± 4 months). The interval between bypass surgery and follow-up PET ranged from 1 to 14 months (mean, 5 ± 4 months). Four patients with MP and 4 patients without MP were part of a previously published dataset.

Inclusion criteria for the present study were as follows: 1) occlusion of the extracranial ICA or occlusion or stenosis (>50% diameter reduction) of the intracranial ICA or MCA as documented by conventional or MR angiography; 2) functional independence in daily life (a modified Rankin Scale score of <3); 3) surgically treated patients with no intervening stroke since the first PET examination; 4) availability and willingness of the patients to return for a PET examination after surgery; and 5) for symptomatic patients, a history of TIA or minor completed stroke in the ICA or MCA distribution. TIA was defined as focal symptoms of presumed ischemic cerebrovascular origin lasting <24 hours. Exclusion criteria were the following: 1) cerebral-cortical, cerebellar, or brain stem infarct detectable on routine MR imaging (T1WI, T2WI, or FLAIR) or CT, 2) unilateral arterial disease with extensive white matter lesions in both hemispheres probably caused by bilateral small-vessel disease, 3) history of taking BZR agonists, and 4) the presence of potential sources of cardiogenic embolism.

Follow-up PET examinations were performed in 8 medically treated patients with MP at baseline (Table 1). They were selected from a cohort of a previously published follow-up PET study for 80 medically treated patients with stenosis or occlusion of the ICA or MCA and no intervening TIA or stroke. Criteria for selection were as follows: 1) occlusion of the extracranial ICA or occlusion or stenosis (>50% diameter reduction) of the intracranial ICA or MCA, and 2) MP at the baseline PET study. The interval between the first and follow-up PET studies ranged from 7 to 26 months (mean, 16 ± 6 months).

For vascular risk factors, the status of hypertension, diabetes mellitus, ischemic heart disease, hypercholesterolemia, and smoking was evaluated from patient histories recorded at the first PET examination. Hypertension, diabetes mellitus, ischemic heart disease, or hypercholesterolemia was judged to be present when there was a history of treatment.

To establish a control database for BZR imaging, we studied 10 healthy control subjects (7 men and 3 women; mean age, 57 ± 7 years) with no previous history of a medical or psychiatric disorder or of taking BZR agonists. Among them, 7 subjects (mean age, 56 ± 8 years, including 4 men and 3 women) underwent follow-up PET examinations. The interval between the first and follow-up PET studies ranged from 38 to 45 months (mean, 41 ± 3 months). All protocols in this study were approved by the ethics committee.

### Table 1: Patient characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Surgically Treated, Misery Perfusion</th>
<th>Medically Treated</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Interval (mean) (mo)</td>
<td>7 ± 5</td>
<td>6 ± 4</td>
</tr>
<tr>
<td>Interval from bypass to PET</td>
<td>6 ± 4</td>
<td>4 ± 3</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>59 ± 9</td>
<td>59 ± 13</td>
</tr>
<tr>
<td>Male sex (No.)</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Symptomatic (No.)</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Cerebral ischemic lesions (No.)</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Qualifying artery (No.)</td>
<td>ICA (occlusion/stenosis)</td>
<td>8 (8/0)</td>
</tr>
<tr>
<td></td>
<td>MCA (occlusion/stenosis)</td>
<td>1 (0/1)</td>
</tr>
<tr>
<td>Other medical illness (No.)</td>
<td>6</td>
<td>3 (3/0)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Ischemic heart disease</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Hypercholesterolemia</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Smoking habit (current and former) (No.)</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Antiplalet agents</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Statins</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Postbypass:</td>
<td>TIA</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Cerebral ischemic lesion</td>
<td>2</td>
</tr>
</tbody>
</table>

Note: NA indicates not applicable. *P < 0.05 vs both surgical groups.
committee of Shiga Medical Center, and all subjects gave written informed consent.

**PET Measurements**

PET scans were performed for each subject by using a whole-body PET scanner, GE Advance (GE Healthcare, Milwaukee, Wisconsin), which permits the simultaneous acquisition of 35 image sections with an intersection spacing of 4.25 mm. After a transmission scan by germanium-68/gallium-68, a series of $^{15}$O-gas studies was performed. Briefly, $^{15}$O$_2$ and $^{15}$O were delivered continuously to the patient via a mask for the duration of a 5-minute scan. CBV was measured by bolus inhalation of $^{15}$O with scanning for 3 minutes. Arterial samples were obtained during scanning. No subject showed substantial changes in the partial pressure of carbon dioxide in arterial blood during scanning.

The $^{15}$O-gas study was followed by a study of $^{11}$C-flumazenil, which was synthesized by $^{11}$C-methylation of demethylated-flumazenil (FMZ) (Hoffmann-La Roche, Basel, Switzerland). After the slow intravenous injection of $^{11}$C-FMZ, a 50-minute dynamic PET scan was initiated.

We used the steady-state method to calculate CBF, the cerebral metabolic rate of oxygen (CMRO$_2$), and OEF. The CMRO$_2$ and OEF were corrected on the basis of CBV. The binding potential (nondisplaceable) of $^{11}$C-flumazenil was calculated by using dynamic data and Logan graphic analysis with reference tissue, with the pons as the reference region.

**Data Analysis**

For $^{15}$O-gas PET scanning analysis, we used a classic ROI analysis. We analyzed 10 tomographic planes, located 46.25–84.5 mm above and parallel to the orbitomeatal line. The lowest plane corresponded to the level of the basal ganglia and the thalamus, and the uppermost plane corresponded to the level of the centrum semiovale. An ROI was selected for CBF images. Each image was examined by compactly placing 10–12 circular ROIs (diameter, 16 mm) over the gray matter of the outer cortex in each hemisphere. According to the atlas, the ROIs in all 10 images covered the distribution of the MCA and the external borderzone regions. The same ROIs were used for the CMRO$_2$, OEF, and CBV images. The mean hemispheric value for the hemisphere affected by ICA or MCA disease was calculated as the average of all the circular ROIs.

Normal control values of the PET variables were obtained from 7 healthy volunteers (4 men and 3 women; mean age, 47 ± 7 years) who underwent routine neurologic examinations and MR imaging. The mean OEF value in the 14 control hemispheres was 44.5% ± 3.8%. Hemispheric OEF values beyond the upper 95% limit defined in healthy subjects (>52.5%) were considered to represent increased OEF. Comparative values for CBF and CBF/CBV in healthy volunteers were 44.6 ± 4.5 mL/100 g/min and 11.4 ± 1.8/min, respectively. Hemispheric CBF and CBV values below 35.0 mL/100 g/min and 7.6/min, respectively, were considered abnormal. Patients with increased OEF, decreased CBF, and decreased CBF/CBV in hemispheres with arterial disease were categorized as having MP. One investigator unaware of the clinical status categorized the patients.

Flumazenil-binding potential parametric images were analyzed by using a 3D-stereotactic surface-projection technique, as previously described. This technique anatomically normalizes individual PET data to a standard brain and compares regional voxel data between patients and controls. In the standard stereotactic system, pixels located on the outer and medial surfaces of both hemispheres and vectors perpendicular to the 3D surface at each pixel are predetermined. For each predetermined surface pixel on an individual’s anatomically standardized PET image set, the algorithm searches along the vector, 6 pixels deep into the cortex, for the highest pixel value and assigns this maximum value to the surface pixel. To correct for fluctuations in whole-brain values and to extract the changes due to ICA or MCA disease, we normalized the pixel values of an individual’s image set to the mean cerebellar value before analysis. Z scores were calculated for each surface pixel as (Mean Normalized Pixel Value for Controls − Normalized Pixel Value for the Patient) / (SD for controls), and were used to quantify decreases in flumazenil-binding potential (FMZ-BP). Thus, a positive z score in a patient represented reduced FMZ-BP relative to the control group. An increased BZR index corresponded to a decreased BZR level, a decreased FMZ-BP level, and thus greater cortical neural damage.

To quantify the degree of abnormal flumazenil-binding potential reduction in each patient, we used the stereotactic estimation method to calculate a BZR index, defined as (% Pixels with z Score >2 × Average z score for those pixels) for the cerebral-cortical MCA distribution affected by ICA or MCA disease. This MCA distribution included the middle and inferior frontal gyri; the precentral gyrus; the superior and inferior parietal gyri; the angular, postcentral, and supramarginal gyri; the superior, middle, inferior, and transverse temporal gyri; and the superior and middle occipital gyri.

At postsurgical or follow-up examinations, total change in the BZR index or the CBF, CMRO$_2$, OEF, CBV, or CBF/CBV values in the MCA distribution with arterial disease was calculated by subtracting the values obtained at the second examination from those obtained at the first examination. In controls, the calculation was performed by using the mean of the bilateral hemispheric values of the BZR index. The mean value of changes in the index in the controls was 0.94 ± 1.38. In patients, an increase of the index beyond the upper 95% limit (the mean plus 60.05 × SD; 60.05, the value of t for P < 0.05 and degree of freedom = 6) defined in healthy subjects (>4.32) was considered an increased BZR index (progression of neuronal damage) during follow-up.

**Statistical Analysis**

The statistical analysis was performed by using StatView (SAS Institute, Cary, North Carolina). Comparisons of clinical backgrounds or PET values among the 3 groups were performed with 1-way ANOVA and a post hoc Scheffe analyses or Kruskal-Wallis tests and post hoc Mann-Whitney U tests as appropriate. PET variable values were compared between examinations by using paired t tests. Relationships between variables were analyzed by using simple or multiple regression analyses. A multivariable linear regression model with a forward stepwise selection procedure was used to test the independent predictive value of the presence of MP at baseline and the change in PET variables during fol-
An increased BZR index corresponds to a decreased BZR level, a decrease in CBF, and an increase in OEF (coefficient, 1.2; standard error, 0.5; t = 2.5; P < .05), and smoking habit (coefficient, 13.3; standard error, 4.7; t = 2.7; P < .05) were positively correlated with changes in the BZR index.

Eight medically treated patients with MP showed a longer interval (P < .05) between the 2 PET examinations than 9 surgically treated patients with MP (Table 1). Other patient characteristics significantly differed between the patients with MP and those without it (Table 1). The value of the BZR index at baseline was not different between the 2 groups (Table 2).

After the operation, patients with MP showed a significant increase in the BZR index (paired t test, P < .001), CBF (P < .01), and CBF/CBV (P < .05) and a tendency toward decreases in the OEF (P = .06), while patients without MP did not show significant changes in any PET variables (Table 2 and Fig 1). The BZR index in the MP group tended to be larger (P = .07) than that in the no-MP group after the operation.

The change in the BZR index in the MP group was significantly larger than that in the no-MP group (P < .001). Eight patients (89%) with MP showed an increase in the BZR index beyond the upper 95% limit defined in healthy subjects, while no patients without MP (0%) showed increases in the BZR index (Fisher exact test, P = .0004). The change in the BZR index was significantly correlated with the value of CBF (r = −0.66, P < .005) or OEF (r = 0.47, P < .05) at baseline (Fig 2).

A multivariable linear regression analysis (forward stepwise selection) was used to investigate the association of changes in the BZR index with the following: 1) the presence of MP at baseline; 2) changes in the values of CBF, CMRO₂, OEF, CBV, or CBF/CBV after bypass; 3) the presence of vascular risk factors; and 4) drug-treatment history (Table 1). Our analysis produced a model that included the presence of MP at baseline, changes in the OEF after surgery, and smoking habit, with a correlation coefficient of 0.905 for the changes in the BZR index after surgery (P < .001). In our model, the presence of MP, changes in the OEF, and smoking habit accounted for 64.5%, 8.2%, and 9.2% of the variance in changes in the BZR index, respectively. The presence of MP (coefficient, 35.2; standard error, 5.2; t = 6.7; P < .0001), changes in the OEF (coefficient, 1.2; standard error, 0.5; t = 2.5; P < .05), and smoking habit (coefficient, 13.3; standard error, 4.7; t = 2.7; P < .05) were positively correlated with changes in the BZR index.

Eight medically treated patients with MP showed a longer interval (P < .05) between the 2 PET examinations than 9 surgically treated patients with MP (Table 1). Other patient characteristics were not significantly different between the medically and surgically treated patients with MP and those without it (Table 1).
hemisphere with arterial disease. The row of arterial disease is reported to be high. Patients with MP have a marginally adequate blood supply relative to metabolic demand, which increases the risk of cerebral ischemia. Therefore, it is reasonable to hypothesize that fluctuations of cerebral hemodynamics during perioperative periods might have caused ischemic neuronal damage in surgically treated patients with MP. Additionally, the presence of MP is reported to be a risk factor for postoperative hyperperfusion after bypass. Cerebral hyperperfusion after carotid endarterectomy was found to result in a decrease in BZRs. To prevent perioperative neuronal damage, therapeutic strategies that limit selective neuronal damage, including careful perioperative management and the use of neuroprotective agents, may be needed for patients with MP. Furthermore, MP may not improve completely or immediately after bypass surgery. Thus, the risk for hemodynamic ischemia may persist during follow-up.

The progression of cortical neuronal damage in surgically treated patients with MP was larger than that in medically treated patients with MP and no intervening ischemic event. Surgically treated patients with MP showed an expected increase in CBF and CBV at baseline or at follow-up were not significantly different between the 2 groups, while at follow-up, the BZR index in the medically treated patients with MP had a tendency to be larger (P = .07) than that in the medically treated patients. The change in the BZR index in surgically treated patients was significantly larger than that in medically treated patients (P < .01). Three medically treated patients (37%) and 8 surgically treated patients (89%) showed an increase in the BZR index beyond the upper 95% limit of changes in the BZR index for the 7 controls. The closed circles indicate patients with misery perfusion, and the open circles indicate patients without misery perfusion.

DISCUSSION
This study demonstrates that the progression of cortical neuronal damage manifests as a decrease in BZRs in the normal-appearing cerebral cortex of surgically treated patients with MP and no perioperative stroke. The degree of progressive cortical neuronal damage in surgically treated patients with MP was significantly larger than that in medically treated patients with MP and no intervening ischemic event.

The precise reason for the progression of cortical neuronal damage in surgically treated patients with MP and no perioperative stroke is unclear from the present study. However, perioperative ischemia, postoperative hyperperfusion, and hemodynamic ischemia during follow-up might cause cortical neuronal damage. In patients with MP, the risk for perioperative ischemic events is reported to be high. Patients with MP have a marginally adequate blood supply relative to metabolic demand, which increases the risk of cerebral ischemia. Therefore, it is reasonable to hypothesize that fluctuations of cerebral hemodynamics during perioperative periods might have caused ischemic neuronal damage in surgically treated patients with MP. Additionally, the presence of MP is reported to be a risk factor for postoperative hyperperfusion after bypass. Cerebral hyperperfusion after carotid endarterectomy was found to result in a decrease in BZRs. To prevent perioperative neuronal damage, therapeutic strategies that limit selective neuronal damage, including careful perioperative management and the use of neuroprotective agents, may be needed for patients with MP. Furthermore, MP may not improve completely or immediately after bypass surgery. Thus, the risk for hemodynamic ischemia may persist during follow-up.

The progression of cortical neuronal damage in surgically treated patients with MP was larger than that in medically treated patients with MP and no intervening ischemic event. Surgically treated patients with MP showed an expected increase in CBF and CBV at baseline or at follow-up were not significantly different between the 2 groups, while at follow-up, the BZR index in the medically treated patients with MP had a tendency to be larger (P = .07) than that in the medically treated patients. The change in the BZR index in surgically treated patients was significantly larger than that in medically treated patients (P < .01). Three medically treated patients (37%) and 8 surgically treated patients (89%) showed an increase in the BZR index beyond the upper 95% limit of changes in the BZR index for the 7 controls. The closed circles indicate patients with misery perfusion, and the open circles indicate patients without misery perfusion.
impairment to improve clinical outcomes. However, reperfusion therapies may have a risk of neuronal damage for patients with hemodynamic impairment. Carotid endarterectomy and bypass surgery for patients with chronic hemodynamic impairment may be associated with cortical neuronal damage that may be correlated with postoperative cognitive impairment. In acute stroke, selective neurological damage may affect salvaged penumbra and hamper functional recovery following reperfusion. The progression of cortical neuronal damage should be considered while treating patients with hemodynamic impairment by means of reperfusion therapy.

**Limitations**
This study had certain limitations, including a small sample size and a cohort with both ICA and MCA disease. It also retrospectively analyzed data from a prospective observational study. Therefore, there is a considerable risk of selection bias, though the clinical backgrounds were not significantly different among the 3 groups. For these reasons, the findings in this study should be confirmed in future randomized and blinded studies. The selection of bypass surgery was left to individual clinical judgment by attending physicians, and unknown clinical factors could have swayed treatment decisions. Nine patients with less marked hemodynamic impairment than is indicative of MP underwent bypass surgery due to regional increases in OEF in addition to other characteristics (poor collateral pathways or recurrent symptoms). Additionally, 2 patients with perioperative stroke were excluded from analysis. One patient with MP had progressing stroke after the first PET examination and underwent emergency bypass surgery, and another without MP showed multiple MR imaging lesions after aneurysm treatment performed in combination with bypass surgery. The interval time between baseline and follow-up PET for medically treated patients with MP was twice as long as that for surgically treated patients with MP, which may make it difficult to accurately compare patient groups. However, the longer interval may lead to the larger, not smaller, increase of cortical neuronal damage in medically treated patients, which may not change the conclusion. The period of patient recruitment ranged from 2002 to 2013, but the PET scan protocol had not changed, so it likely offered the same precision and accuracy. We could not systematically investigate cognitive differences among these patients. Although decreases in BZR are objective markers of neuronal injury that may be associated with cognitive impairment, the functional effect of neuronal damage should have been determined by neuropsychological testing.

**CONCLUSIONS**
The progression of cortical neuronal damage manifests as a decrease in BZR in the normal-appearing cerebral cortex of surgically treated patients with MP and no perioperative stroke. The degree of progressive cortical neuronal damage in surgically treated patients with MP may be larger than that in medically treated patients with MP and no intervening ischemic event. The progression of cortical neuronal damage should be considered while treating patients with MP with bypass surgery.

**REFERENCES**

Cognitive Implications of Deep Gray Matter Iron in Multiple Sclerosis

E. Fujiwara, J.A. Kmech, D. Cobzas, H. Sun, P. Seres, G. Blevins, and A.H. Wilman

ABSTRACT

BACKGROUND AND PURPOSE: Deep gray matter iron accumulation is increasingly recognized in association with multiple sclerosis and can be measured in vivo with MR imaging. The cognitive implications of this pathology are not well-understood, especially vis-à-vis deep gray matter atrophy. Our aim was to investigate the relationships between cognition and deep gray matter iron in MS by using 2 MR imaging–based iron-susceptibility measures.

MATERIALS AND METHODS: Forty patients with multiple sclerosis (relapsing-remitting, n = 16; progressive, n = 24) and 27 healthy controls were imaged at 4.7T by using the transverse relaxation rate and quantitative susceptibility mapping. The transverse relaxation rate and quantitative susceptibility mapping values and volumes (atrophy) of the caudate, putamen, globus pallidus, and thalamus were determined by multiatlas segmentation. Cognition was assessed with the Brief Repeatable Battery of Neuropsychological Tests. Relationships between cognition and deep gray matter iron were examined by hierarchic regressions.

RESULTS: Compared with controls, patients showed reduced memory (P < .001) and processing speed (P = .02) and smaller putamen (P < .001), globus pallidus (P = .002), and thalamic volumes (P < .001). Quantitative susceptibility mapping values were increased in patients compared with controls in the putamen (P = .003) and globus pallidus (P = .003). In patients only, thalamus (P < .001) and putamen (P = .04) volumes were related to cognitive performance. After we controlled for volume effects, quantitative susceptibility mapping values in the globus pallidus (P = .03; trend for transverse relaxation rate, P = .10) were still related to cognition.

CONCLUSIONS: Quantitative susceptibility mapping was more sensitive compared with the transverse relaxation rate in detecting deep gray matter iron accumulation in the current multiple sclerosis cohort. Atrophy and iron accumulation in deep gray matter both have negative but separable relationships to cognition in multiple sclerosis.

ABBREVIATIONS: DGM — deep gray matter; GP — globus pallidus; NPtotal — composite z-score across neuropsychological tests; PASAT — Paced Auditory Serial Addition Test; QSM — quantitative susceptibility mapping; R2* — transverse relaxation rate; SDMT — Symbol Digit Modalities Test; SPART — 10/36-Spatial Recall Test; SRT — Selective Reminding Task

Cognitive problems occur in 40%–65% of individuals with multiple sclerosis, predominantly affecting information processing speed and episodic memory. Subcortical atrophy, particularly in the thalamus, is well-known to predict cognitive deficits in MS. Elevated levels of iron accumulation in deep gray matter (DGM) nuclei in MS have also been reported using different iron-sensitive MR imaging measures, with studies focusing particularly on the large basal ganglia nuclei (caudate, putamen, globus pallidus [GP]), and the thalamus). Excess iron catalyzes production of free radicals, promoting neurodegeneration. This affects the DGM in both healthy aging and different CNS disorders. DGM iron accumulation in MS may be an epiphenomenon of structural atrophy caused by cell death, but others reported no relationships between DGM iron, global/regional brain volumes, or lesion load, suggesting potentially independent pathologies. The functional implications of DGM iron accumulation relative to other DGM pathologies in MS need further examination. Previous studies have examined some aspects of cognitive functions and DGM iron in MS with 4 different MR techniques.

Among the MR imaging measures used, only the gradient-echo
transverse relaxation rate (R2*) and quantitative susceptibility mapping (QSM) have been validated against postmortem iron assessment, both in non-MS12,13 and in MS populations.14,15

Three MS studies assessed different aspects of cognition along with R2*, 5,10,11 in Khalil et al,5 R2* in the basal ganglia (but not in the thalamus) was related to processing speed in patients with clinically isolated syndrome and those with MS. In Pinter et al,10 a neuropsychological composite score of cognitive efficiency/processing speed (but not memory) of patients with clinically isolated syndrome and those with MS was reported. This was predicted by R2* relaxation rates averaged across basal ganglia nuclei, along with caudate volume and T2 lesion load. Schmalbrock et al11 recently cross-examined QSM and R2* measures against performance in 2 inhibitory cognitive tasks (a Stroop Task and an Erik-sen Flanker Task) in patients with relapsing-remitting MS, imaged at 7T. Inhibition in the Flanker Task (but not the Stroop Task) was related to caudate and anterior putamen iron assessed with QSM, but performance in neither task was related to R2* measures. Thus, only 1 study11 directly compared the cognitive correlates of R2* and QSM-based iron measures in MS, but it did not control for atrophy in the same DGM regions.

The objective of our study was to determine whether cognition in MS, measured by the Brief Repeatable Battery of Neuropsychological Tests, is related to DGM iron accumulation measured with R2* and QSM at a high field strength (4.7T). The core hypothesis was that iron (R2* and QSM) in DGM nuclei correlates with decreased cognitive performance in MS, irrespective of atrophy.

MATERIALS AND METHODS
Participants
This study was approved by the local research ethics board, and all participants provided written informed consent. Forty patients diagnosed with MS were recruited from the Northern Alberta Multiple Sclerosis Clinic in Edmonton, Alberta, Canada. Sixteen patients had relapsing-remitting MS, 15 patients had secondary-progressive MS, and 9 patients were diagnosed with primary-progressive MS. Patients were at least 18 years of age and were diagnosed with MS on the basis of the 2010 McDonald criteria.16 Twenty-seven healthy controls were recruited through word of mouth, on-line (Kijiji; http://www.kijiji.ca/h-alberta/9003), and print advertising (local daily newspaper). All participants were alert, lucid, able to communicate verbally, and understood the test instructions and purpose of the testing; they had normal or corrected-to-normal vision and hearing and were fluent in English. Exclusion criteria were major neurologic or psychiatric illnesses apart from MS in the patient group (eg, stroke, encephalitis, or meningitis; head injury with loss of consciousness of >5 minutes; psychosis), diabetes, learning disabilities, and contraindications to MR imaging (pacemakers, nonremovable metal clips, major dental work, and so forth). Six participants were excluded due to MR imaging artifacts (2 patients), incomplete MR imaging data (1 control), or incomplete cognitive data (2 controls, 1 patient). The gap between MR imaging and cognitive testing was 24.48 ± 22.25 days for patients and 10.44 ± 20.01 days for controls. No patients were imaged during a time of relapse. The earliest time from re-lapse to imaging was 5.57 months (6.17 months to cognitive testing).

MR Imaging
Participants were imaged by using a 4.7T Varian Inova MR imaging system (Agilent Technologies, Santa Clara, California). Two MR imaging sequences were collected to enable quantitative measurements of volume, R2*, and QSM. A 3D longitudinal relaxation time (T1)-weighted sequence used inversion recovery rapid gradient-echo (84 sections; 2-mm thick; in-plane, 0.9 × 0.9 mm2; acquisition time, 4.8 minutes). The R2*/QSM sequence used 3D multiecho gradient-echo (80 sections; 2-mm thick; in-plane, 1× 1 mm2; acquisition time, 9.4 minutes), with TE parameters (10 echoes; first echo, 2.9 ms; echo spacing, 4.1 ms). Both R2* and QSM were reconstructed from the same multiecho images by using previously validated methods.17,18 Briefly, R2* used a 3D linear field gradient correction to compensate for air-tissue susceptibility effects and then a monoexponential fit.17 For QSM, a field map was estimated from the multiecho data, followed by background field removal by using Regularization-Enabled Sophisti-cated Harmonic Artifact Reduction for Phase data (RESHARP),18 and dipole inversion by using total variation regularization.19-21 The imaging protocol also included axial T2-weighted and FLAIR imaging, both with 4-mm section thickness, which were used to estimate lesion burden.

Neuropsychological Assessment
The Brief Repeatable Battery of Neuropsychological Tests was conducted on all participants.22 The battery includes verbal and visual memory tests: the Selective Reminding Task (SRT), the 10/36-Spatial Recall Test (SPART), information-processing speed/working memory tests (Symbol Digit Modalities Test [SDMT]), and the Paced Auditory Serial Addition Test [PASAT], 2- and 3-second versions), and a phonemic fluency test (word list generation). Administration time was approximately 25 minutes. For 3 patients with limited hand or arm movement, the test admin-strator placed the checkers in the nontimed SPART, as di-reccted by the patient. The PASAT was attempted but not completed in all participants due to noncompliance or fatigue.

Image Analysis
R2* and QSM data from each participant were rigidly aligned with the T1-weighted images and interpolated to the same resolution. Bias field-intensity normalization for T1 images was performed by using the N4 method23 as part of the Advanced Normalization Tools package (stnava.github.io/ANTs/). We segmented 4 DGM nuclei (caudate, putamen, GP, and thalamus) with a multiatlas segmentation method by means of both T1 and QSM images,24 taking advantage of the high DGM contrast available on QSM (Fig. 1), which is particularly beneficial to reliably segment the GP. Ten manually segmented volumes from healthy controls (2 controls from the current study, 8 from another ongoing study) were used as atlases. Following a standard multiatlas segmentation method,25 the atlases from the 10 healthy controls were propagated to each individual dataset by using automatic nonlinear registration on multimodal T1, R2*, and QSM data.26 Each of the registered anatomic labels propagated from the 10 atlases was fused by using a probabilistic label-fusion method proposed by Wang et al27 to produce optimal segmentation of each dataset. Using these anatomic segmentation labels, we then extracted vol-
and manual tracing\(^{24}\): caudate, 82%; putamen, 87%; globus pallidus, 83%; thalamus, 87%.

We made by the senior author, an imaging expert with 25 years of experience. Lesion measures were defined as hyperintensities on the outline of each lesion on T2-weighted images and multiplying by the section thickness. Lesion measures were more educated. Our patient cohort comprised 60% patients with progressive MS with a median Expanded Disability Status Scale\(^{28}\) score of 5.25 and disease duration of 14.9 ± 8.9 years. Eighteen patients (all patients with relapsing-remitting MS and 2 with progressive MS) were taking disease-modifying medications. Additional clinical details are outlined in the Table. After we controlled for the education differences, ANCOVAs on the Brief Re- presentations, successively adding predictors in blocks (see also Pinter et al\(^{10}\)). Models were run separately for each of the 4 DGM structures and separately in the controls and patients. For these regressions, cognitive scores were first z-transformed on the basis of the control group and then combined into a composite neuropsychological z score (NP\(_{\text{total}}\)) comprising 6 test scores: 2 SRT measures (continuous long-term retrieval, delayed recall), 2 SPART measures (immediate and delayed recall), SDMT, and word list generation. The PASAT was excluded from NP\(_{\text{total}}\) due to excessive missing data (Table). For each of the 8 models per group, the first block of predictors included age, sex, and education. The second block added linear regression models (hierarchical regressions). Results

### Statistical Analyses

We first compared patients with controls in cognitive scores and DGM parameters (volume, iron) by using ANCOVAs, controlling for age, sex, and/or education as applicable. Next, to predict cognition by the DGM parameters, we conducted nested sets of

| Participant demographics and cognitive performance\(^{a}\) |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| **Controls**    | **Patients**    | **P Value**     |
| **Age (yr)**    | 47.51 ± 10.09   | 49.08 ± 10.03   | .53             |
| **Sex**         | 18 F, 9 M       | 27 F, 13 M      | .94             |
| **Education (yr)** | 17.37 ± 3.87   | 13.66 ± 2.05    | <.001           |
| **Median Expanded Disability Status Scale score (range)** | NA               | 5.25 (2.0–9.0)  | NA              |
| **Mean disease duration** | NA              | 14.90 ± 8.91 years | NA              |
| **Disease-modifying therapy** | NA              | None: n = 22; glatiramer acetate: n = 12; interferon β-1a: n = 6 | NA              |
| **Lesion load** | NA              | 6.23 ± 8.78 mL  | NA              |
| **SRT immediate recall** | 56.15 ± 6.20    | 44.40 ± 10.77   | <.001           |
| **SRT continuous long-term retrieval** | 42.63 ± 12.12   | 23.00 ± 14.66   | <.001           |
| **SRT delayed recall** | 9.11 ± 2.10     | 6.70 ± 2.69     | <.001           |
| **SPART immediate recall** | 20.37 ± 5.10    | 20.25 ± 5.20    | .90             |
| **SPART delayed recall** | 7.11 ± 2.55     | 6.98 ± 2.39     | .76             |
| **SDMT**        | 58.74 ± 10.52   | 49.56 ± 16.08   | .02             |
| **PASAT 3-second\(^{b}\)** | 49.04 ± 8.51    | 43.94 ± 11.73   | .55             |
| **PASAT 2-second\(^c\)** | 36.52 ± 8.62    | 32.15 ± 10.78   | .49             |
| **Phonemic fluency (word list generation)** | 30.52 ± 6.94    | 27.28 ± 8.57    | .25             |

| **Note:** —NA indicates not applicable. |

\(^{a}\) Group comparisons in cognition are education-adjusted. Data are means unless otherwise indicated.

\(^{b}\) Available from 35 patients with MS and 26 controls.

\(^{c}\) Available from 34 patients with MS and 25 controls.

DGM volumes (age- and sex-corrected). The third block added either R2* or QSM measures (age- and sex-corrected).
fence (1.5 interquartile ranges), and outliers are shown by dots outside the fence. Cau = caudate; Put = putamen; GP = globus pallidus; Tha = thalamus; double asterisks = $P < .001$; asterisk = $P < .01$.

**FIG 2.** Group differences in age-, sex-, and intracranial volume-normalized deep gray matter volumes (A) and iron based on QSM (B) and $R^2*$ (C). Boxplots show ranges of the first-to-third quartiles, circles indicate means, lines inside the boxes indicate medians, dotted lines indicate the fence (1.5 interquartile ranges), and outliers are shown by dots outside the fence. Cau = caudate; Put = putamen; GP = globus pallidus; Tha = thalamus; double asterisks = $P < .001$; asterisk = $P < .01$.

**FIG 3.** Prediction of cognition (NP$_{total}$) by age, sex, and intracranial volume-normalized putamen (A) and thalamus (B) volumes in patients with MS after correcting for age, sex, and education.

**Group Differences in DGM**

Bilateral DGM volumes were correlated with age and sex in some regions; thus, age-, and sex-adjusted volumes were used. We observed volumetric reductions in patients with MS in the putamen ($P < .001$), GP ($P = .002$), and thalamus ($P < .001$) (Fig 2A). For QSM, patients showed higher values than controls in the putamen ($P = .003$) and GP ($P = .003$), with a trend in the caudate ($P = .066$) (Fig 2B). $R^2*$ increases in patients showed trends similar to those in QSM, but they were not significant (Fig 2C).

With partial correlations correcting for age and sex, measures of $R^2*$ and QSM were highly correlated with each other within each region in patients (caudate: $r = 0.66$; putamen: $r = 0.86$; GP: $r = 0.75$; thalamus: $r = 0.60$; all, $P < .001$). Except for the thalamus, these correlations were similar, albeit weaker, in the control group (caudate: $r = 0.46$, $P = .02$; putamen: $r = 0.59$, $P = .002$; GP: $r = 0.78$, $P < .001$; thalamus: $r = 0.03$, $P = .89$). Notably, $R^2*$ and QSM values were uncorrelated with structural volumes within each of the 4 DGM nuclei, both in patients and controls, permitting their use in the hierarchic regressions.

**Cognition, DGM Volumes, and Iron Susceptibility**

Regressions examined predictors of cognitive performance within each of the DGM regions separately for patients and controls. The first block of predictors included only demographic variables: age, sex, and education years. In the second block, DGM volumes were added. In the third block, QSM or $R^2*$ values were included.

Within the patient group, compared to a model with demographic variables alone ($R^2 = 0.15$, $P = .12$), putamen volumes significantly increased prediction of NP$_{total}$ ($R^2 = 0.25$, $\Delta R^2 = 0.1$, model $F = 2.89$, $P = .046$; putamen $\beta = 0.32$, $t = 2.15$, $P = .04$; Fig 3A). These results were similar but more pronounced with thalamic volumes ($R^2 = 0.39$, $\Delta R^2 = 0.24$, model $F = 5.50$, $P = .002$; thalamus $\beta = 0.49$, $t = 3.68$, $P < .001$; Fig 3B). Volumetric data did not contribute significantly to relationships with NP$_{total}$ in the healthy controls (On-line Table 1).

In a third block, QSM values were then included. In patients, after controlling for demographic influences and GP volumes, NP$_{total}$ was related to GP QSM ($R^2 = 0.29$, $\Delta R^2 = 0.1$, model $F = 2.76$, $P = .03$; GP QSM $\beta = -0.32$, $t = 2.22$, $P = .03$; Fig 4B). Thus, GP QSM explained an additional $10\%$ of variance ($\Delta R^2$) in cognitive performance, after controlling for demographic variables and for GP volume. None of the other QSM measures were substantially related to NP$_{total}$ in patients or in controls (see On-line Tables 1 and 2 for details). Analyses with $R^2*$ showed a similar but nonsignificant result for the GP only ($R^2 = 0.25$, $\Delta R^2 = 0.06$, model $F = 2.26$, $P = .071$; GP $\beta = -0.23$, $t = -1.70$, $P = .099$; Fig 4A). No other $R^2*$ measure was significantly or at trend-level related to NP$_{total}$ in patients. Controls did not show any significant relationships between iron measures and cognition, including in the GP iron models (QSM-GP: GP $\beta = -0.077$, $t = -0.44$, $P = .663$; $R^2*$-GP: GP $\beta = -0.15$, $t = -0.86$, $P = .399$).

When we inspected the NP$_{total}$ 2 tests were individually sensitive to DGM changes. Thalamic volume correlated with SPART--
delayed recall (SPART: \( \beta = 0.45, P = .003 \)) and with SDMT (\( \beta = 0.45, P = .003 \)). QSM in the GP was negatively related to phonemic fluency (\( \beta = -0.40, P = .01 \)). Bonferroni correction by 6, the number of individual cognitive tests in the NP\(_{\text{total}}\) (\( P = .05/6 = 0.008 \)), only retained the thalamic volume correlations with SDMT and SPART. Disability (Expanded Disability Status Scale) and disease duration were unrelated to any of the MR imaging measures in this cohort, and only trend-level correlated with NP\(_{\text{total}}\) (\( r = -0.37, P = .07 \)), possibly due to limited sensitivity of the Expanded Disability Status Scale in higher ranges (focusing on motor symptoms but not cognitive symptoms) and the long duration and variability in disease durations in the current cohort.

**DISCUSSION**

We examined cognitive correlates of DGM atrophy and iron accumulation in a mixed cohort of patients with progressive MS and relapsing-remitting MS. Irrespective of atrophy, cognition was negatively related to iron accumulation in the GP assessed with QSM, with a similar trend in the R\(^2\*\) measure. In turn, volumes of the thalamus and putamen were related to cognition, replicating previous findings.\(^2\) These correlations were also irrespective of iron accumulation in those regions. Within each investigated DGM structure, volumetric and iron measures were unrelated, suggesting that atrophic and iron-related pathologies in the DGM may exert separable influences on cognitive functions in MS.

Our findings extend the limited literature on relationships between DGM iron and cognition in MS. Only 2 previous studies concurrently accounted for regional\(^5\) or global atrophy\(^10\) and reported, similarly, moderate relationships between DGM iron and cognition. Modica et al\(^6\) used susceptibility-weighted imaging to assess DGM iron. Excessive iron was indicated by the mean phase below the normal mean phase. In the caudate, putamen, GP, and pulvinar thalamus, this iron measure predicted processing speed. However, after they adjusted for regional volumes, iron–cognition links were no longer significant. Thus, unlike in the current study using QSM and R\(^2\*\), Modica et al\(^6\) found no relationship between their iron measure and cognition when regional atrophy was controlled. Pinter et al\(^11\) reported correlations between a processing speed (“cognitive efficiency”) composite score (SDMT, PASAT) and a combined basal ganglia R\(^2\*) measure, while controlling for global brain measures (normalized brain volume, lesion load, magnetization transfer ratio for normal-appearing brain tissue). Methodologic differences between the behavioral and iron measures and the normalization procedures likely caused differences among studies, but generally, these findings emphasize the importance of examining cognitive correlates of DGM iron pathology in MS vis-a-vis atrophy, as was done here.

Because iron imaging in MS is relatively new and rapidly evolving, another important aspect here was the assessment of different iron MR imaging parameters, combined with cognition. Schmalbrock et al\(^11\) also tested both QSM and R\(^2\*) reporting a relationship between Flanker Task performance and caudate/anterior putamen iron assessed with QSM, but not R\(^2\*\), in patients with relapsing-remitting MS. The Brief Repeatable Battery of Neuropsychological Tests was applied but not included, and the cognitive effects of iron accumulation were not further examined against other DGM pathologies such as atrophy. Nevertheless, the findings are comparable with ours insofar as the QSM measure was a better predictor of cognition (ie, inhibition in a Flanker Task in Schmalbrock et al) than R\(^2\*\). While R\(^2\*) and QSM are complementary measures, they have distinct differences. First, the iron sensitivity of R\(^2\*) is highly field-dependent, with higher fields being advantageous,\(^29\) while QSM is largely field-insensitive.\(^30\) Second, R\(^2\*) is more susceptible to water content with inflammation weakening the R\(^2\*) signal but having little effect on QSM. Third, demyelination has opposing effects, with the QSM signal increasing from demyelination and R\(^2\*) decreasing.\(^31\) Thus, QSM increases by demyelinating effects adding to iron accumulation, while these 2 events oppose each other for R\(^2\*\). Thus, our findings that QSM shows a tighter coupling to cognition than R\(^2\*) may relate to both the dominant iron accumulation in DGM and the additive effects of DGM demyelination on QSM.

We observed a specific role of GP iron accumulation in global cognitive functions, irrespective of GP atrophy, implying that iron accumulation in the GP may have a unique role in globally affecting cognitive processes in this MS cohort. In an early study, Brass et al\(^9\) had approximated iron accumulation by examining T2 hypointensities (at 1.5T). The authors also reported hypointensities in the GP—but no other DGM region—to be the only significant predictor of a composite cognitive score. Similar to our findings, tests that were individually related to GP hypointensities included verbal fluency and the SDMT, but not memory. The GP has the highest iron concentration in the human brain, exceeding that in all other DGM nuclei,\(^12,32\) and it is a target region of several neurodegenerative diseases with primary brain iron accumulation etiologies.\(^33\) Functionally, the GP is the major input region to the thalamus within all the frontostriatal-thalamic loops, including lateral prefrontal and motor/supplemental motor cortex targets and serving a range of cognitive, emotional, motor, and oculomotor functions.\(^34\) A finer segregation within the GP would be valuable to delineate further whether specific psychomotor functions in MS are particularly vulnerable to iron accumulation.
and/or regional atrophy. As it stands, our findings only speak to a relationship between GP iron and global cognitive functions in MS. Thus, future studies should extend DGM imaging and add basal ganglia–specific tasks that also probe motor functions, considering the putamen-GP iron pathology observed here.

Among the study limitations, sample size for both patients and controls was relatively small so that finer graded analyses of DGM–cognition links within MS subtypes could not be performed reliably. Combining relapsing ($n = 16$) and progressive MS subtypes ($n = 24$) in the current study may have biased the results toward patients with progressive MS. Qualitatively, the relapsing–remitting MS subgroup took an intermediate place between patients with progressive MS and controls in all of the neuropsychological measures and in most of the DGM measures. However, larger scale studies with equally sized groups of MS subtypes should be performed. These may uncover differential relationships to cognition with more dominant inflammatory (relapsing MS) or atrophic/demyelinating (progressive MS) features. Our study also did not account for lesion burden, which has shown a relationship with cognitive decline and $R^2*$ in some, but not other studies. None of the lesions were visible in the DGM ROIs in the current study, by using manual tracing on T2/FLAIR images. Because the focus of the current study was to disambiguate the cognitive significance of iron pathology vis-à-vis atrophy specifically in the DGM and the MR parameters were not optimized for identification of DGM lesions, potential additive or separable effects of such lesions on cognition remain to be clarified in the future.

In addition, many patients were prescribed medications for symptom management. Although we did not observe an interaction between these medications and cognitive performance here, a formal assessment of mental health and other central nervous factors would have been preferable. Finally, the healthy controls had a higher educational level than the patient group. Besides controlling for education levels in each analysis, other aspects of the data speak against a strong influence of education on the final results. For example, On-line Tables 1 and 2 show that education never emerged as a significant predictor in any of the regression models in either example. On-line Tables 1 and 2 show that education never emerged as a significant predictor in any of the regression models in either example. On-line Tables 1 and 2 show that education never emerged as a significant predictor in any of the regression models in either example.

**CONCLUSIONS**

Increased iron in the GP, measured by QSM, was moderately associated with a lower cognitive composite score in this MS cohort. This effect was unrelated to atrophy of the GP. Whereas thalamic atrophy was the strongest predictor of cognitive performance in patients with MS, this outcome, in turn, was not further modulated by thalamic QSM/R2* iron measures. Our findings suggest separable and negative relationships among cognition, DGM iron, and DGM atrophy in MS.

**ACKNOWLEDGMENTS**

We would like to thank Leah White for her help with patient recruitment and chart review.

**REFERENCES**


tibility mapping with Perls’ iron staining for subcortical gray mat-
ABSTRACT

BACKGROUND AND PURPOSE: Neuromyelitis optica spectrum disorders are inflammatory demyelinating disorders with optic neuritis and/or longitudinally extensive transverse myelitis episodes. We now know that neuromyelitis optica spectrum disorders are associated with antibodies to aquaporin-4, which are highly concentrated on astrocytic end-feet at the blood-brain barrier. Immune-mediated disruption of the blood-brain barrier may manifest as contrast enhancement on brain MR imaging. We aimed to delineate the extent and frequency of contrast enhancement on brain MR imaging within 1 month of optic neuritis and/or longitudinally extensive transverse myelitis attacks and to correlate contrast enhancement with outcome measures.

MATERIALS AND METHODS: Brain MRIs of patients with neuromyelitis optica spectrum disorders were evaluated for patterns of contrast enhancement (periependymal, cloudlike, leptomeningeal, and so forth). The Fisher exact test was used to evaluate differences between the proportion of contrast enhancement in patients who were seropositive and seronegative for aquaporin-4 antibodies. The Mann-Whitney test was used to compare the annualized relapse rate and disease duration between patients with and without contrast enhancement and with and without seropositivity.

RESULTS: Brain MRIs of 77 patients were evaluated; 59 patients (10 males, 49 females) were scanned within 1 month of optic neuritis and/or longitudinally extensive transverse myelitis attacks and were included in the analysis. Forty-eight patients were seropositive, 9 were seronegative, and 2 were not tested for aquaporin-4 antibodies. Having brain contrast enhancement of any type during an acute attack was significantly associated with higher annualized relapse rates ($P = .03$) and marginally associated with shorter disease duration ($P = .05$). Having periependymal contrast enhancement was significantly associated with higher annualized relapse rates ($P = .03$).

CONCLUSIONS: Brain MRIs of patients with neuromyelitis optica spectrum disorders with contrast enhancement during an acute relapse of optic neuritis and/or longitudinally extensive transverse myelitis are associated with increased annual relapse rates.

ABBREVIATIONS: AQP4 = aquaporin-4; ARR = annualized relapse rate; CE = contrast enhancement; IgG = immunoglobulin G; LETM = longitudinally extensive transverse myelitis; NMO = neuromyelitis optica; NMOSD = NMO spectrum disorders; ON = optic neuritis

Enhancing Brain Lesions during Acute Optic Neuritis and/or Longitudinally Extensive Transverse Myelitis May Portend a Higher Relapse Rate in Neuromyelitis Optica Spectrum Disorders


Received December 22, 2016; accepted January 14, 2017.

From the Division of Neuroradiology (G.O., Y.P., II.), Russell H. Morgan Department of Radiology, and Department of Neurology (M.M., M.L.), The Johns Hopkins University School of Medicine, Baltimore, Maryland; Department of Radiology (K.Y.W.), Baylor College of Medicine, Houston, Texas; and Biostatistics Center (C.B.T.), Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland.

Carol Thompson received a National Institutes of Health–Clinical and Translational Science Awards grant 1UL1TR001079 (to Johns Hopkins). Maureen Mealy received honors from the International Organization of Multiple Sclerosis Nurses and EMD Serono. Michael Levy received research support from the Guthy-Jackson Charitable Foundation, ViroPharma, Acorda Therapeutics, Sanofi, Neuristem, and Genentech and serves as a consultant for Chugai Pharmaceuticals, GlaxoSmithKline, Alexion Pharmaceuticals, and MedImmune.

Paper previously presented in part as an oral presentation at: Annual Meeting of the American Society of Neuroradiology and the Foundation of the ASNR Symposium, April 25–30, 2015, Chicago, Illinois. The presentation number was O-444.

Please address correspondence to Izlem Izbudak, MD, Division of Neuroradiology, Russell H. Morgan Department of Radiology, The Johns Hopkins University School of Medicine, 600 N Wolfe St, Phipps B-126B, Baltimore, MD 21287; e-mail: izbudal@jhmi.edu

*Indicates open access to non-subscribers at www.ajnr.org

http://dx.doi.org/10.3174/ajnr.A5141


Neuromyelitis optica (NMO) is an inflammatory demyelinating disorder of the central nervous system, characterized by recurrent episodes of longitudinally extensive transverse myelitis (LETM) and/or optic neuritis (ON). Discovery of an NMO-specific autoantibody, NMO–immunoglobulin G (IgG), and its target autoantigen, aquaporin-4 (AQP4), have differentiated NMO from multiple sclerosis as a distinct disease entity. Moreover, given the high specificity of AQP4-IgG serology for clinically diagnosed NMO, such seropositivity was incorporated into the revised diagnostic criteria for NMO in 2006. The term “NMO spectrum disorders” (NMOSD) was introduced in 2007 to encompass broader phenotypes, including seropositive patients with coexisting autoimmune disorders and patients with limited or inaugural...
forms of NMO. The terms NMO and NMOSD were unified under a revised NMOSD definition in 2015. The unifying NMOSD diagnostic criteria allowed the diagnosis of NMOSD in patients without clinical involvement of the optic nerves or the spinal cord and stratified the diagnosis according to those with or without AQP4-IgG positivity.

While NMO was traditionally thought to be a disease exclusively involving the optic nerves and spinal cord, imaging abnormalities within the brain have been reported in a significant proportion of patients seropositive for AQP4-IgG, in regions with both high and low AQP4 expression. Lesions involving the diencephalon, area postrema, corpus callosum, hemispheric white matter, and corticospinal tracts have been reported. Specific patterns of contrast enhancement (CE) within the brain have also been reported in NMO, including pencil-thin, cloudblke, leptomeningeal, and perivascular enhancement (Figs 1 and 2). The current literature suggests a relatively low incidence of contrast-enhancing brain lesions in NMO.

However, in a large proportion of brain MRIs in those studies, whether they were acquired during an acute phase of the disease versus at any time point was not specified. Furthermore, the incidence of contrast-enhancing lesions in the brain during acute relapses of ON and/or LETM in patients with NMOSD has not been examined before, to our knowledge. Prior studies investigating predictors of relapse in patients with NMOSD have addressed factors that are either clinical or biochemical in nature, including AQP4-IgG seropositivity, female sex, and older age of onset. In contrast, no MR imaging parameters have been shown to be associated with disease outcome.

In the current study, we aimed to delineate the extent and frequency of CE in the brain during acute attacks of ON and/or LETM. We also sought to determine whether detection of brain CE was associated with specific outcome measures, including disease duration and the annualized relapse rate (ARR).

**MATERIALS AND METHODS**

**Patients**

A retrospective chart review was performed to identify patients with contrast-enhancing brain lesions between September 2001 and November 2013 at the Johns Hopkins NMO center. All patients identified were diagnosed with NMO or NMOSD based on the Wingerchuk et al 2006 or 2007 revised criteria, respectively. Institutional review board approval was obtained for the study. Electronic patient records were reviewed for demographic information, history of relapse, AQP4-IgG status, age at diagnosis, age at last follow-up, and the number of relapses.

**Neuroimaging**

MR imaging examinations were performed by using either 1.5T or 3T scanners (Philips Healthcare, Best, the Netherlands; GE Healthcare, Milwaukie, Wisconsin; and Siemens, Erlangen, Germany). T1WI, fast spin-echo T2WI, fast spin-echo FLAIR, and postgadolinium T1WIs were performed. A gadolinium contrast agent of 0.1 mL/kg was intravenously administered followed by a 20-mL saline injection. T1-weighted axial and coronal images were acquired without any delay after intravenous injection. The sagittal T1WIs were obtained with the following parameters: TR range = 520–696 ms, TE range = 4.6–14 ms, matrix size range = 192 × 192 to 512 × 196, FOV range = 190 × 190 mm to 240 × 240 mm, section thickness/spacing range = 1/1 to 5/7 mm. Axial T2WI was performed with the following parameters: TR range = 2500–7000 ms, TE range = 83–112 ms, matrix size range = 236 × 184 to 448 × 335, FOV range = 159 × 200 mm to 240 × 240 mm, section thickness/spacing range = 2/2 to 5/5 mm. A FLAIR sequence was obtained with the following parameters: TR = 6000 ms, TE = 120 ms, TI = 2000 ms, section thickness = 5 mm, FOV = 23 cm, matrix size = 256 × 256.

All brain MRIs were evaluated in consensus by 2 radiologists, a
Table 1: Association between the proportion of patients with CE and the presence of AQP4-IgG seropositivity

<table>
<thead>
<tr>
<th>AQP4-IgG</th>
<th>CE Present</th>
<th>CE Absent</th>
<th>No. of Patients (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>31 (64.6%)</td>
<td>17 (35.4%)</td>
<td>48</td>
</tr>
<tr>
<td>Negative</td>
<td>4 (55.6%)</td>
<td>5 (44.4%)</td>
<td>9</td>
</tr>
<tr>
<td>Total</td>
<td>35 (63.2%)</td>
<td>21 (36.8%)</td>
<td>57 (2 not tested)</td>
</tr>
</tbody>
</table>

a Fisher exact test, P = .7.
b One patient was not tested.

Table 2: Comparison of disease duration and ARR between patients with and without CE during an acute attack

<table>
<thead>
<tr>
<th>Disease duration (mean) (yr)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE Present</td>
<td>CE Absent</td>
</tr>
<tr>
<td>4.76 ± 4.81 (n = 21)</td>
<td>7.26 ± 5.75 (n = 38)</td>
</tr>
<tr>
<td>ARR (mean)</td>
<td>P Value</td>
</tr>
<tr>
<td>1.15 ± 0.73 (n = 21)</td>
<td>0.73 ± 0.52 (n = 38)</td>
</tr>
</tbody>
</table>

* P < .05, based on Mann-Whitney test.

Table 3: Comparison of disease duration and ARR between patients with and without PCE during an acute attack

<table>
<thead>
<tr>
<th>Disease duration (mean) (yr)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCE Present</td>
<td>PCE Absent</td>
</tr>
<tr>
<td>4.76 ± 4.81 (n = 21)</td>
<td>7.26 ± 5.75 (n = 38)</td>
</tr>
<tr>
<td>ARR (mean)</td>
<td>P Value</td>
</tr>
<tr>
<td>1.30 ± 0.83 (n = 14)</td>
<td>0.74 ± 0.49 (n = 45)</td>
</tr>
</tbody>
</table>

Note:—PCE indicates periependymal contrast enhancement.
a P < .05, based on the Mann-Whitney test.

Table 4: Distribution of brain CE patterns among 59 patients with ON and/or LETM

<table>
<thead>
<tr>
<th>Type of CE</th>
<th>No. of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Periependymal</td>
<td>14</td>
</tr>
<tr>
<td>Cloudlike</td>
<td>7</td>
</tr>
<tr>
<td>Leptomeningeal</td>
<td>2</td>
</tr>
<tr>
<td>Isolated</td>
<td>4</td>
</tr>
<tr>
<td>Ring</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
</tr>
<tr>
<td>Absent</td>
<td>28</td>
</tr>
</tbody>
</table>

RESULTS

Brain MRIs of 77 patients (11 males, 66 females) were evaluated for contrast enhancement. Fifty-nine patients (10 males, 49 females) underwent brain MR imaging within 1 month of the onset ON and/or LETM attack and were included in the final analysis. The mean age of patients was 47.8 years (range, 6–78 years). There were 35 African-American, 18 white, and 6 Hispanic (individuals from Mexico) individuals. Forty-eight patients were AQP4-IgG seropositive, 9 were seronegative, and the AQP4-IgG status was not checked in 2 of them. The ARR was not available for 1 patient.

Table 1 depicts the proportions of patients with CE in those with or without AQP4-IgG seropositivity during acute attacks. The Fisher exact test did not demonstrate significantly different proportions of CE in patients with or without AQP4-IgG seropositivity during acute attacks (P = .7). No significantly different proportions were noted when stratified by specific enhancement patterns (P = .7, data not shown).

Tables 2 and 3 depict the association between the detection of CE during an acute phase and either disease duration or ARR. When imaged during the acute phase, patients demonstrating periependymal CE had significantly higher ARRs compared with those without periependymal CE (P = .03). Moreover, patients demonstrating any type of CE during the acute phase had significantly higher ARRs (P = .03) than those without.

On the basis of the regression analyses, the unadjusted difference in ARRs between those with periependymal CE and those without it was 0.56 (95% CI, 0.07–1.05; P = .03). After we adjusted for age, sex, race, and AQP4 status, the difference was 0.60 (95% CI, 0.08–1.13; P = .03). The unadjusted difference in ARRs between those with any CE and without was 0.42 (95% CI, 0.04–0.80; P = .03). After we adjusted for age, sex, race, and AQP4 status, the difference was 0.41 (95% CI, 0.02–0.81; P = .04).

Table 4 shows the distribution of brain CE patterns among 59 patients who were scanned within 1 month of an ON and/or LETM attack. Brain CE was categorized and evaluated in 6 specific patterns of enhancement: periependymal, cloudlike, leptomeningeal, isolated, ring, or other (Figs 1 and 2).

DISCUSSION

The current literature on NMO is limited in its description of neuroimaging features that may predict the outcome of disease. Most asymptomatic NMO brain lesions have not been shown to demonstrate enhancement, and the frequency of acute lesion-associated enhancement remains to be determined. This study demonstrates that approximately 63% of patients during an acute attack of ON and/or LETM may also show CE within the brain parenchyma. CE within the brain, when identified during an acute phase, is associated with a significantly increased ARR. The relapse rate during the first 2 years of the disease strongly determines the risk of an unfavorable outcome as defined by severe disability or death. Brain enhancement in patients during an acute ON and/or LETM may reflect a more severe underlying disease process compared with those without brain CE.

We found no significant difference in the propensity for CE in patients who were AQP4-IgG seropositive (64.6%) and seronegative (55.6%) (P = .7, Table 1). CE patterns of brain lesions in the
current literature were described mostly in patients seropositive for AQP4-IgG and have been reported to range from 3% to 56%, excluding small case reports. Our study revealed a much higher proportion of CE, approximating 64.6% and 55.6% in patients seropositive and seronegative, respectively. A study investigating contrast-enhancing LETM lesions reported CE in 94% and 71% of seropositive and seronegative patients, respectively, though the authors did not specify the location of CE as within either the brain or spinal cord. To our knowledge, the current study is the largest cohort to report the frequency of contrast-enhancing brain lesions in both seropositive and seronegative patients during active ON and/or LETM relapse of NMO. The small sample of patients seronegative for AQP4-IgG in the current study may be contributing to lack of detection for a significant difference in the proportion of CE seen in seropositive and seronegative patients; however, that might be the case in prior studies failing to show a difference as well. Nevertheless, the number of patients seronegative for AQP4-IgG is always low compared with those who are seropositive; therefore, multi-institutional studies are needed to increase the sample size.

That AQP4 is highly expressed on astrocytic foot processes at the BBB and contributes to the maintenance of BBB integrity is well-described. Binding of AQP4-IgG to AQP4 in vitro has been shown to alter BBB permeability and astrocyte killing. Disruption of the BBB manifests as CE on brain MR imaging. Periependymal white matter is one of the most AQP4-rich regions of the brain; hence, the high prevalence of periependymal contrast enhancement in our cohort. Furthermore, AQP4-IgG is thought to be pathogenic only in proximity to CNS parenchyma, as evidenced by NMO-like histopathology in animal models in those that received direct administration of AQP4-IgG into the CNS. In contrast, peripheral administration had no effect. The presence and levels of AQP4-IgG in CSF are associated and correlated with those in serum during acute relapses. For example, AQP4-IgG is detectable in the CSF of most seropositive patients with serum titers of >1:250 during an acute relapse. Moreover, the amount of CSF AQP4-IgG is correlated with astrocyte damage and BBB breakdown. Therefore, it is possible that those with lower serum titers or those not in acute relapse in the current study may not have detectable or significant CSF levels of AQP4-IgG to lead to the BBB disruption and consequent CE. Most interesting, there have been reports of patients with NMOSD who are AQP4-IgG positive in the serum for many years before the onset of symptomatic disease. The poor correlation between the presence and level of serum and CSF titers of AQP4-IgG may be contributing to the lack of significantly different proportions of CE in seropositive and seronegative patients in our study. It may also be contributing to the high interstudy variability in the reports of the percentage of CE observed in patients with NMOSD.

Most important, the association between CE during the acute phase and ARR may be confounded because those who underwent brain MRIs during an acute phase may have warranted more immediate imaging because they may have been inherently sicker. Findings may be further confounded by other clinical characteristics. For example, longer intervals between the first and second attack, older age at onset, patients of African origin, female sex, and AQP4-IgG seronegativity are associated with worse outcomes and/or higher relapse rates. However, other studies failed to find AQP4-IgG status as a predictor of outcome. Jarius et al found that AQP4-IgG status did not differ significantly with regard to time to relapse or ARR. Jiao et al found that the effect of seropositive status on the relapse rate and disability outcome did not differ. Responses to plasmapheresis based on AQP4-IgG were also not significantly different. Regardless of the discrepant prognostic findings of AQP4-IgG in the existing literature, in our study, a periependymal pattern of CE and the presence of any pattern of CE in the acute phase remained significant predictors of higher ARRs after adjusting for AQP4 status, as well as age, sex, and race, on multivariable analysis.

Enhancement patterns of brain lesions in NMO have some unique features and sometimes, in the presence of characteristic T2 lesions, might aid in making a specific diagnosis. Patchy CE with blurred margins, so-called “cloudlike enhancement,” is the most commonly reported enhancement pattern in the literature. More recently, linear periependymal CE, so called “pencil-thin enhancement,” and leptomeningeal CE were proposed as more specific patterns than cloudlike enhancement. Isolated CE and ring and open-ring CE are considered specific to MS, and they are rarely seen in patients with NMOSD. However, although rare, these intense, well-defined CE patterns have been described before, especially in seronegative patients with NMOSD.

The main limitation of our study is the retrospective design, and factors that have been described in the literature associated with outcomes such as seropositivity status, sex, race, and age at onset may be potential confounders and were not accounted for. There is a possible selection bias based on a group of patients with NMOSD who required brain MR imaging, which may reflect a different subpopulation than that not requiring brain MRIs. The threshold of 1 month as the criterion for an acute attack may be arbitrary, given the lack of information in records available to more accurately assess the patients’ clinical statuses and may thus misrepresent these statuses in the current study. Furthermore, the current study was originally conducted before the introduction of the more inclusive revised diagnostic criteria for NMOSD of 2015. Rather, included patients were based on the 2006 diagnostic criteria; therefore, the current study does not account for patients who may now qualify as diagnostic for NMOSD under the 2015 criteria.

CONCLUSIONS
Detection of CE in postgadolinium T1-weighted brain imaging within 1 month of onset of an acute ON and/or LETM is associated with higher ARRs. CE is an important marker reflecting the underlying pathogenic process of NMOSD. Although no significant association was found between CE and AQP4-IgG serostatus, the strong interplay among the BBB disruption, AQP4-IgG deposition, and CE warrants further investigation with a larger multicenter cohort to determine the prognostic role that CE may play as a predictor of outcome and its correlation with clinical severity.

Disclosures: Carol B. Thompson—RELATED: Grant: National Institutes of Health Clinical and Translational Science Awards UL1TR000107 to Johns Hopkins. Comments: The grant provides limited free statistical consultation to researchers. Mau- reen Mealy—UNRELATED: Payment for Lectures Including Service on Speakers
REFERENCES


Can MRI Visual Assessment Differentiate the Variants of Primary-Progressive Aphasia?


ABSTRACT

BACKGROUND AND PURPOSE: Primary-progressive aphasia is a clinically and pathologically heterogeneous condition. Nonfluent, semantic, and logopenic are the currently recognized clinical variants. The recommendations for the classification of primary-progressive aphasia have advocated variant-specific patterns of atrophy. The aims of the present study were to evaluate the sensitivity and specificity of the proposed imaging criteria and to assess the intra- and interrater reporting agreements.

MATERIALS AND METHODS: The cohort comprised 51 patients with a root diagnosis of primary-progressive aphasia, 25 patients with typical Alzheimer disease, and 26 matched control participants. Group-level analysis (voxel-based morphometry) confirmed the proposed atrophy patterns for the 3 syndromes. The individual T1-weighted anatomic images were reported by 3 senior neuroradiologists.

RESULTS: We observed a dichotomized pattern of high sensitivity (92%) and specificity (93%) for the proposed atrophy pattern of semantic-variant primary-progressive aphasia and low sensitivity (21% for nonfluent-variant primary-progressive aphasia and 43% for logopenic-variant primary-progressive aphasia) but high specificity (91% for nonfluent-variant primary-progressive aphasia and 95% for logopenic-variant primary-progressive aphasia) in other primary-progressive aphasia variants and Alzheimer disease (sensitivity 43%, specificity 92%). MR imaging was least sensitive for the diagnosis of nonfluent-variant primary-progressive aphasia. Intrarater agreement analysis showed mean κ values above the widely accepted threshold of 0.6 (mean, 0.63 ± 0.16). Pair-wise interobserver agreement outcomes, however, were well below this threshold in 5 of the 6 possible interrater contrasts (mean, 0.41 ± 0.09).

CONCLUSIONS: While the group-level results were in precise agreement with the recommendations, semantic-variant primary-progressive aphasia was the only subtype for which the proposed recommendations were both sensitive and specific at an individual level.

ABBREVIATIONS: AD = Alzheimer disease; lvPPA = logopenic-variant PPA; nfvPPA = nonfluent-variant PPA; PPA = primary-progressive aphasia; svPPA = semantic-variant PPA

Primary-progressive aphasia (PPA) is a clinically and pathologically heterogeneous condition characterized by insidious onset and gradual worsening of language due to degeneration of brain language areas. Clinical heterogeneity, compounded by the evolution of signs and symptoms, makes accurate classification of patients a challenging task. Making a reliable clinical diagnosis, on the other hand, is important. Despite lack of a one-to-one relationship between the clinical diagnosis and the underlying pathology, previous clinicopathologic series have identified probabilistic associations among the 3 recognized clinical presentations of PPA and certain pathologies. There are established associations between semantic-variant PPA (svPPA) and frontotemporal lobar degeneration–TAR DNA binding protein 43 (TDP-43); nonfluent-variant PPA (nfvPPA) and frontotemporal lobar degeneration-tau; and logopenic-variant PPA (lvPPA) and Alzheimer pathology.1–5

The recommendations on clinical subtyping of PPA have proposed that clinical classification can be supported by imaging according to the pattern of regional atrophy or metabolic impairment.6 Left posterior frontoinsular atrophy in nfvPPA, anterior temporal atrophy in svPPA, and left posterior peri-Sylvian or parietal atrophy in lvPPA are the recommended atrophy patterns.
Remarkably, these imaging recommendations are derived from studies that either used group-averaged data—which though highly replicated, are not necessarily valid for single-patient diagnosis—or were based on observed atrophy in convenience samples without qualification of sensitivity, specificity, or reliability. Little is known about whether individual patients, as opposed to groups, fulfilling the clinical criteria for these variants reliably present with the prescribed patterns of atrophy and whether these patterns have sufficient reliability to be exploited to arrive at an accurate syndromic diagnosis.

The aims of the present study were the following: 1) to evaluate the utility of the proposed imaging criteria for the diagnosis of PPA variants by contrasting the patterns of atrophy in individual patients with PPA, as reported by senior neuroradiologists, with the recommendations from the criteria; and 2) to assess the neuroradiologists’ intra- and interrater agreement, which, in turn, would be an indication of the robustness of the observed abnormalities.

**MATERIALS AND METHODS**

**Participants**

The cohort comprised 51 patients with a root diagnosis of PPA, 25 patients with mild typical Alzheimer disease (AD) as a neurodegenerative control group, and 26 healthy age- and education-matched control participants. The breakdown of the subjects with PPA based on clinical variants was 21 with svPPA, 14 with nfvPPA, 14 with mixed PPA, and 2 with lvPPA. Clinical diagnoses were made in accordance with the published criteria for the diagnosis of PPA and probable Alzheimer disease. The diagnosis of PPA variants was based on a quantitative application of the consensus recommendations as detailed elsewhere. No patients had pedigrees to suggest an autosomal dominant genetic cause.

The mixed-PPA group, however, was designated “mixed” on the basis of strict application of the proposed clinical criteria. The patients almost certainly corresponded, however, to what others have designated lvPPA in that they had neither svPPA nor nfvPPA, and they had the same group-level atrophy pattern as in previous lvPPA cohorts. Furthermore, some researchers have proposed to diagnose lvPPA through a hierarchic decision tree in which the key feature of this group is that they are neither svPPA nor nfvPPA, and that they had the same group-level atrophy pattern as in previous lvPPA cohorts. Applying such an algorithm to the present mixed cases would also have them classified as lvPPA. The subjects with mixed PPA in this study should, therefore, be considered analogous to those with lvPPA and are referred to henceforth as such.

The study was approved by the institutional review board of Cambridge University hospitals, UK.

**Neuropsychological Battery and Connected Speech Analysis**

All patients underwent comprehensive neuropsychological and connected speech assessment before imaging, full details of which have been published previously. The On-line Table provides a summary of some of these data.

**Imaging**

**Image Acquisition.** Study participants were scanned within an average of 1.6 ± 0.8 months from cognitive assessment. All MR imaging was performed on the same Magnetum Trio 3T system (Siemens, Erlangen, Germany). T1-weighted anatomic images were acquired by using 3D MPRAGE with the following imaging parameters: TR/TE/T1/flip angle = 2300/2.86/900 ms/90°, 144 sections, 192 × 192 matrix dimensions, and 1.25 × 1.25 × 1.25 mm³ voxel size. Receiver bandwidth and echo spacing were 240 Hz/pixel and 6.7 ms, respectively.

**Data Processing and Group-Level Data Analysis**

All obtained T1 volumes were preprocessed as reported previously. Preprocessing and warping procedures need reasonable initial estimates; hence, the origin of each structural volume was set manually to the anterior commissure before preprocessing. All volumes were then spatially normalized and segmented by using the unified segmentation model in statistical parametric mapping 5 (SPM5) (http://www.fil.ion.ucl.ac.uk/spm/software/spm5/). The segments were also modulated to compensate for volumetric differences introduced into the warped images. Finally, gray matter segments were smoothed by using an 8-mm full width at half maximum isotropic Gaussian kernel. Total intracranial volumes were calculated by using the automated SPM technique as described elsewhere, and the obtained values, along with age, were fed into the statistical models as nuisance covariates. Following these steps, a 2-sample t test implemented in SPM5 contrasted the gray matter volumes of the patient groups against those in controls. The statistical maps were thresholded at P < .01, corrected for multiple comparisons (false discovery rate = .01).

**Visual Reporting of Individual Scans**

Three senior neuroradiologists who were blinded to the clinical diagnoses of the study participants separately reported all unprocessed T1 sequences displayed by using the FMRIB Software Library (FSL, Version 4.1.2; http://www.fmrib.ox.ac.uk/fsl). Thirty-five scans (%7 for each diagnostic group including controls) were duplicated, bringing the total number of scans to 137 to assess intrarater agreement.

The neuroradiologists were asked to report the scans for the presence and patterns of disproportionate regional or global atrophy in 2 stages. In the first stage, the outcome of which was used for calculations of intra- and interrater agreement and sensitivity, the radiologists were asked to report the scans in their own preferred styles. Subsequent calculations for this stage were based on the reported lobar distribution of the abnormalities. “Global atrophy” and “no atrophy” were also accepted as valid entries. In agreement with the published recommendations for AD and PPA, the following lobar distributions were deemed consistent with the syndromic PPA variants and Alzheimer disease: temporal lobe atrophy for svPPA; left frontal or left frontotemporal atrophy for nfvPPA; left temporal, left parietal, or left temporoparietal atrophy for lvPPA; and temporal, parietal, or temporoparietal atrophy for typical AD. Rating a scan as showing “global atrophy” was not deemed acceptable for any of the syndromic variants of PPA or for AD because the reporting radiologists had not observed “disproportionate” atrophy of a target region. In the second stage, the outcome of which was the basis for specificity and separate sensitivity and agreement calculations, however, the raters were specifically asked to comment on whether there was “disproportionate” left posterior frontoinsular atrophy (indicative of...
nfvPPA), anterior temporal lobe atrophy (indicative of svPPA), left posterior peri-Sylvian or parietal lobe atrophy (indicative of lvPPA), and medial temporal or parietal atrophy (indicative of typical AD). Instructions for this second stage were only given after stage 1 was completed to ensure that the initial ratings were not biased by expected atrophy patterns.

### Statistical Considerations

Predictive Analytics Software (PASW, Version 18; IBM, Armonk, New York) and SPM5 were used for statistical analysis of the data. One-way ANOVA with a 2-tailed significance level of .05 was used to compare the demographic and neuropsychological measures. A pair-wise k was used to assess the intra- and interobserver agreement in the reports of atrophy. Group-level comparisons of the imaging data were made with a 2-sample t test implemented in SPM5, with age and total intracranial volume included as nuisance covariates. Group-level results are reported at a false discovery rate–corrected P < .01.

### RESULTS

Demographic data for all participant groups are summarized in Table 1. Neuropsychological and language assessment results are presented in the On-line Table.

![Figure 1](https://example.com/figure1.png)

**FIG 1.** Group-level patterns of atrophy in identical axial, sagittal, and coronal sections of the brain. Images are displayed in neurologic orientation. Asterisks demonstrate the section most representative for the particular groups. All comparisons were made at false discovery rate–corrected P < .01.

<table>
<thead>
<tr>
<th>Demographics</th>
<th>svPPA Mean (Range)</th>
<th>nfvPPA Mean (Range)</th>
<th>lvPPA Mean (Range)</th>
<th>AD Mean (Range)</th>
<th>Control Mean (Range)</th>
<th>Omnibus Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at test (yr)</td>
<td>67 (60–79)</td>
<td>68.9 (53–79)</td>
<td>70.8 (60–83)</td>
<td>68 (60–79)</td>
<td>67.5 (51–80)</td>
<td>NS</td>
</tr>
<tr>
<td>Disease duration (mo)a</td>
<td>56 (24–108)</td>
<td>38.57 (18–60)</td>
<td>48.7 (24–108)</td>
<td>58 (24–96)</td>
<td>NA</td>
<td>NS</td>
</tr>
<tr>
<td>Education (yr)a</td>
<td>13.6 (10–18)</td>
<td>12.6 (10–20)</td>
<td>11.6 (9–16)</td>
<td>12.5 (10–19)</td>
<td>12.8 (10–20)</td>
<td>NS</td>
</tr>
<tr>
<td>Sex</td>
<td>1 (0–5)</td>
<td>0.53 (0–4)</td>
<td>0.68 (0–4)</td>
<td>2.3 (0–6)</td>
<td>11 M, 15 F</td>
<td>–</td>
</tr>
<tr>
<td>ADL-Q</td>
<td>0 (0–5)</td>
<td>0.53 (0–4)</td>
<td>0.68 (0–4)</td>
<td>2.3 (0–6)</td>
<td>11 M, 15 F</td>
<td>–</td>
</tr>
</tbody>
</table>

**Note:** —ADL-Q indicates Activities of Daily Living Questionnaire; Sig, significant; NA, not applicable; NS, not significant.

*a* Nonparametric test.
lvPPA, by left posterior temporoparietal atrophy; and typical AD showed bilateral hippocampal and patchy temporoparietal atrophy.

Concerning the single-subject visual reporting outcomes, sensitivity calculations based on the lobar distribution of the abnormalities revealed almost perfect results in the svPPA group (mean sensitivity, 98% ± 2.9%) but low sensitivity of the imaging markers in the other PPA variants and the typical AD group (Table 2). The proposed imaging markers were least sensitive for the diagnosis of nfvPPA (mean, 29% ± 21%). Sensitivity values for the lvPPA and typical AD groups were modest at 57% and 53%, respectively. Sensitivity figures based on the prescribed patterns of atrophy revealed slightly lower values compared with the above figures but a similar pattern overall: high sensitivity for svPPA and low values for the other study groups (Table 2). Specificity figures were, however, consistently high for all diagnostic groups with no discernible difference (Table 3).

Tables 4 and 5 provide values for intra- and interobserver agreement for all diagnostic groups. Intraobserver agreement values for the reported lobar distribution of atrophy were consistently above the widely accepted threshold of 0.6 (mean, 0.75 ± 0.18), indicating substantial agreement. The recommendation-based values, however, fell below the 0.6 threshold for 2 of the 3 reporting radiologists (mean, 0.63 ± 0.16). In the pair-wise interobserver agreement values, while κ was just below the 0.6 threshold (mean, 0.56 ± 0.08) for the reported lobar distribution of atrophy (first round), it dropped considerably to 0.41 ± 0.09 for the recommendation-based outcomes; 0.4 is generally considered the minimum threshold for moderate agreement.

**DISCUSSION**

This study provides an objective assessment of the utility of the proposed MR imaging markers in supporting the diagnosis of various PPA variants. The group-level atrophy patterns were in precise agreement with the proposed imaging criteria for different PPA subtypes. Assessing the reliability of these measures at a single-subject level is, however, much more relevant—indeed mandatory—for determining the diagnostic utility of the proposed imaging markers in clinical practice.
imaging criteria in further classification of individual patients with PPA. In the absence of reliable automated single-subject statistical measures capable of detecting the abnormalities at an individual level, assessment of the consistency and reliability of neuroradiologists’ reports along with the level of agreement constitutes a suitable substitute. Moreover, it mirrors real-life clinical practice in which visual rating of scans remains the standard reporting method.

The group-level voxel-based morphometry–gray matter atrophy patterns for each of the PPA variants (Fig 1) were consistent with those in past studies.7–9 This finding was important to confirm because it was precisely this group-level atrophy pattern that led to the recommendations for imaging-supported diagnoses. The results of the visual rating suggested a dichotomized pattern of high sensitivity and specificity of the proposed imaging markers for svPPA, but less reliable outcomes for the other 2 PPA subtypes, with a low sensitivity but rather high specificity. Agreement analyses for the whole group revealed substantial intrarater but only moderate interrater agreement values (mean $\kappa$, 0.63 and 0.41, respectively) for the recommendation-based atrophy patterns. This finding is now discussed in more detail for each variant.

**svPPA**

svPPA is characterized by an amodal loss of knowledge that consistently presents as a reduction of expressive vocabulary and word comprehension.27 Various studies have emphasized the importance of the anterior temporal lobes as hubs of semantic knowledge.28,29 In agreement with the proposed recommendations, group-average voxel-based morphometry analysis of the svPPA participants revealed predominant bilateral anterior temporal lobe atrophy. The high sensitivity and specificity of the proposed imaging markers (means, 0.98% and 93%, respectively) demonstrated that the presence of temporal lobe atrophy offered robust support for the diagnosis of svPPA. This finding is not unexpected because, though not always assessed systematically, atrophy of rostral-inferior temporal structures has been consistently reported in svPPA (also known as semantic dementia) both at a group level and individually.7,10,30–35 Given the uniform pattern of atrophy seen in this consecutively recruited cohort of 21 patients with svPPA, it can be argued that the diagnosis of svPPA should be seriously questioned in the absence of this atrophy pattern. Gil-Navarro et al36 found the same consistent presence of anterior temporal lobe atrophy in a study of 29 patients with PPA that included 5 with svPPA. As already mentioned, most patients with svPPA have frontotemporal lobar degeneration–TDP-43 pathology, but frontotemporal lobar degeneration–tau pathology is found occasionally. Previous work has indicated that the atrophy pattern does not discriminate between these 2 pathologic substrates.31

**nfvPPA**

Clinical features of nfvPPA include effortful, halting speech with sound distortions and/or grammatic errors in language production.6 Degeneration of the left frontal operculum and rostral insula is the culprit lesion in nfvPPA. More recent studies have also highlighted involvement of premotor32 and basal ganglia37 regions. Like svPPA, the group-average voxel-based morphometry findings in our nfvPPA cohort were largely compatible with the proposed diagnostic recommendations. As demonstrated in Fig 1, voxel-based morphometry analysis clearly identified disproportionate left-sided atrophy in the frontal operculum and insula. There was, in addition, evidence of further atrophy in the left basal ganglia region, but no atrophy was visible in the premotor area.

Concerning the single-subject outcomes, none of the patients with nfvPPA had unanimous reports of the prescribed atrophy pattern by all 3 neuroradiologists. In fact, “left posterior frontal and insular atrophy” was only the third most commonly observed report in this group with “no atrophy” and “left posterior peri-Sylvian atrophy” being the first and second, respectively (data not shown). “No atrophy” was reported by at least one of the neuroradiologists in 10 of 14 (71%) individuals with nfvPPA and in 22 of the total 42 (14 × 3) reports (52%). Low sensitivity values (mean, 21% ± 7%) further corroborated the above findings. Given the abundance of no-atrophy reports, the most plausible explanation for this result seems to be that atrophy in patients with nfvPPA is often very subtle. Also, clinical heterogeneity inherent in the recommended features of nfvPPA (ie, requiring the presence of either abnormal speech or agrammatism) and more white than gray matter burden are other potential explanations for the observed discrepancies. High specificity values, however, indicated the potential utility of the prescribed pattern of atrophy in the diagnosis of nfvPPA if present.

Further evidence for the inconsistency of imaging findings in nfvPPA comes from previous single case studies reporting widely discrepant findings, ranging from no atrophy38 to left hemispheric atrophy39 to left frontotemporal atrophy,14 bifrontal atrophy,40 and generalized atrophy.41 Even group-level findings, using parametric analysis techniques such as SPM, have been inconsistent, with different studies showing evidence of: left-sided inferior frontal and insular atrophy42 and hypometabolism43; atrophy in a wide distribution comprising the left inferior frontal, superior temporal, and inferior parietal areas, with7,9 and without additional atrophy in the premotor area; and finally abnormalities in the prefrontal cortex and left basal ganglia.8,12,44 One previous study reported a considerably higher sensitivity for the MR imaging–defined atrophy pattern of nfvPPA (76%).76 The discrepancy, however, likely relates to the study design in that the raters had to expressly classify scans for the 3 proposed atrophy patterns and in a group comprising only patients with PPA (there were neither healthy controls nor controls with dementia). This difference is important, given the high prevalence of no-atrophy reports in our nfvPPA group (see above). Considering that the main challenge in the diagnosis of degenerative aphasia is at the mildest stages when it is difficult to distinguish degenerative aphasia from a normal variation, inclusion of scans with normal findings makes our study a closer reflection of real-life situations.

**lvPPA**

LvPPA is the PPA variant that is highly associated with Alzheimer pathology.7 Impaired single-word retrieval in spontaneous speech and impaired sentence repetition are the recommended features. While a number of studies have failed to demonstrate the utility of these features in the diagnosis of lvPPA,17,45 atrophy of the left
temporoparietal lobe has been consistently emphasized in Alzheimer disease—related aphasia.\textsuperscript{17,19,45} In the “Materials and Methods” section, we mentioned our rationale for applying the lvPPA label to the group of patients with PPA whom we had previously reported as having mixed PPA. Turning to the individual visual reporting outcomes, we found low-to-moderate sensitivity (mean, 49% ± 7%) for the prescribed atrophy patterns. This was consistent with the results of the only previous study looking at the same metrics that found a sensitivity of 57% for MR imaging atrophy.\textsuperscript{39} Given the high specificity value (mean, 95% ± 2%), it can be inferred that while the presence of the prescribed pattern of “left posterior peri-Sylvian or parietal” atrophy is highly suggestive of lvPPA, its absence does not exclude the diagnosis of lvPPA. In addition, assessment of individual reports revealed that contrary to what might be expected, a typical AD atrophy pattern (ie, medial temporal or parietal atrophy) was reported in only 8% of the ratings of patients with lvPPA; an lvPPA atrophy pattern was the most common (38%), while an nfvPPA atrophy pattern was the second most frequently reported outcome (20%). This radiologic finding resonates with the previously reported difficulty in distinguishing nfv- and lvPPA variants on clinical and neuropsychological grounds.\textsuperscript{19,45} None of the radiologists had reported global atrophy for the n = 16 lvPPA cohort. This is an important negative, given possible concerns about the severity of dementia in this group.

CONCLUSIONS

This study provides an objective assessment of the utility of the proposed MRI recommendations for supporting the diagnosis of 3 PPA variants. Our findings are largely compatible with the only previous study on the subject.\textsuperscript{36} Moreover, to our knowledge, this article is the first to report the intra- and interrater agreement of the reporting radiologists and the specificity of the MR imaging markers for the diagnosis of PPA variants. Our study provides compelling evidence for the utility of the proposed imaging recommendations for the diagnosis of svPPA. On the basis of the findings of the current and previous studies, it could even be argued that lack of anterior temporal lobe atrophy should exclude the diagnosis of svPPA. The results were less consistent in the other groups. While high specificity values observed in all groups indicate the potential utility of the recommendations for patients in whom the atrophy patterns can be identified, low sensitivity and modest agreement values suggest that absence of the proposed atrophy patterns is common in the nonsemantic PPA subtypes.

Disclosures: S. Ahmad Sajjadi—RELATED: Grant: Donald Forrester Trust. Comments: grant to the University of Cambridge* Justin Cross—RELATED: Grant: National Institute for Health Research Biomedical Research Centre. Comments: part funding of salary. *Money paid to the institution.

REFERENCES

and structural MR image analysis and implementation as FSL. Neuroimage 2004;23(suppl 1):S208–19 CrossRef Medline


Entorhinal Cortex: Antemortem Cortical Thickness and Postmortem Neurofibrillary Tangles and Amyloid Pathology


ABSTRACT

BACKGROUND AND PURPOSE: The entorhinal cortex, a critical gateway between the neocortex and hippocampus, is one of the earliest regions affected by Alzheimer disease–associated neurofibrillary tangle pathology. Although our prior work has automatically delineated an MR imaging–based measure of the entorhinal cortex, whether antemortem entorhinal cortex thickness is associated with postmortem tangle burden within the entorhinal cortex is still unknown. Our objective was to evaluate the relationship between antemortem MRI measures of entorhinal cortex thickness and postmortem neuropathological measures.

MATERIALS AND METHODS: We evaluated 50 participants from the Rush Memory and Aging Project with antemortem structural T1-weighted MR imaging and postmortem neuropathologic assessments. Here, we focused on thickness within the entorhinal cortex as anatomically defined by our previously developed MR imaging parcellation system (Desikan-Killiany Atlas in FreeSurfer). Using linear regression, we evaluated the association between entorhinal cortex thickness and tangles and amyloid-β load within the entorhinal cortex and medial temporal and neocortical regions.

RESULTS: We found a significant relationship between antemortem entorhinal cortex thickness and entorhinal cortex (P = .006) and medial temporal lobe tangles (P = .002); we found no relationship between entorhinal cortex thickness and entorhinal cortex (P = .09) and medial temporal lobe amyloid-β (P = .09). We also found a significant association between entorhinal cortex thickness and cortical tangles (P = .003) and amyloid-β (P = .01). We found no relationship between parahippocampal gyrus thickness and entorhinal cortex (P = .31) and medial temporal lobe thickness (P = .05).

CONCLUSIONS: Our findings indicate that entorhinal cortex–associated in vivo cortical thinning may represent a marker of postmortem medial temporal and neocortical Alzheimer disease pathology.

ABBREVIATIONS: AD = Alzheimer disease; EC = entorhinal cortex; SE = standard error

The human entorhinal cortex (EC) plays an integral role in memory formation and serves as the critical gateway between the hippocampus and neocortex.1 Located in the medial temporal lobe, the EC constitutes the anterior portion of the parahippo...
pocampal gyrus and is localized in vivo laterally by the rhinal sulcus, anteriorly by the amygdala and hippocampus, and posteriorly by the posterior portion of the parahippocampal gyrus.\textsuperscript{2,3} The EC is one of the earliest affected regions in Alzheimer disease (AD). Tau-associated neurofibrillary tangle pathology in AD follows a defined topographic and hierarchical pattern, first affecting the EC and then progressing to anatomically connected limbic and association cortices; in contrast, amyloid-\(\beta\)-associated pathology does not involve the EC in the earliest stages of AD but selectively affects the neocortical regions.\textsuperscript{4,5} Most important, the association of neuronal volume loss in the EC has been shown to parallel \(\tau\)-associated tangle pathology (neurofibrillar tangles and neuritic plaques) but not senile plaques seen with amyloid-\(\beta\) deposition.\textsuperscript{6}

Structural MR imaging provides visualization and quantification of volume loss and has been extensively investigated in AD.\textsuperscript{6-9} Using manually delineated assessments, early studies have shown that volumetric measures of the EC can identify individuals without dementia in the earliest stages of the AD process.\textsuperscript{7-9} Within the past decade, rapid advances in MR imaging postprocessing have led to the development of software tools for automatic quantification of human subcortical and neocortical regions.\textsuperscript{10} We have previously developed an MR imaging–based parcellation atlas for the human cerebral cortex, which has automatically delineated the entorhinal cortex.\textsuperscript{11} The EC ROI from our parcellation atlas correlates with CSF levels of \(\tau\), amyloid,\textsuperscript{12} and Apolipoprotein \(E\).\textsuperscript{13} and has been used to identify cognitively healthy\textsuperscript{14} and cognitively impaired individuals without dementia who are most likely to progress to clinical AD.\textsuperscript{15,16} However, whether our antemortem MR imaging–based measure of the EC is associated with established postmortem measures of AD pathology is still unknown.

In this study, we evaluated the relationship between antemortem MR imaging–based automated measurements of EC thickness and postmortem measures of neurofibrillary tangle and amyloid-\(\beta\) pathology. To assess the specificity of our EC ROI (anterior portion of the parahippocampal gyrus), we also evaluated the relationship between the thickness of the posterior parahippocampal gyrus and EC amyloid-\(\beta\) and \(\tau\) pathology.

**MATERIALS AND METHODS**

**Participants**

We evaluated participants from the Rush Memory and Aging Project, a community-based longitudinal study of aging, which began in 1997.\textsuperscript{17} Details of the clinical and neuropathologic evaluations have been reported previously.\textsuperscript{17,18} Briefly, all participants underwent a uniform structured clinical evaluation that included a medical history, physical examination with emphasis on neurologic function, and neuropsychological testing (including the Mini-Mental State Examination and 20 other tests). All participants were evaluated in person by a neuropsychologist and a physician with expertise in the evaluation of older individuals with cognitive impairment. On the basis of physician evaluation and review of the cognitive testing and the neuropsychologist’s opinion, participants were classified with respect to AD and other common conditions with the potential to impact cognitive function according to the recommendations of the joint working group of the National Institute of Neurologic and Communicative Disorders and Stroke and the Alzheimer’s Disease and Related Disorders Association.\textsuperscript{19} In this article, we focused on 50 participants, clinically defined at baseline (on study entry) as being cognitively healthy (\textit{n} = 25), having mild cognitive impairment (\textit{n} = 18), and having probable AD (\textit{n} = 7) (Table), with concurrent antemortem MR imaging and postmortem neuropathologic assessments. The Rush University Medical Center institutional review board approved the study, and all participants gave written informed consent and signed an Anatomic Gift.

**Demographic information on participants in the current study**

<table>
<thead>
<tr>
<th></th>
<th>Healthy (\textit{n} = 25)</th>
<th>MCI (\textit{n} = 18)</th>
<th>AD (\textit{n} = 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Female</td>
<td>64%</td>
<td>55%</td>
<td>71%</td>
</tr>
<tr>
<td>Education (yr)</td>
<td>15.6 (3.6)</td>
<td>14.9 (3.2)</td>
<td>13.2 (1.2)</td>
</tr>
<tr>
<td>Age at MRI (yr)</td>
<td>87.4 (5.0)</td>
<td>88.3 (5.4)</td>
<td>85.2 (4.4)</td>
</tr>
<tr>
<td>Age at death (yr)</td>
<td>90.3 (4.7)</td>
<td>91.1 (5.4)</td>
<td>88.7 (4.9)</td>
</tr>
<tr>
<td>Years between MRI and death</td>
<td>2.8 (1.3)</td>
<td>2.8 (1.2)</td>
<td>3.6 (1.5)</td>
</tr>
<tr>
<td>Entorhinal cortex thickness</td>
<td>1.50 (0.30)</td>
<td>1.49 (0.27)</td>
<td>1.40 (0.41)</td>
</tr>
<tr>
<td>EC tangle density</td>
<td>8.9 (6.2)</td>
<td>12.8 (10.3)</td>
<td>20.1 (12.9)</td>
</tr>
<tr>
<td>EC amyloid-(\beta) load</td>
<td>5.5 (3.1)</td>
<td>7.1 (5.3)</td>
<td>7.5 (4.9)</td>
</tr>
</tbody>
</table>

**Note:** MCI indicates mild cognitive impairment. All values are expressed as mean (SD).

**Imaging Assessments**

We assessed previously obtained T1-weighted anatomic data by using a 1.5T MR imaging scanner (GE Healthcare, Milwaukee, Wisconsin). For the current study, all antemortem MR imaging data were acquired by using a 3D magnetization-prepared rapid acquisition of gradient echo sequence with the following parameters: TE = 2.8 ms, TR = 6.3 ms, preparation time = 1000 ms, flip angle = 8\(^o\), FOV = 24 \times 24 cm, 160 sections, 1-mm section thickness, a 224 \times 192 acquisition matrix reconstructed to 256 \times 256, and 2 repetitions. The MR imaging data were automatically segmented with FreeSurfer 5.0 (http://surfer.nmr.mgh.harvard.edu; for additional details see McKhann et al\textsuperscript{17}). Here, we focused on intracranial volume–corrected average thickness of the entire entorhinal cortex and posterior parahippocampal gyrus (average of the left and right hemispheres) as delineated with our previously developed automated cortical parcellation atlas (Desikan-Killiany atlas in Freesurfer) (Fig 1).\textsuperscript{11} In secondary analyses, we also evaluated baseline intracranial volume–corrected hippocampal volumes (average of the left and right hemispheres).\textsuperscript{18}

**Neuropathologic Assessments**

We used results from previously obtained neuropathologic evaluations and focused on amyloid and neurofibrillary tangle pathology within the entorhinal cortex and medial temporal (entorhinal and hippocampal) regions (for additional details on neuropathologic measures from the Rush Memory and Aging Project/Rush University Medical Center, please see Bennett et al\textsuperscript{21,22} and Barnes et al\textsuperscript{23}). Briefly, at least 2 tissue blocks from the entorhinal cortex and hippocampus (CA1/subiculum) were dissected from 1-cm coronal slabs fixed for 48–72 hours in 4% paraformaldehyde, embedded in paraffin, and cut into 20-\(\mu\)m sections. Amyloid-\(\beta\) was labeled with MO0872 (1:100; Dako,
Carpinteria, California), and paired helical filament \(\tau\) was labeled with AT8 (1:800 in 4% horse serum; Innogenex, San Ramon, California), an antibody specific for phosphorylated \(\tau\). Images of amyloid-\(\beta\)-stained sections were captured for quantitative analysis by using a systematic random-sampling scheme, and calculation of the percentage area occupied by amyloid-\(\beta\) immunoreactive pixels was performed. Quantification of tangle density per square millimeter was performed with a stereologic mapping station. We used composite summary measures of the percentage area occupied by amyloid-\(\beta\) and the density of neurofibrillary tangles by averaging the values for each lesion within the entorhinal cortex and medial temporal regions (entorhinal cortex and hippocampus) (for additional details see Bennett et al\(^{21,22}\) and Barnes et al\(^{23}\)). To evaluate neuropathology within the neocortex and to minimize multiple comparisons, we used a composite measure of tangle density and amyloid-\(\beta\) load within the midfrontal cortex, inferior temporal gyrus, inferior parietal cortex, calcarine cortex, cingulate region, and superior frontal gyrus (cortical tangle density and cortical amyloid-\(\beta\) load).

**Statistical Analysis**

Using linear regression, we evaluated the association between entorhinal cortex thickness and average tangle and amyloid-\(\beta\) load within the entorhinal cortex and medial temporal regions (entorhinal cortex + hippocampus). We also evaluated the relationship between entorhinal cortex thickness and cortical tangle density and amyloid-\(\beta\) load. In secondary analyses, we assessed the association between (posterior) parahippocampal gyrus thickness and tangles and amyloid-\(\beta\) load within the entorhinal cortex and medial temporal regions. In all analyses, we controlled for the effects of age at death, sex, and clinical diagnosis.

**RESULTS**

We found a relationship between ante-mortem entorhinal cortex thickness and postmortem tangle density within the entorhinal cortex (\(\beta\)-coefficient = \(-11.04\), standard error [SE] = 3.78, \(P\) value = .006) and medial temporal regions (\(\beta\)-coefficient = \(-16.67\), SE = 5.04, \(P = .002\)); lower EC thickness was associated with increased EC and medial temporal lobe tangle density (Fig 2). Even after controlling for the effects of hippocampal volume, the relationship between entorhinal cortex thickness and tangle density within the entorhinal cortex (\(\beta\)-coefficient = \(-9.36\), SE = 2.57, \(P = .03\)) and medial temporal regions (\(\beta\)-coefficient = \(-12.89\), SE = 5.64, \(P = .02\)) remained significant. We found no relationship between entorhinal cortex thickness and amyloid-\(\beta\) load within the entorhinal cortex (\(\beta\)-coefficient = \(-4.69\), SE = 2.70, \(P = .09\)) and medial temporal regions (\(\beta\)-coefficient = \(-3.27\), SE = 1.90, \(P = .09\)) (Fig 2).

We found a relationship between entorhinal cortex thickness and both cortical tangle density (\(\beta\)-coefficient = \(-37.5\), SE =
DISCUSSION

Our results demonstrate that quantitative in vivo volumetric MR imaging measurements of the EC are associated with postmortem measures of entorhinal and neocortical AD pathology. Specifically, these results indicate that lower EC thickness predicts higher postmortem tangle load. We also found a similar association between EC thickness and postmortem tangle load within the medial temporal regions (EC + hippocampus). Most important, this effect remained significant after controlling for hippocampal volume. Finally, we found a robust relationship between EC thickness and cortical tangle and amyloid-β pathology. Rather than representing a specific measure of EC tangle pathology, our combined findings suggest that an antemortem MR imaging measure of the entorhinal cortex likely captures Alzheimer-associated pathology within the medial temporal and neocortical regions.

Using neuropathologic assessments from the same individuals and building on our prior work, this study suggests that EC-associated cortical thinning in AD may represent a marker of τ-associated pathology within both the entorhinal cortex and neocortex. Consistent with the known pattern of amyloid-β deposition in AD, which heavily involves the neocortex with relative sparing of the entorhinal cortex, an association was found between EC thickness and neocortical amyloid-β load but not entorhinal amyloid pathology, compatible with the hypothesis that amyloid deposition within the neocortex, rather than the EC may represent an early component of Alzheimer pathobiology.24,25 Our results also suggest that entorhinal cortex thickness may provide independent information about AD pathology even after accounting for hippocampal volume, further illustrating the importance of evaluating EC thickness as an early marker of in vivo Alzheimer neurodegeneration. Most interesting, even among cognitively healthy older adults, we found a selective relationship between entorhinal cortex thickness and tangle pathology but not amyloid pathology, suggesting the potential usefulness of quantitative MR imaging measures in preclinical AD.

Accumulating evidence suggests that τ pathology is closely associated with cognitive performance, particularly in the early stages of disease.20,27 These findings suggest that automated measures of entorhinal cortex atrophy may reflect regional τ pathology, which will be clinically useful for early AD detection and disease monitoring. Additionally, quantitative measures of EC and neocortical structures may be combined with genetic, fluid (CSF or plasma), and cognitive parameters for risk stratification, which may become increasingly relevant for AD prevention and therapeutic trials. With the advent and use of novel agents for detecting in vivo τ deposition, volumetric MR imaging could be integrated with PET imaging to determine whether regional measures of entorhinal atrophy and τ deposition provide independent or complementary information. Finally, beyond AD, automated assessments of the entorhinal cortex can be useful in other disorders of the medial temporal lobe such as medial temporal sclerosis, vascular dementia, and frontotemporal lobar degeneration.

A potential limitation of our study is that our imaging and neuropathology datasets were not coregistered; thus, this drawback limits the precise correspondence between antemortem and postmortem measures of both EC and neocortical τ and amyloid deposition.

The use of MR imaging to evaluate entorhinal tissue is important for two main reasons. First, the EC may represent an early component of Alzheimer pathobiology and may provide independent information about AD pathology even after accounting for hippocampal volume. Second, automated assessments of the EC may be easily integrated with PET imaging to determine whether regional measures of EC atrophy and τ deposition provide independent or complementary information.
postmortem definitions of the EC. Another limitation is the need for validation of our results in independent, community-based samples.

CONCLUSIONS
We found a strong association between an automated, antemortem MR imaging–based measure of EC thickness and postmortem neurofibrillary tangle burden within the entorhinal cortex, medial temporal lobe, and neocortex. We additionally detected a relationship between EC thickness and mild cognitive impairment and Alzheimer’s disease. Brain 2009; 132:2048–57 CrossRef Medline


Barnes LL, Schneider JA, Boyle PA, et al. Memory complaints are related to Alzheimer disease pathology in older persons. Neurology 2006;67:1581–85 CrossRef Medline


ABSTRACT

BACKGROUND AND PURPOSE: Both diffusion tensor imaging and the apparent transverse relaxation rate have shown promise in differentiating Parkinson disease from atypical parkinsonism (particularly multiple system atrophy and progressive supranuclear palsy). The objective of the study was to assess the ability of DTI, the apparent transverse relaxation rate, and their combination for differentiating Parkinson disease, multiple system atrophy, progressive supranuclear palsy, and controls.

MATERIALS AND METHODS: A total of 106 subjects (36 controls, 35 patients with Parkinson disease, 16 with multiple system atrophy, and 19 with progressive supranuclear palsy) were included. DTI and the apparent transverse relaxation rate measures from the striatal, midbrain, limbic, and cerebellar regions were obtained and compared among groups. The discrimination performance of DTI and the apparent transverse relaxation rate among groups was assessed by using Elastic-Net machine learning and receiver operating characteristic curve analysis.

RESULTS: Compared with controls, patients with Parkinson disease showed significant apparent transverse relaxation rate differences in the red nucleus. Compared to those with Parkinson disease, patients with both multiple system atrophy and progressive supranuclear palsy showed more widespread changes, extending from the midbrain to striatal and cerebellar structures. The pattern of changes, however, was different between the 2 groups. For instance, patients with multiple system atrophy showed decreased fractional anisotropy and an increased apparent transverse relaxation rate in the subthalamic nucleus, whereas patients with progressive supranuclear palsy showed an increased mean diffusivity in the hippocampus. Combined, DTI and the apparent transverse relaxation rate were significantly better than DTI or the apparent transverse relaxation rate alone in separating controls from those with Parkinson disease/multiple system atrophy/progressive supranuclear palsy; controls from those with Parkinson disease; those with Parkinson disease from those with multiple system atrophy/progressive supranuclear palsy; and those with Parkinson disease from those with multiple system atrophy; but not those with Parkinson disease from those with progressive supranuclear supranuclear palsy or those with progressive supranuclear palsy.

CONCLUSIONS: DTI and the apparent transverse relaxation rate provide different but complementary information for different parkinsonisms. Combined DTI and apparent transverse relaxation rate may be a superior marker for the differential diagnosis of parkinsonisms.

DIFFERENTIATING these diseases on the basis of clinical symptoms alone can be challenging.1,2 Despite similar clinical symptoms, each 

Received September 22, 2016; accepted after revision January 11, 2017.
From the Departments of Neurology (G.D., M.M.L., N.W.S., L.H., X.H.), Radiology (S.K., X.H.), Pharmacology (M.M.L., X.H.), Neurosurgery (X.H.), Bioengineering (X.H.), and Public Health Sciences (L.K.), Pennsylvania State University–Milton S. Hershey Medical Center, Hershey, Pennsylvania; Department of Public Health (L.H.), Shanxi Medical University, Taiyuan, China; and Department of Statistics (L.L.), Pennsylvania State University, University Park, Pennsylvania.
This work was supported in part by the National Institute of Neurological Disorders and Stroke (NS082151); the Hershey Medical Center General Clinical Research Center (National Center for Research Resources, grant UL1 RR033184); the National Center for Advancing Translational Sciences, grant UL1 TR000127; the National Institute on Drug Abuse (PS0 DA019383 and PS0 DA036107), the National Nature Science Foundation of China (16690031 and 16690015), the Pennsylvania Department of Health Tobacco CURE Funds, the Michael J. Fox Foundation for Parkinson’s Research, the Alzheimer’s Association, Alzheimer’s Research UK, and the Weston Brain Institute.
Please address correspondence to Guangwei Du, MD, PhD, Department of Neurology, 500 University Dr, Hershey, PA 17033; e-mail: guangweidu@hmc.psu.edu; Xumei Huang, MD, PhD, Department of Neurology, 500 University Dr, Hershey, PA 17033; e-mail: xumei@psu.edu.
Indicates open access to non-subscribers at www.ajnr.org.
Indicates article with supplemental on-line appendix and table.
Indicates article with supplemental on-line photos.
http://dx.doi.org/10.3174/ajnr.A5136
disorder has distinct gross and microscopic pathologies. PD is marked by the loss of dopamine neurons in the substantia nigra (SN). 3 MSA is characterized neuropathologically by glial and neuronal cytoplasmic inclusions in many basal ganglia and cerebellar related structures, 4 whereas PSP has neuronal loss, gliosis, and neurofibrillary tangles in both the basal ganglia and cerebellum that may extend to limbic areas. 5, 6

Two MR imaging modalities, diffusion tensor imaging and the apparent transverse relaxation rate (R2*), have been studied intensively in recent decades with the goal of detecting the distinct pathologic patterns in PD, MSA, and PSP and differentiating them from each other. 7-13 DTI has been suggested to reflect the disruption of microstructural integrity (eg, cell death and associated myelin changes), whereas R2* has been used to estimate iron accumulation in brain tissue. 14, 15 There has been little effort, however, to directly compare DTI and R2* in the differential diagnosis of PD and atypical parkinsonism, and in testing whether they can provide complementary information regarding pathology and/or discriminability of those diseases. 16, 17

In the current study, we compared the pattern of DTI and R2* changes among the different parkinsonian diseases and a control group in multiple ROIs that included striatal-, midbrain-, limbic-, and cerebellar-related structures. The performance of DTI, R2*, and their combination to discriminate controls from patient groups and patient groups from each other also was assessed by using an Elastic-Net machine learning approach with a nested 10-fold cross-validation.

MATERIALS AND METHODS

Subjects

A total of 106 individuals (16 with MSA parkinsonian subtype [MSA-P], 19 with PSP [13 with Richardson subtype and 6 with parkinsonian subtype], 35 with PD, and 36 healthy controls) were included in this study from an ongoing longitudinal case-control cohort established in 2012. Patients were recruited from a tertiary movement disorders clinic, and controls were recruited from the spouse population of the clinic or the local community. All patients were free of major neurologic/medical issues other than PD, MSA-P, or PSP, and all controls were free of any known neurologic/psychiatric diagnoses. Patient diagnoses were initially established according to published criteria 16-18 by a movement disorder specialist and updated (August 2016) before the analysis of the current data according to the most recent clinical assessment and postmortem pathology if available (5 PD and 3 PSP cases were confirmed by postmortem pathology results). Two subjects (1 with PD and 1 with PSP) were excluded from later analyses due to severe motion artifacts. Disease duration was defined as the number of years between the date when a parkinsonian syndrome was first diagnosed by a medical professional and the study visit date. All participants were administered the Movement Disorder Society Unified Parkinson’s Disease Rating Scale part III (UPDRS-III) for motor function assessment and the Montreal Cognitive Assessment (MoCA) for global cognitive function. 19 UPDRS-III and MoCA scores and MR imaging scans were collected for patients in an “on” state. The study was approved by the institutional review board at the Pennsylvania State University-Milton S. Hershey Medical Center. All subjects provided written informed consent.

MR Imaging Data Acquisition

Brain MRIs were obtained from all participants by using a 3T MR imaging system (Magnetom Trio; Siemens, Erlangen, Germany) with an 8-channel phased array head coil. The MR imaging examination included multi-gradient-echo (for R2*) and diffusion tensor imaging sequences, along with high-resolution T1-weighted and T2-weighted images for segmentation. Detailed imaging parameters are described in the On-line Appendix.

DTI and R2* Maps

Diffusion tensor images were processed using DTIPrep (Neuro Image Research and Analysis Laboratory, University of North Carolina, Chapel Hill, North Carolina). In DTIPrep, a thorough quality control for diffusion-weighted images was performed by intersection and intervolume correlation analysis, eddy currents, and motion artifact correction. Fractional anisotropy (FA) and mean diffusivity (MD) maps were then estimated for subsequent analysis.

For R2*, an affine registration was used to align 6 magnitude images to an averaged mean magnitude image for potential head motion correction in multi-gradient-echo images. The R2* maps then were generated by using a voxelwise nonlinear Levenberg-Marquardt algorithm to fit a monoexponential function ($s = s_0 e^{-TE \times R^2*}$) by using an in-house Matlab (MathWorks, Natick, Massachusetts) tool.

ROI Segmentation

The segmentation of ROIs was performed by using the Advanced Normalization Tools software package (ANTS; http://stnava.github.io/ANTS/) 20 and an atlas-based segmentation pipeline implemented in AutoSeg (http://www.nitrc.org/projects/autoseg/), 21 along with an in-house atlas. An unbiased, age-appropriate template was generated from T1-weighted images from all controls with ANTs. 22 The following 13 ROIs, including striatal and related structures (putamen [PUT], caudate nucleus [CN], and globus pallidus), midbrain (anterior SN, posterior SN, red nucleus [RN], and subthalamus nucleus [STN]), limbic (hippocampus and amygdala), and cerebellar structures (dentate nucleus, cerebellar hemisphere, superior cerebellar peduncle, and middle cerebellar peduncle) were defined on the cohort-specific T1-weighted and T2-weighted templates by an experienced neuroimager (G.D.). Segmented ROIs are illustrated in On-line Fig 1. ROIs for each subject were then parcellated by using AutoSeg with ANTs as a warping option 21, 23 (see the On-line Appendix for details regarding the segmentation process). On-line Fig 3 illustrates the segmentation quality for small structures (SN, RN, and superior cerebellar peduncle).

B0 images for DTI and mean magnitude images for R2* then were coregistered to individual T2-weighted images using ANTs. The resulting transformations were then applied to FA, MD, and R2* maps by using a B-spline interpolation to bring FA, MD, and R2* images into the same space as the segmented ROIs, where the mean values of FA, MD, and R2* for each ROI were calculated for subsequent analyses.

Statistical Analysis and Modeling

The difference in sex frequency among groups was evaluated by using the $\chi^2$ test. Age and disease duration were compared by using 1-way analysis of variance. MoCA and UPDRS-III scores
among groups were assessed by using 1-way analyses of covariance with adjustments for age and sex.

Each MR imaging measurement in patients with PD, MSA-P, and PSP was compared with that of controls by using univariate ANCOVAs with age and sex as covariates for each of the 13 ROIs. For MR imaging measurements, the Bonferroni method was used to correct for multiple comparisons, with a resulting *P* value ≤ .0038 (0.05/13 independent tests) considered significant.

One major challenge for multimodal MR imaging studies is the high dimensionality of potential predictors generated from different MR imaging measurements and brain structures, which can result in overfitting and collinearity among variables, causing traditional analyses to fail. In this study, we used an Elastic Net regularized logistic regression approach with a nested 10-fold cross-validation scheme to unravel the high-dimensional problem. Two hyperparameters need to be defined in Elastic-Net regularized regression. In our study, *α* was fixed to 0.2 empirically and *λ* was selected by an inner layer 10-fold cross-validation that was independent of the outer layer 10-fold cross-validation used for performance evaluation. This nested cross-validation setting was implemented to alleviate potential overfitting.24

Regularized logistic models were built from all ROI measurements including R2*, DTI (including both FA and MD), and the combined measures (R2*, FA, and MD) for discriminating the following: 1) controls from those with PD/MSA-P/PSP, 2) those with PD from those with MSA-P/PSP, 3) controls from those with PD, 4) those with PD from those with MSA-P, 5) those with PD from those with PSP, and 6) those with MSA-P from those with PSP. Receiver operating characteristic (ROC) curves were generated by using outer layer 10-fold cross-validation models for each MR imaging technique and their combination. A bootstrap approach was used to test the differences among ROC curves.25 ROC curve comparisons were performed between the combined marker and DTI because DTI was better or equal to R2* in all 6 scenarios mentioned above. Sensitivity, specificity, positive predictive value, and negative predictive value were generated by using the Youden method.

Statistical analyses were performed by using the open-source statistical software package R (Version 3.0.3; http://www.r-project.org). Elastic-Net regularized logistic regression was conducted by using the R package glmnet (http://web.stanford.edu/~hastie/glmnet/glmnet_alpha.html),26 whereas the ROC curve analyses were performed by using the R package pROC (https://cran.r-project.org/web/packages/pROC/index.html).27

### RESULTS

#### Demographic Data

Demographic characteristics for subjects are shown in Table 1. No significant overall differences in sex distribution or age were detected among the control, PD, PSP, and MSA-P groups. Post hoc pair-wise analysis showed trending differences between MSA-P and PSP in both sex (*P* = .072) and age (*P* = .065).

#### DTI and R2* Comparison between Parkinsonian Disease and Control Groups

Compared with controls, patients with PD showed changes in the posterior SN and RN in both DTI and R2*, though only the R2* value in the RN survived correction for multicomparsions. Patients with both MSA and PSP showed more widespread changes (after correction for multicomparsions) involving structures both within and outside the midbrain. The pattern of changes, however, was different between the 2 groups. Namely, patients with MSA-P showed increased MD values in the PUT, globus pallidus, cerebellum, and middle cerebellar peduncle, a decreased FA value in the STN, and increased R2* values in the STN and middle cerebellar peduncle. Patients with PSP, however, showed increased MD and R2* values in the posterior substantia nigra but no changes in the STN or any other basal ganglia structures. Patients with PSP had significantly increased MD values in the dentate nucleus, cerebellum, and superior cerebellar peduncle, but not in the middle cerebellar peduncle (Table 2 and On-line Table).

#### Discriminative Analysis

We compared the discriminative ability of DTI and R2* measures and their combination under 6 different scenarios by using Elastic-Net regularized logistic regression and ROC curves (Table 3 and On-line Fig 2). The combined models (DTI + R2*) were better than DTI or R2* alone (*P* < .05) in discriminating controls from those with PD/MSA-P/PSP, controls from those with PD, those with PD from those with MSA-P/PSP, and those with PD from those with MSA-P. When we considered the separation of controls from subjects with PD, the combined model was improved dramatically compared with either measure alone (from area under the curve = 0.82 to area under the curve = 0.91, *P* = .001).

The DTI model, however, showed strong discriminability when differentiating PD from PSP (area under the curve = 0.97) or MSA-P from PSP (area under the curve = 0.96), and adding R2* did not significantly improve the performance of the model. Nevertheless, R2* alone showed decent discriminative ability.
DISCUSSION

First, we confirmed that DTI and R2* differentiate parkinsonian syndromes and controls. In addition, our studies demonstrated that DTI and R2* can capture the distinct pathologic patterns of the different parkinsonian syndromes and may provide complementary information about each disease. Individually, DTI showed better discriminability among the disease groups, whereas R2 added significant value in separating controls from those with parkinsonian syndromes and those with PD from those with MSA-P/PSP or MSA-P.

**DTI and R2* Changes in PD**

The pathologic hallmark of PD is neuronal loss in the SN pars compacta. Our study may capture this pathology by demonstrating decreased FA and increased R2* in the posterior SN.28,29 The inclusion of additional ROIs in our study, however, requires a rather conservative Bonferroni correction; thus, the detected difference did not reach statistical significance. Future studies are needed to confirm these findings in light of a recent meta-analysis suggesting that nigral FA changes in patients with PD vary widely.30 In the current study, patients with PD also demonstrated increased R2* values in the RN. This result is consistent with the notion that the RN may be involved in the primary cerebellar motor pathway, which has been shown to be affected in PD.31,32

**DTI and R2* Changes in MSA-P**

We also found significantly increased MD values in the PUT, globus pallidus, cerebellum, and middle cerebellar peduncle of patients with MSA-P, consistent with previous neuroimaging results.7,12,33,34 On the basis of previous studies, DTI MD changes in the CN have been controversial. For example, Seppi et al35 reported significantly increased MD values in the CN, whereas others have found no changes in CN MD values.12,34 We did not find significant MD changes in the CN, consistent with these later reports. One study reported MD changes in the SN of patients with MSA-P;12 however, we could not replicate this finding. Pathology studies have reported robust changes in the PUT but more variable changes in other basal ganglia regions.4,36 This varying pathology may contribute partly to the inconsistent DTI findings in the CN and SN in the current study and previous ones.9,12,13,34,35

<table>
<thead>
<tr>
<th>Table 2: Individual MRI measurements in PD, MSA-P, and PSP compared with controls in different structures</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Striatal and related structures</strong></td>
</tr>
<tr>
<td>PUT</td>
</tr>
<tr>
<td>CN</td>
</tr>
<tr>
<td>GP</td>
</tr>
<tr>
<td>Midbrain structures</td>
</tr>
<tr>
<td>antSN</td>
</tr>
<tr>
<td>postSN</td>
</tr>
<tr>
<td>RN</td>
</tr>
<tr>
<td>STN</td>
</tr>
<tr>
<td>Limbic structures</td>
</tr>
<tr>
<td>Hipp</td>
</tr>
<tr>
<td>AM</td>
</tr>
<tr>
<td>Cerebellar structures</td>
</tr>
<tr>
<td>DN</td>
</tr>
<tr>
<td>CB</td>
</tr>
<tr>
<td>SCP</td>
</tr>
<tr>
<td>MCP</td>
</tr>
<tr>
<td><strong>Table 3: ROC analysis of individual and combined MRI modalities</strong></td>
</tr>
<tr>
<td><strong>AUC</strong></td>
</tr>
<tr>
<td>C vs PD/MSP-P/PSP</td>
</tr>
<tr>
<td>DTI + R2*</td>
</tr>
<tr>
<td>DTI</td>
</tr>
<tr>
<td>R2*</td>
</tr>
<tr>
<td>PD vs MSA-P/PSP</td>
</tr>
<tr>
<td>R2*</td>
</tr>
<tr>
<td>PD vs MSA-P</td>
</tr>
<tr>
<td>DTI + R2*</td>
</tr>
<tr>
<td>DTI</td>
</tr>
<tr>
<td>R2*</td>
</tr>
<tr>
<td>MSA-P vs PSP</td>
</tr>
</tbody>
</table>

**Note:**—AntSN indicates anterior substantia nigra; postSN, posterior substantia nigra; Hipp, hippocampus; AM, amygdala; CB, cerebellum; DN, dentate nucleus; GP, globus pallidus; MCP, middle cerebellar peduncle; STN, superior cerebellar peduncle.

<table>
<thead>
<tr>
<th><strong>Table 3: ROC analysis of individual and combined MRI modalities</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AUC</strong></td>
</tr>
<tr>
<td>C vs PD/MSP-P/PSP</td>
</tr>
<tr>
<td>DTI + R2*</td>
</tr>
<tr>
<td>DTI</td>
</tr>
<tr>
<td>R2*</td>
</tr>
<tr>
<td>PD vs MSA-P/PSP</td>
</tr>
<tr>
<td>R2*</td>
</tr>
<tr>
<td>PD vs MSA-P</td>
</tr>
<tr>
<td>DTI + R2*</td>
</tr>
<tr>
<td>DTI</td>
</tr>
<tr>
<td>R2*</td>
</tr>
<tr>
<td>MSA-P vs PSP</td>
</tr>
</tbody>
</table>

**Note:**—Sens indicates sensitivity; Spec, specificity; PPV, positive predictive value; NPV, negative predictive value; AUC, area under the curve; C, controls.

*ROC curves were compared between the models, including all MRI measurements and that with DTI measurements only.

**Values**

when differentiating PD from PSP (area under the curve = 0.87) and MSA-P from PSP (area under the curve = 0.89).
Patients with MSA-P consistently demonstrated increased R2* values in the PUT.10-12 The current study, however, failed to detect R2* changes in the PUT of these patients. Although the exact reason for the discrepancy is unknown, we postulate the following 2 possibilities: First, heterogeneous cohort characteristics may have contributed to the different results. For example, previous studies had significantly younger patients with MSA (mean ages, 58–62 years) compared with our cohort (mean age, 68 years). Age significantly affects iron and R2* values in basal ganglia structures.27 Thus, these age effects may mask the disease-related changes in the PUT. Second, the different R2* techniques used among the studies may influence the results.10,12,38 For example, Lee et al11 used 8 echoes and a TR = 24 ms, whereas Barbagallo et al12 used 6 echoes with repetition and a TR = 100 ms; and we used 6 echoes and a TR = 54 ms. In addition to imaging parameters, each study used different curve-fitting techniques: Lee et al11 used linear fitting after log-transformation of the original signal, whereas the current study used nonlinear curve-fitting to a monoexponential function similar to that in Barbagallo et al.12

Most interesting, we detected a decreased FA value in the STN of patients with MSA-P, along with an increased R2* value, which has not been reported by any previous MR imaging studies, to our knowledge. It is unclear whether the lack of significant STN findings arises from a lack of focus on this structure or whether no differences were found. The neuronal/glial cytoplasmic inclusions that typically are found in basal ganglia regions are less common in the STN of patients with MSA.4 One pathology study, however, noted increased microglia in the STN of patients with MSA-P,46 which may reflect a reactive or compensatory process instead of the primary pathology. Thus, the STN changes we detected may reflect these reactive or compensatory changes, though future studies focused on the STN are warranted to verify this.

**DTI and R2* Changes in PSP**

Consistent with previous studies, we found significant DTI (MD) changes in midbrain (posterior SN and cerebellar [cerebellum and superior cerebellar peduncle]) structures of patients with PSP, with the most robust change seen in the superior cerebellar peduncle.7,8,35,39 Whereas most studies reported increased MD values in the PUT of patients with PSP,35,39,40 we did not detect MD changes in the PUT or other basal ganglia structures (CN and globus pallidus) in the current study. Consistent with our findings, Tsukamoto et al44 reported no MD changes in the PUT of patients with PSP. Additional studies are needed to clarify the discrepancies.

In the past, both pathologic and neuroimaging studies with free-water imaging suggested changes in the STN of patients with PSP.6,11 Pathologic studies also reported both neuronal and oligodendroglia loss in the STN of patients with PSP. Using traditional DTI measures (FA and MD), the current study did not detect significant changes in the STN of patients with PSP. It is possible that the mixed microscopic pathology may have complex or opposing effects on these traditional DTI measurements at the macroscopic level. Change in the STN of patients with PSP by means of the free-water measure derived from a bi-tensor model13 suggests that free-water may be a more sensitive marker for PSP-related pathology in the STN. Future studies are needed to further confirm the links between PSP-related pathology and different MR imaging contrasts.

In the current study, we also detected an increased MD value in the dentate nucleus of patients with PSP. Although this finding is new, it is in line with pathologic results of neuronal loss in the dentate nucleus of patients with PSP.6 In addition, patients with PSP demonstrated significantly increased MD values in the hippocampus and a trending change in the amygdala. These results are consistent with previous volumetric studies suggesting pathologic involvement of the hippocampus in PSP5,41 and early cognitive issues that often are detected in patients with PSP clinically. These findings are inconsistent, however, with previous pathologic studies indicating that the hippocampus and amygdala are spared from pathology in patients with PSP.12 A growing literature supports the heterogeneity of PSP and mixed pathologic findings across different tauopathies39,43; thus, the value of using differential imaging patterns to subtype the patient with PSP will be evaluated in the future.

Previous studies on R2* in the PUT, CN, and globus pallidus in patients with PSP have been controversial because some studies showed significantly increased R2* values in these structures,11,44 whereas others did not.10 The current results are consistent with no R2* changes in the PUT, CN, and globus pallidus. Patients with PSP, however, had significantly increased R2* values in the SN and RN. This finding is consistent with previous PSP pathologic studies indicating that τ pathology–related neuronal and oligodendroglia loss is involved in both the SN and RN.42

**Discriminative Analysis**

Many promising MR imaging markers have been suggested to differentiate patients with PD from those with atypical parkinsonism.8,9,13,45,46 Systematic comparison and validation of those markers in the same subjects are needed before translating these findings into a clinical setting. The current study is the first to systematically compare DTI, R2*, and their combination by using Elastic-Net regularized logistic regression. When we compared DTI and R2* measures under 6 clinically relevant scenarios, our results suggested the following: 1) that DTI measures overall are better or comparable with R2* values in differentiating parkinsonisms, and 2) that R2* provides complementary information in most scenarios except when differentiating PD from PSP or MSA-P from PSP.

**Limitations**

The current study has some limitations. First, among 70 patients with parkinsonism, only 8 cases were confirmed by postmortem pathology. Despite updating the clinical diagnosis by integrating more longitudinal clinical information right before conducting the current analysis, diagnosis error inevitably exists and might bias the results. Additionally, we included controls with positive UPDRS-III scores as high as 14. It is possible that controls with high UPDRS-III scores have a preclinical parkinsonian syndrome. Nonetheless, a recent study has demonstrated that parkinsonian signs are common in older adults, even without a clinical diagnosis of disease.47 Second, this study is case-control in nature and does not simulate clinical practice, which would include other diseases potentially confused with PD such as essential tremor,
corticobasal degeneration, dementia with Lewy bodies, and psychogenic disorders. In addition, we did not separate PSP subtypes. Re-analyzing the data to include only patients with PSP Richardson subtype \((n = 13)\) did not change the results demonstrably from those including the entire PSP cohort. Finally, in the current study, all data were collected while patients were on anti-parkinsonian medications, and the MR imaging measures may be affected by the drugs. Further prospective studies that mimic clinical practice are warranted to further test the potential of these markers in clinical practice.

Technically, recent advances in MR imaging markers for PD and atypical parkinsonism have suggested that 2 new measures (free-water and quantitative susceptibility) may be useful for discriminating patient groups and are derived from the same MR imaging data (DTI and \(R2^*\), respectively). Quantitative susceptibility has been suggested to improve the \(R2^*\) signal by reducing imaging data (DTI and \(R2^*\), respectively). Quantitative susceptibility (free-water and quantitative susceptibility) may be useful for differentiating PD and atypical parkinsonism.4,6,8,10,12,14,15,29,34,35,39 The current study did not include these new measures, and future work validating and comparing them is warranted. Finally, this study did not compare our models with conventional MR imaging clues used by radiologists in these disorders, such as the “hummingbird” and “hot cross bun” signs, midbrain atrophy, and putaminal T2-weighted hypointensity.50–52 Notably, Reiter et al.3 with visual rating of dorsolateral nigral hyperintensity in susceptibility-weighted images, showed promising discriminability in differentiating those with parkinsonian syndromes from controls. It will be important to discern the additional value a quantitative MR imaging marker derived from combining DTI and \(R2^*\) provides compared with the best medical knowledge. In this study, we adopted an Elastic-Net regularized regression as the multivariate classification method. Even though we used a nested 10-fold cross-validation for model selection and performance evaluation, the models still may be overly optimistic due to the small sample size.24

CONCLUSIONS

Our findings are consistent with those in previous neuroimaging and postmortem pathologic studies reporting significant involvement of striatal-, midbrain-, and cerebellar-related structures in PD and atypical parkinsonism.4,6,8,10,12,14,15,29,34,35,39 The exact location and MR imaging measures in striatal and midbrain-related structures between previous studies and the current study, however, vary.34,35,39 This study demonstrated that DTI and \(R2^*\) reflect different-yet-complementary information that can be used for discriminating controls and patients with PD, MSA, and PSP. Further refinement of this approach, including the use of novel measures that assess other aspects of disease pathology and the extension to whole-brain feature space, could lead to an optimized tool that can diagnose and differentiate PD from atypical parkinsonism. We envision applying this approach to a large prospective cohort, including a more diverse patient population (PD, MSA, PSP, essential tremor, corticobasal degeneration, and dementia with Lewy bodies), that simulates a real clinical setting to further test its utility in clinical practice.

ACKNOWLEDGMENTS

We thank the MR imaging technical support of Jeffery Vesek and the assistance of study coordinators Melissa Santos and Ragha Clayiff.

Disclosures: Guangwei Du—RELATED: Grant: National Institute of Neurological Disorders and Stroke and Michael J. Fox Foundation for Parkinson’s Research.* Mechelle M. Lewis—RELATED: Grant: National Institute of Neurological Disorders and Stroke, and Michael J. Fox Foundation for Parkinson’s Research UNRELATED: Grants/Grants Pending: National Institute of Environmental Health Sciences*. Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed: Parkinson’s Disease Biomarkers Program (PDBP) Annual Consortium Meeting, Comments: every year, the PDBP meets in Bethesda, Maryland, for 1.5 days. The PDBP pays for hotel accommodations, travel, and per diem meals for me to attend the meeting. Nicholas W. Sterling—UNRELATED: Support for Travel to Meetings for the Study or Other Purposes: American Neurological Association, Comments: travel grant.* Lan Kong—RELATED: Grant: National Institutes of Health and Michael J. Fox Foundation for Parkinson’s Research, Comments: I am the co-investigator on the following grants: National Institutes of Health grant: Multimodal MRI Markers of Nigrostriatal Pathology in Parkinson’s Disease; Michael J. Fox Foundation for Parkinson’s Research: Benchmark Structural MRI Markers for Parkinsonian Syndromes.* Ruize Li—RELATED: Grant: National Science Foundation and National Institutes of Health*. UNRELATED: Grants/Grants Pending: National Science Foundation and National Institutes of Health.* Xuehui Huang—RELATED: Grant: National Institutes of Health, Comments: National Institute of Neurological Disorders and Stroke, and Michael J. Fox Foundation for Parkinson’s Research*, UNRELATED: Consultancy: National Institute of Environmental Health Sciences, Comments: This is payment for my consultation for the National Institute of Environmental Health Sciences about Parkinson disease diagnosis, Grants/Grants Pending: National Institute of Environmental Health Sciences, Comments: NIHES-E029872.* Money paid to the institution.

REFERENCES


ABSTRACT

BACKGROUND AND PURPOSE: Clinical brain MR imaging registration algorithms are often made available by commercial vendors without figures of merit. The purpose of this study was to suggest a rational performance comparison methodology for these products.

MATERIALS AND METHODS: Twenty patients were imaged on clinical 3T scanners by using 4 sequences: T2-weighted, FLAIR, susceptibility-weighted angiography, and T1 postcontrast. Fiducial landmark sites (n = 1175) were specified throughout these image volumes to define identical anatomic locations across sequences. Multiple registration algorithms were applied by using the T2 sequence as a fixed reference. Euclidean error was calculated before and after each registration and compared with a criterion standard landmark registration. The Euclidean effectiveness ratio is the fraction of Euclidean error remaining after registration, and the statistical effectiveness ratio is similar, but accounts for dispersion and noise.

RESULTS: Before registration, error values for FLAIR, susceptibility-weighted angiography, and T1 postcontrast were 2.07 ± 0.55 mm, 2.63 ± 0.62 mm, and 3.65 ± 2.00 mm, respectively. Postregistration, the best error values for FLAIR, susceptibility-weighted angiography, and T1 postcontrast were 1.55 ± 0.46 mm, 1.34 ± 0.23 mm, and 1.06 ± 0.16 mm, with Euclidean effectiveness ratio values of 0.493, 0.181, and 0.096 and statistical effectiveness ratio values of 0.573, 0.352, and 0.929 for rigid mutual information, affine mutual information, and a commercial GE registration, respectively.

CONCLUSIONS: We demonstrate a method for comparing the performance of registration algorithms and suggest the Euclidean error, Euclidean effectiveness ratio, and statistical effectiveness ratio as performance metrics for clinical registration algorithms. These figures of merit allow registration algorithms to be rationally compared.

ABBREVIATIONS: ANTs = advanced normalization tools; AOI = algorithm of interest; CC = cross-correlation; EER = Euclidean effectiveness ratio; LM = landmarks; MI = mutual information; SER = statistical effectiveness ratio; SWAN = susceptibility-weighted angiography; T1C = T1 postcontrast; TRE = target-to-registration error

Image registration is an essential step in the analysis of brain MR imaging data from multiple images because it ensures the spatial correspondence of anatomy across complementary information sources for diagnosis and treatment. Most commercially available MR image–analysis software packages have some implementation of image registration, and such techniques have a thorough, well-documented grounding in the literature.1–6

Research publications about new registration methods for MR images of the brain nearly always include quantitative assessments of their performance, while commercial registration solutions are often released without disclosing the performance metrics of the vendor. Furthermore, due to the proprietary nature of these commercial algorithms, the explicit transformations are often not disclosed; thus, there are relatively few publicly available figures of
Merit to assess the performance of these heavily used commercial solutions that are essential to clinical neuroimaging.

Performance assessments of these products from commercial vendors would be clinically useful, however, such information is seldom available. In a study that used a widely accepted, neurosurgical commercial package, Hoelper et al placed 25 anatomic landmarks in T1 and T2 brain volumes to test the registration error for 39 patients, a rare example of a publicly available assessment for a commercial registration solution. They demonstrated that whole-brain volume registrations could have errors ranging from 0.7 to 2 mm, depending on the region of the brain, and they therefore recommended using a volume of interest to improve local registrations when a particular area was important.7 Such knowledge of the behavior of a commercial product can improve its use in the clinic, but if commercial vendors do not use objective metrics to characterize thoroughly the performance of their products and then make these results available, rational choices and improvements cannot be made. This situation is to the ultimate detriment of the patient undergoing treatments that rely on these algorithms being highly accurate.

Within the academic research world, various prior methods have been used to assess the performance of registration algorithms. External, invasive skull-implanted markers have been used as fiducial landmarks to assess CT-MR imaging and PET-MR imaging registrations,4 the 8 corner voxels of a box around the head have been used to assess MR-MR brain registrations,6 and 256 anatomic landmarks throughout the brain have been used to assess inter-subject MR-MR registrations.9 An on-line data base of MR imaging and sonography brain volumes also contains 19–40 landmarks per patient to assess registration accuracy.10

In addition to fiducial landmarks, other quantitative and semi-quantitative methodologies for assessment have been used. Examples include tissue edge distances as measured by the Haussdorff distance,11 comparing the resulting transformations with the criterion standard reference transformations,12 using the amount of tissue overlap for equivalent regions,13 using image-similarity measures calculated between images,14 and using visual assessments by human observers.15 Extensive neuroimaging algorithm comparison studies have shown that registration performance is minimally affected by the many variations of labeling protocols and overlap measures.9

In light of these examples, there is a real need for simple, objective metrics to serve as figures of merit for clinically used registration algorithms from academic and commercial vendors. The aim of this project was therefore to demonstrate the feasibility of a performance-testing methodology for modern registration algorithms by using fiducial landmark sites in routinely used clinical images. This method was then used to assess the performance of both commercial and open-source registration algorithms as applied to a set of intrasubject, multisequence, MR images of the brain. Internal, anatomic landmark-based fiducials served as the criterion standard against which to measure performance. By calculating several objective metrics of performance from the results of these registrations on clinically acquired data, we show how various methods of registration can be meaningfully compared by end users. We use a limited set of algorithms in demonstration, but any registration algorithm could be substituted and similarly assessed (the authors could be contacted to arrange this).

Materials and methods

This study was a retrospective analysis of data acquired as part of a Health Insurance Portability and Accountability Act–compliant, institutional review board–approved clinical protocol that required signed consent from study participants.

Images

Patients were consecutively recruited on the basis of specific criteria for inclusion (18 years of age or older, candidate for cerebral tumor resection with suspected or biopsy-proved primary brain tumor) and exclusion (prior brain tumor treatment, including surgical resection, radiation therapy, or chemotherapy). From February 2013 to October 2015, 20 patients (mean age, 45.3 years; range, 21–75 years) were imaged for surgical-planning purposes on Signa HDxt 3T or Discovery MR750 3T clinical scanners (GE Healthcare, Milwaukee, Wisconsin). This cohort included 11 women (mean, 40.7 years; range, 21–75 years) and 9 men (mean, 50.9 years; range, 28–67 years). The imaging protocol (On-line Table 1) included a high-resolution T2-weighted scan (voxel size, 0.5469 × 0.5469 × 2 mm), a FLAIR scan (voxel size, 0.5 × 0.5 × 1 mm), a susceptibility-weighted angiography (SWAN) scan (voxel size, 0.3906 × 0.3906 × 1 mm), and a T1 postcontrast (T1C) scan (voxel size, 0.4688 × 0.4688 × 3.5 mm) obtained after injecting 0.1 mmol/kg of either gadopentetate dimeglumine or gadobutrol (Magnevist; Bayer HealthCare Pharmaceuticals, Wayne, New Jersey; or Gadavist; Bayer Schering Pharma, Berlin, Germany, respectively) at 5 mL/s, followed by 30 mL of saline at 5 mL/s. Because 2 separate contrast doses were needed for the scanning session, the total dosage was 20 mL of Magnevist or Gadavist and 60 mL saline. DICOM image files were converted into the NIfTI file format (https://nifti.nimh.nih.gov/nifti-1) with functions from the Insight ToolKit (https://lhncbc.nlm.nih.gov/project/insight-toolkit).16 Images were skull-stripped by using the Brain Extraction Tool (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET)17 from the FMRIB Software Library (fsl.fmrib.ox.ac.uk), followed by manual refinement of the mask with Amira3D (Version 6.0; FEI, Hillsboro, Oregon) and application of the mask by radiologist with 15 years’ clinical experience (D.S.).

Landmarks

Fifteen landmarks (LMs) were manually specified per patient sequence, meaning that 60 independent landmark sites were specified across the 4 imaging sequences. For the entire set of 20 patients, 1200 LM points were planned. Landmarks were placed at anatomically distinct locations across the entire brain volume (Fig 1, On-line Fig 1, and On-line Table 2) by using the Amira3D software. All fiducial sites were reviewed by multiple observers (I.S.L., D.S.) with expertise in neuroanatomy, including a neuroradiologist with 15 years’ clinical experience (D.S.).

Landmark fiducial sites were chosen for their unambiguous appearance across imaging sequences (ie, vessel intersections, bifurcations, inflection points, and unique geometries), so as to
ensure correspondence between homologous sites (On-line Fig 2). The general locations of landmarks were similar across patients, but the exact placements differed from patient to patient because of natural anatomic variations. Landmarks were intended to be spatially distributed within the brain, to capture registration error in many different regions. However, most suitable fiducial sites ended up being in the midaxial area of the brain, due to plentiful, easily identifiable anatomy being located there. Key sites were therefore chosen in extreme anterior, posterior, inferior, and superior locations to round out the placement of landmarks.

The standardized landmark coordinates shown in On-line Table 2 were obtained by registering the T2 volume to the International Consortium for Brain Mapping 152 Nonlinear Symmetric 2009b template, thereby generating coordinates in the right-anterior-superior convention with the anterior commissure as the origin. This template volume was only used for determining these coordinates and was not part of the actual registration experiments.

Registrations
Each patient’s T2 image volume served as the fixed image, and the FLAIR, SWAN, and T1C images served as the moving image for registration procedures.

Landmark registration was performed as a criterion standard reference. Rigid (6 df) and affine (12 df) transformations with the fiducial LM sites as input were performed by using convert3D (c3d). These registrations disregarded all imaging content and focused only on minimizing the gap between corresponding landmarks, thereby creating a lower bound for target-to-registration error (TRE), defined as the Euclidean distance between 2 points in space.

The Volume Viewer software package (Version 11.3 Ext. 14; GE Healthcare) available on the Advantage Workstation Server (Version 2; GE Healthcare) was used for its Integrated Registration module with its specialized Neuro Registration mode. Registrations for GE were performed on a stand-alone server dedicated to this software.

The open-source Advanced Normalization Tools software package (ANTS; http://stnava.github.io/ANTs/) was used to perform multiple registrations for each of the FLAIR-T2, SWAN-T2, and T1C-T2 image pairs, by using both different similarity measures (cross-correlation [CC] versus mutual information [MI]) and different df s for image movement (rigid versus affine). All registrations were performed on a Linux workstation (Xeon X5675 CPU @ 3.07GHz with 24 cores, 96 GB RAM; Intel, Santa Clara, California).
Analysis
Euclidean TRE values between fixed and moving images were calculated at baseline (preregistration) and for 7 different registration experiments (Table 1). To create independent TRE measurements for analysis, we averaged together multiple TRE values within each patient by sequence. For location-dependent analyses, TRE values were averaged across patients on the basis of coordinate locations.

Baseline TRE values were analyzed on the basis of the time of image acquisition and location in space. Postregistration TRE values were also analyzed, with Shapiro-Wilk test results used to determine the appropriate ANOVA test for comparing means of groups for main effects (On-line Fig 3). When we compared TRE values of 8 different groups for a single sequence, if all groups passed the Shapiro-Wilk test, the repeated measures ANOVA was used, with post hoc testing with the Tukey test to adjust values for multiple comparisons (α = .05). Otherwise, the Friedman nonparametric ANOVA was used, with post hoc testing by using the Dunn test (α = .05). Pair-wise comparisons were also performed between the best-performing algorithm for each sequence and its runners-up (On-line Fig 4), with the paired t test used if results for both groups passed the Shapiro-Wilk test. Otherwise, the Wilcoxon signed-rank test was used. All statistical testing was performed by GraphPad Prism software (Version 6.07, 2015; GraphPad Software, San Diego, California) with P < .05 denoting significance.

The Euclidean effectiveness ratio (EER, On-line Fig 5) represents the fraction of Euclidean error remaining after registration, defined as 1 TRE gap (between the results of an algorithm of main interest and the affine LM algorithm [LM12]) divided by another TRE gap (between baseline and LM12 results). The EER scale will always be between 0 and 1 with values at the boundary interpreted as significance. The statistical effectiveness ratio (SER, On-line Fig 5) was calculated as the ratio of 1 Cohen d (between baseline and LM12 results) divided by another Cohen d (between base-line and LM12 results). The Cohen effect size d was calculated between algorithms by using

\[ d_{A-B} = \frac{\Delta TRE_{A-B}}{s_{A-B}} \]

where the 2 algorithms being compared are A and B, the pooled SD, \( s_{A-B} = \sqrt{\frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2}} \), where \( s_A \) is the SD of group A, and \( n_A \) is the number of samples in group A.

RESULTS
Images
Table 2 contains demographic information about patients from this clinical trial. The “high movers” and “low movers” subgroups are defined in “Baseline TRE as a Function of Time.”

Landmarks
Of the 1200 fiducial LM sites planned, 1175 were realized. Because SWAN volumes in 17 of 20 patients had limited superior/inferior head coverage, LM sites 14 and/or 15 had to be omitted, depending on the patient, for a deficit of 25 SWAN landmark sites (n = 10 omitted for LM site 14, and n = 15 omitted for LM site 15). The resulting 1175 points translated to 875 LM pairs (300 for FLAIR-T2, 275 for SWAN-T2, and 300 for T1C-T2). In 3 of 20 patients, the SWAN imaging volume was located too superiorly to include the normal locations of LM sites 1–4 (the next most inferior landmarks, after LM site 14). In these cases, alternative fiducial sites were chosen, superior enough to still allow a valid site pairing with the T2 volume.

Registrations and Analysis
Baseline TRE as a Function of Time. Baseline TRE increased as a function of time in the magnet for the entire population (Fig 2, left), with a best-fit line significantly different from the zero slope (F = 17.80, P < .0001, R² = 0.2348). Using z scores based on T1C TRE values, we dichotomized patients into high movers (z > 0.35) and low movers (z < 0.35). The line for the high movers was significantly different from the zero slope (Fig 2, right; F = 33.64, P < .0001), but the line for the low movers was not (F = 0.004419, P = .9474).

Baseline TRE as a Function of Location. We analyzed the high movers group, and found that their predominant motion appeared to be rotatory in nature, around an axis passing through the dens (On-line Fig 6). TRE values were therefore plotted against their distance from a point on this fulcrum axis (Fig 3). For the T1C sequence (Fig 3, lower row), the baseline TRE increased with the distance from this point, with a

<table>
<thead>
<tr>
<th>Registration Experiment</th>
<th>Inputs</th>
<th>Software</th>
<th>Similarity Measure</th>
<th>df</th>
<th>Outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-registration</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>TRE, EER, SER</td>
</tr>
<tr>
<td>GE</td>
<td>Images</td>
<td>GE</td>
<td>Proprietary</td>
<td></td>
<td>TRE, EER, SER</td>
</tr>
<tr>
<td>Rigid CC</td>
<td>Images</td>
<td>ANTs</td>
<td>CC</td>
<td></td>
<td>TRE, EER, SER</td>
</tr>
<tr>
<td>Rigid MI</td>
<td>Images</td>
<td>ANTs</td>
<td>MI</td>
<td></td>
<td>TRE, EER, SER</td>
</tr>
<tr>
<td>Affine CC</td>
<td>Images</td>
<td>ANTs</td>
<td>CC</td>
<td></td>
<td>TRE, EER, SER</td>
</tr>
<tr>
<td>Affine MI</td>
<td>Images</td>
<td>ANTs</td>
<td>MI</td>
<td></td>
<td>TRE, EER, SER</td>
</tr>
<tr>
<td>Rigid LM</td>
<td>LM</td>
<td>c3d</td>
<td>LM</td>
<td></td>
<td>TRE, EER, SER</td>
</tr>
<tr>
<td>Affine LM</td>
<td>LM</td>
<td>c3d</td>
<td>LM</td>
<td></td>
<td>TRE, EER, SER</td>
</tr>
</tbody>
</table>

Note: —NA indicates not applicable; c3d, Convert3D; GE, GE Volume Viewer.

| Table 2: Age, race, and frontal lobe tumor involvement of patient groups |
|---------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
|                          | All (n = 20)        | Female (n = 11)     | Male (n = 9)        | High Movers (n = 9) | Low Movers (n = 11) |
| Age (mean)               | 45.3 ± 16.4         | 40.7 ± 17.7         | 50.9 ± 14.3         | 38.1 ± 15.3         | 512 ± 15.5          |
| Age range (yr)           | 21–75               | 21–75               | 28–67               | 21–66               | 29–75               |
| Race (White/Black/Hispanic/Asian) (No.) | 15/2/2/1 | 8/2/1/0 | 7/0/1/1 | 6/1/1/1 | 9/1/1/0 |
| No. of patients with tumor with frontal lobe involvement | 11 (55%) | 7 (64%) | 4 (44%) | 8 (89%) | 3 (27%) |
best-fit line significantly different from the zero slope ($F_{H11005}^{4.4212, P_{H11005}^{.0074}}$). However, the best-fit lines for FLAIR ($F_{H11005}^{8.8311, P_{H11005}^{.0634}}$) and SWAN ($F_{H11005}^{1.331, P_{H11005}^{.3322}}$) were not significantly different from the zero slope, suggesting a threshold effect for movement as time increases.

TRE and Euclidean Effectiveness Ratio after Registration. For FLAIR TRE and EER values (Fig 4 upper row and On-line Table 3), the rankings were the following: rigid MI < rigid CC < affine CC < GE < affine MI. Pair-wise TRE comparisons between the best algorithm and its runners-up did not reveal significant differences for the rigid MI–rigid CC pair ($P_{H11005}^{.0552}$ paired $t$, effect size $d_{H11005}^{0.1492}$) or the rigid MI–affine CC pair ($P_{H11005}^{.0532}$ paired $t$, effect size $d_{H11005}^{0.3620}$). Significant differences did exist for the rigid MI–GE pair ($P_{H11005}^{.0323}$ paired $t$, effect size $d_{H11005}^{0.4627}$).

For SWAN TRE and EER values (Fig 4 middle row and On-line Table 3), the rankings were the following: affine MI < GE < rigid MI < affine CC < rigid CC. Pair-wise TRE comparisons between the best algorithm and its runners-up revealed significant differences for the affine MI–GE pair ($P_{H11005}^{.0136}$, Wilcoxon signed-rank test, effect size $d_{H11005}^{0.4969}$) and the affine MI–rigid MI pair ($P_{H11005}^{.0121}$, Wilcoxon signed-rank test, effect size $d_{H11005}^{0.6454}$).

For T1C, TRE, and EER values (Fig 4 lower row and On-line Table 3), the rankings were: GE < affine MI < affine CC < rigid MI < rigid CC. Pair-wise TRE comparisons between the best algorithm and its runners-up did not reveal significant differences for the GE–affine MI pair ($P_{H11005}^{.6640}$ paired $t$, effect size $d_{H11005}^{0.0669}$), but they did reveal significant differences for the GE–affine CC pair ($P_{H11005}^{.0042}$ Wilcoxon signed-rank test, effect size $d_{H11005}^{0.7720}$).

The SER rankings were the same as rankings based on TRE and EER for the FLAIR and SWAN sequences. For the T1C sequence, the SER ranking was different, selecting rigid CC as the top algorithm (On-line Table 4).

TRE Values before Registration: FLAIR versus SWAN versus T1C. Before registration, the mean TRE rankings were as follows: FLAIR < SWAN < T1C. Differences among these 3 groups were statistically significant (Friedman ANOVA, $F_{H11005}^{9.300, P_{H11005}^{.0096}}$), and post hoc testing showed significant differences for the FLAIR–T1C pair, but not for the FLAIR-SWAN or SWAN-T1C pairs ($P_{H11005}^{.0133}, P_{H11005}^{.0531}$, and $P_{H11005}^{>.9999}$, respectively; Dunn test).

TRE Values after the Affine LM Registration: FLAIR versus SWAN versus T1C. Affine LM results, representing the criterion standard minimum possible TRE, consistently had the smallest TRE values
DISCUSSION

The main findings from this study were the following:

1) The FLAIR, SWAN, and T1C image volumes, on average, all had lower TRE values after registrations that corrected for spatial errors due to patient motion. For open-source methods, MI outperformed CC registrations and affine usually outperformed rigid registrations.

2) Spatial error values after registration were comparable with or better than values found in the literature.

3) The unregistered spatial error increased as a function of time in the magnet, and a subpopulation of patients, most with frontal lobe tumor involvement, was responsible for most of the time-dependence of the spatial error.

4) Better correction was possible between sequences with similar planes of acquisition. If images are acquired in different planes, out-of-plane distortion corrections should be applied.

5) The best algorithms with the EER metric were the same as the best algorithms by TRE values.

The FLAIR, SWAN, and T1C image volumes, on average, all had lower TRE values after registrations that corrected for spatial errors due to patient motion, a finding that is compatible with conventional wisdom regarding spatially aligning images before analysis. For open-source methods, MI generally outperformed CC registrations, given the same similarity measures and image sequence; this outcome agrees with existing literature. Affine usually outperformed rigid transformations, given the same similarity measure and image sequence; this finding makes sense, given that the latter is a special case of the former.

The smallest TRE values for FLAIR (1.55 mm), SWAN (1.34 mm), and T1C (1.06 mm) were comparable with or better than average nonlinear registration errors for other body parts and imaging modalities, including CT-CT lung (1.0 mm, 2.05 mm), CT-CT liver (1.8 mm), MR-CT liver (3.9 mm), and MR-MR prostate (2.3 mm).

Spatial error increased as a function of time in the magnet; this finding supports previous observations about image misalignments increasing with time during a scanning session and which likely occurs due to patient restlessness expressed by repositioning of the head. The high movers subpopulation, responsible for most of the time-dependent error, was younger than the overall cohort (18.1 versus 45.3 years), and 8 of 9 patients had tumors with involvement in the frontal lobe, an area of the brain associated with motor impulse control.

The FLAIR, SWAN, and T1C image volumes, on average, all had lower TRE values after registrations that corrected for spatial errors due to patient motion, a finding that is compatible with conventional wisdom regarding spatially aligning images before analysis. For open-source methods, MI generally outperformed CC registrations, given the same similarity measures and image sequence; this outcome agrees with existing literature. Affine usually outperformed rigid transformations, given the same similarity measure and image sequence; this finding makes sense, given that the latter is a special case of the former.

The smallest TRE values for FLAIR (1.55 mm), SWAN (1.34 mm), and T1C (1.06 mm) were comparable with or better than average nonlinear registration errors for other body parts and imaging modalities, including CT-CT lung (1.0 mm, 2.05 mm), CT-CT liver (1.8 mm), MR-CT liver (3.9 mm), and MR-MR prostate (2.3 mm).

Spatial error increased as a function of time in the magnet; this finding supports previous observations about image misalignments increasing with time during a scanning session and which likely occurs due to patient restlessness expressed by repositioning of the head. The high movers subpopulation, responsible for most of the time-dependent error, was younger than the overall cohort (38.1 versus 45.3 years), and 8 of 9 patients had tumors with involvement in the frontal lobe, an area of the brain associated with motor impulse control.

Analysis of error versus location suggests that most head movement during the scanning session occurs as rotation about a fulcrum that is in line with the dens.

Better spatial corrections were possible between sequences with similar planes of acquisition, because differences in native orientations lead to differences in residual, out-of-plane distortions. Specifically, FLAIR had left/right distortions (with the brain pinched along the left/right axis, especially near the superior/inferior edges of the volume), while T2, SWAN, and T1C did not have such distortions (because reverse-pin-cushion-shaped corrections had restored the brain to its proper shape and size).
than the smallest TRE values for SWAN and T1C. Rigid also out-performed affine registrations for the FLAIR sequence, because affine transformations attempted to recreate the nonlinear effects of distortion correction and unintentionally worsened FLAIR-T2 matching. Out-of-plane distortion corrections should therefore be applied by MR imaging scanners, whenever available, to counter this problem.

The best algorithms by EER values (FLAIR, rigid MI; SWAN, affine MI; T1C, GE) were the same as the best algorithms by TRE values, which makes sense given that the EER is a normalization of TRE that preserves the relative rankings of different algorithms while facilitating comparisons across different sequences. The SER, however, gave different rankings for T1C and is a less intuitive metric to interpret but attractive from a statistical perspective due to its incorporation of noise.

Both single-voxel landmarks and labeled regions have been used to assess registration accuracy for various purposes, with the best assessment method ultimately dictated by the application area and, to some degree, the resources available for the time-intensive, manual label, and/or landmark dataset curation process to create the criterion standard control. We believe that single-voxel landmarks were an appropriate choice for our study, relative to labeled regions and their derived metrics. First, volume and surface overlap metrics, as used in other literature, are well-known to be biased by the total volume/surface and also ignore misregistrations within the labeled regions themselves because no landmarks exist inside those areas to assess correspondence. Also, using volume size as a metric is reasonable in the context of nonlinear registration algorithms that can locally deform the image volume and change the size of labeled anatomic regions, but our affine registrations create minimal regional size changes, making volume size an inappropriate metric and further motivating the use of pinpoint, single-voxel landmarks in a manner that extends prior work. Additionally, surface distance, which describes registration success by using the average distance between points on one surface to the closest points on another surface, disregards whether homologous points are being compared and highlights a possible lack of spatial precision.

This article targets the radiologic clinic, where patient care involves the affine registration of a single patient’s images across modalities and/or techniques to assess pathology. Because of the guaranteed anatomic correspondence across these images, single-voxel landmarks can be successfully used to assess registration performance. Additionally, spatially precise landmarks are preferred over a labeled region overlap measure in this context because the use of a single person’s images guarantees that a homologous point can be found. Moreover, if widely distributed throughout the brain volume, these single-voxel landmarks can give a sense of the regional registration error not obtainable by using labeled regions. However, if the registration goal is to spatially align the images from different patients (as is frequently the case in the grouping of functional MR imaging data), it makes sense to use volume overlap as the metric of success, particularly given anatomic variations among patients.

Multiple limitations existed for this study. The placement of fiducial landmark sites was subject to user error, voxel size limitations, and image deformations caused by patient motion. A neuroanatomic expert therefore evaluated all landmark sites, followed by adjustments, if needed. Fiducial sites were also limited to those locations that could be reliably identified across images with widely different contrast mechanisms, with most sites ending up in the central region of the brain. To help compensate for this limitation, we sought a well-rounded distribution of landmarks in the anatomy that was usable and also placed landmarks at extreme anterior, posterior, superior, and inferior locations. Measurement precision was limited by using voxels (discretized representations of anatomy) to measure spatial error on a continuous scale. All voxel sizes and error values as originally calculated are therefore reported, to allow readers to form their own judgments. Our patient number was low and focused on a treatment-naïve patient population with gliomas. Further studies are needed to investigate biases in the registration accuracy that may be influenced by the major disease phenotypes seen in a general patient population, including brain metastases, stroke, neurodegeneration, post-surgery, and postradiation.

CONCLUSIONS

In summary, we developed and evaluated a methodology to quantify the registration accuracy of registration algorithms. The method could be used to test any algorithm, providing easy-to-interpret figures of merit that allowed meaningful comparisons with other algorithms in clinical practice. We advocate the publication of figures of merit such as these for all clinical registration algorithms, to better inform the choices of clinical users and allow the future development of improved algorithms for clinical use.


REFERENCES


Quantitative Assessment of Variation in CT Parameters on Texture Features: Pilot Study Using a Nonanatomic Phantom

K. Buch, B. Li, M.M. Qureshi, H. Kuno, S.W. Anderson, and O. Sakai

ABSTRACT

SUMMARY: Our aim was to evaluate changes in texture features based on variations in CT parameters on a phantom. Scans were performed with varying milliampere, kilovolt, section thickness, pitch, and acquisition mode. Forty-two texture features were extracted by using an in-house-developed Matlab program. Two-tailed t tests and false-detection analyses were performed with significant differences in texture features based on detector array configurations (Q values = 0.001–0.006), section thickness (Q values = 0.0002–0.001), and acquisition mode (Q values = 0.003–0.006). Variations in milliampere and kilovolt had no significant effect.

ABBREVIATIONS:

GLCM = gray-level co-occurrence matrix; GLGM = gray-level gradient matrix; GLRL = gray-level run length; MDCT = multidetector row CT; RP = run percentage

Image texture describes a complex visual pattern within an image that consists of simpler subpatterns with characteristic features that may be evaluated through quantitative analysis known as a texture analysis. Texture analysis is a set of quantitative, postprocessing, image-analysis algorithms that are being increasingly used within the field of radiology. The texture analysis may be applied to any imaging technique including CT, MR imaging, and sonography, among others. The texture analysis is composed of a series of mathematical algorithms that extract texture descriptors from an image, thus allowing the mathematic detection of subtle changes in pixel intensity throughout an image.

Within radiology, texture analysis has been used to detect subtle pathologic changes in an image that are not easily quantifiable by the human eye in a variety of areas of the body such as the liver, brain, and cartilage. One of the most prevalent areas for use of texture analysis within radiology is for tumor evaluation and characterization. Despite a growing number of publications on the use of texture analysis for tumor imaging, direct correlations between texture analysis features and histopathologic analysis are still being investigated.

Recently, the upsurge in genomic discoveries and advanced imaging technologies on publicly available data bases, such as The Cancer Imaging Archive (http://www.cancerimagingarchive.net/) and The Cancer Genomic Atlas (https://cancergenome.nih.gov/), large, multi-institutional, quantitative genomic and imaging-based studies, are becoming areas of increased interest. Particularly, within the field of tumor imaging, texture analysis is becoming increasingly used; however, the lack of standardized scanning protocols poses a significant limitation to the use of texture analysis in these instances.

The purpose of this study was to evaluate how changes in CT parameters (milliampere, kilovolt[peak], section thickness, pitch, and acquisition mode) could affect variations in the texture analysis features irrespective of the internal architecture of the item being scanned.

MATERIALS AND METHODS

This study used a phantom for all image acquisitions, precluding the requirement for institutional review board approval.

Phantom Development

A nonanatomic phantom was constructed to investigate whether texture analysis features change with various CT parameters (milliampere, kilovolt[peak], section thickness, pitch, and acquisition mode). The phantom was constructed from a composition of cereal (Cheerios; General Mills, Minneapolis, Minnesota) and commercially available mayonnaise (Hellmann’s; Unilever US, Englewood Cliffs, New Jersey). The content of the phantom was chosen on the basis of a regularly repeating geometric pattern composed of different internal densities, including air, fat, and grain (of
slightly higher density). The phantom measured 15.88 × 23.50 × 31.75 cm (Fig 1).

**CT Protocol**

Noncontrast CT scans were obtained on the phantom on a 64-detector row CT scanner (Lightspeed VCT; GE Healthcare, Milwaukee, Wisconsin) and a 16-detector row CT scanner (LightSpeed, 16-section; GE Healthcare). In addition to scanning the phantom on different detector-row CT scanners, we also scanned the phantom with different CT parameters, including variations in the milliampere, kilovolt(peak), section thickness, and acquisition mode.

Each of the aforementioned CT scanning parameters was altered one at a time, while the remaining 5 texture features were held constant as shown in the Table.

The variations for the acquisition mode included an axial scan compared with a helical scan. Variations in milliampere ranged from 80 mA to 140 mA in increments of 20 mA. Similarly, variations in the kV ranged from 80 kV up to 140 kV in increments of 20 kV. Serial scans were obtained with variations in the thickness of the reconstructed axial datasets of 0.625, 1.25, 2.5, and 5 mm. Changes in helical pitch were also investigated by using pitches of 0.51, 0.98, 1.37, and 1.75.

**Image Segmentation and Texture Analysis**

The internal content of the phantom was manually contoured on an equal number of sections by a fourth-year diagnostic radiology resident on each CT scan. Segmentation was performed by using a dedicated workstation (Advantage Workstation; GE Healthcare) with a semiautomated graphical user interface.

Each contour was then imported into in-house-developed Matlab (MathWorks, Natick, Massachusetts) texture analysis software. The texture analysis software was developed by the co-author (B.L.), and the use of this texture analysis program has been previously reported in the literature. In total, 42 texture features, including 13 histogram features, 5 gray-level co-occurrence matrix (GLCM) features, 11 gray-level run-length (GLRL) features, 4 gray-level gradient matrix (GLGM) features, and 9 Law’s features, were computed and averaged over the images per dataset.

The use of this in-house-developed Matlab program and the specific details of the texture analysis features calculated in this program have been previously published. For additional details on the mathematic equations proposed by Haralick et al and the GLRL matrix defined by Tang, please refer to the Online Appendix.

**Statistical Analysis**

The CT scanning parameters were incrementally changed, and the textures features were compared among the scans by using the Student t test for independent samples. To adjust for multiple comparisons, we performed a false discovery rate correction and calculated the false discovery rate–corrected P values (termed Q values) in addition to raw P values with the Benjamini-Hochberg method described in the literature. Statistical computations were performed by using SAS 9.1.3 software (SAS Institute, Cary, North Carolina). The PROC MULTTEST function in SAS was used to calculate the Q values. A 2-tailed P value of < .05 was used to evaluate statistical significance.

**RESULTS**

The results of the texture analysis features with variations in CT parameters are shown in Online Tables 1–8.

---

**Outline of CT scanning protocol**

<table>
<thead>
<tr>
<th>Varying Parameters</th>
<th>Scan No.</th>
<th>CT Type</th>
<th>mA</th>
<th>kV</th>
<th>Section Thickness</th>
<th>Pitch</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDCT</td>
<td>1</td>
<td>Axial</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Axial</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Axial</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>0.98</td>
</tr>
<tr>
<td>Acquisition Mode</td>
<td>4</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>0.98</td>
</tr>
<tr>
<td>Milliampere</td>
<td>7</td>
<td>Helical</td>
<td>120</td>
<td>120</td>
<td>5</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>0.98</td>
</tr>
<tr>
<td>Kilovolt</td>
<td>11</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>Helical</td>
<td>140</td>
<td>140</td>
<td>5</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>0.98</td>
</tr>
<tr>
<td>Pitch</td>
<td>15</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>1.37</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>1.75</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>0.625</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>1.25</td>
<td>0.98</td>
</tr>
<tr>
<td>Section thickness</td>
<td>19</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>2.5</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>Helical</td>
<td>140</td>
<td>120</td>
<td>5</td>
<td>0.98</td>
</tr>
</tbody>
</table>

*Eight CT scanning parameters were varied during serial CT acquisitions.*
Multidetector Row CT Scanner

Significantly larger values were seen in the histogram features of second deviation ($P = 0.0042$, $Q = 0.018$) and range ($P = 0.0046$, $Q = 0.018$) as shown in On-line Table 1. There were no significant differences in the GLCM texture features with the exception of the texture feature correlation, which demonstrated lower values for the 64-multidetector row CT (MDCT) compared with 16–MDCT ($P = 0.012$, $Q = 0.039$). Overall, higher texture features were seen in all the Law’s features for the 64–MDCT scanner compared with the 16–MDCT scanner (for all Law’s features, $P < 0.0001$, $Q = 0.0005$). No statistically significant differences were seen in the GLRL or GLGM texture features.

Changes in Milliampere

All CT parameters except for milliampere were held constant on a 64–MDCT scanner. The milliampere varied from 80 to 140 mA in increments of 20 mA. For all included texture features, there were no statistically significant differences between the texture features based on the variations in the milliampere, as shown in On-line Table 1.

Changes in Kilovolt

All CT parameters except for kilovolt were held constant for these serial scans of the phantom on the 64–MDCT scanner. Kilovolt varied from 80 to 140 kV in increments of 20 kV. For all included texture features, there were no statistically significant differences between the texture features based on the variations in kilovolt as shown in On-line Table 1.

Changes in Section Thickness

All CT parameters except for the section thickness were held constant for serial scans of the phantom on the 64–MDCT scanner. Serial helical acquisitions of the phantom were performed by using section thicknesses of 0.625, 1.25, 2.5, and 5 mm.

While there was no statistically significant difference in the histogram feature of mean ($P = 0.86$, $Q = 0.86$), statistically significant differences were seen in the other histogram features, including median ($P = 0.0004$, $Q = 0.0005$), SD ($P = 0.0006$, $Q = 0.0008$), second SD ($P < 0.0001$, $Q = 0.0002$), geometric mean ($P = 0.0013$, $Q = 0.0016$), and harmonic mean ($P < 0.0001$, $Q = 0.0002$), as shown in On-line Table 1. Additionally, all the GLCM texture features demonstrated statistically significant differences with variations in the section thicknesses.

Changes in Pitch

For variations in the pitch of helical scans, there were several significant differences in the Law’s features of L5, L6, and L7 ($P = 0.021$, $P = 0.001$, $P = 0.014$, respectively); however, after false discovery rate correction only the L6 feature remained statistically significant ($Q = 0.034$). No statistically significant differences were seen in the histogram, GLCM, GLRL, or GLGM texture features as shown in On-line Table 1.

Changes in Acquisition

Statistically significant differences were seen in the histogram features of second SD ($P = 0.0015$, $Q = 0.0045$) and range ($P = 0.0016$, $Q = 0.0045$), which exhibited lower values in an axial acquisition compared with a helical acquisition. No significant differences were seen in the GLCM features in an axial-versus-helical scan. Lower GLRL values were seen in the axial acquisition compared with the helical acquisition, with significant differences in the GLRL features of short-run emphasis ($P = 0.0024$, $Q = 0.63$), long-run emphasis ($P = 0.006$, $Q = 0.0028$), gray-level nonuniformity ($P = 0.0011$, $Q = 0.0042$), and run-length nonuniformity ($P = 0.0007$, $Q = 0.0029$). No statistically significant differences were seen in the Law features or GLGM texture features shown in On-line Table 1.

Discussion

The results of this study demonstrate statistically significant changes in the texture features based on the use of different CT parameters. CT texture features were not dependent on variations in milliampere and kilovolt. Variations in section thickness resulted in significant differences in the largest number of texture features, most of which were histogram texture features. To a lesser extent, the histogram features of second SD and range were both significantly affected by changes in the MDCT and an axial-versus-helical scan. Differences in CT texture features based on variations in scanning protocols have been suggested by the work of Fave et al; however, to date, there remains no standardization for CT texture analysis or a comprehensive study investigating how CT parameters may influence the various texture parameters.

The use of texture analysis has been increasing in prevalence throughout the radiology literature, particularly as an adjunct aiding in diagnosis, lesion characterization, and even in the evaluation for treatment-related response. Preliminary works investigating the use of a texture analysis to detect and characterize stages of hepatic fibrosis have been reported in the literature. Similarly, texture analysis has also been used for examining potential differences in tissue architecture on CT in oropharyngeal squamous cell carcinomas and for help in evaluating changes in cartilage. In these studies, a single CT protocol scanning algorithm has been used, limiting potential variations in the texture analysis feature related to scanning technique. This has future implications in multi-institutional research, such as for the imaging-based studies from The Cancer Imaging Archive, in which different scanning techniques could potentially reflect quantitative differences related to scanning techniques across different equipment manufacturers.

The GLCM texture features were most affected by changes in slice thickness. This dependency may be related to an increase in the likelihood of partial volume effect as the section thickness increases. Partial volume effect occurs when a structure partially intrudes the x-ray beam. With thicker sections, due to the volume-averaging effect, the contrast of the anatomy with its background decreases (ie, decreased GLCM contrast) and, visually, the image becomes more “homogeneous” (ie, increased GLCM homogeneity). Clinically, this is an important consideration because several GLCM features, particularly the GLCM texture feature of entropy, have been described as being important for tumor imaging.

Multiple prior studies have highlighted the potential importance and ease of use of a quantitative texture analysis to evaluate
subtle changes in pixel intensity, which may not be evident to the human eye. The potential clinical applications of CT texture analysis include disease and lesion characterization, prognosis and treatment prediction, and treatment-response evaluation. However, a precise relationship between texture analysis features and a histopathologic understanding of correlate changes in tissue microstructure is less well-defined. Additionally, a prior study noted changes in texture features based on changes in CT parameters thought to reflect a degree of image heterogeneity. In this study, we developed a physically heterogeneous phantom simulating varying densities to closely investigate how changes in CT scanning parameters influence texture analysis features. The impact of CT acquisition parameters on texture analysis features is an important consideration for the development and understanding of texture-based features when applied to CT images of clinical patients or research subjects. This is particularly important in cases in which texture analysis is used to evaluate treatment-related responses, longitudinal studies, and cross-institutional studies whereby the CT scanning parameters may vary, therefore further complicating the texture analysis results because measurements of image heterogeneity may be confused with biologic responses.

A previous study used a water phantom with no internal structural heterogeneity and concluded that the CT texture features were relatively insensitive to CT parameters such as tube voltage, tube current, and section thickness. The results of this study also demonstrated that all examined texture analysis features were insensitive to tube voltage and tube current, but not scanning parameters or section thickness. The observed insensitivity of the texture features to changes in the tube voltage and current may reflect the phantom used in our study comprising similar material densities. Future investigation will be directed at using a more complex phantom with a greater variation in material densities. This study serves as an initial pilot investigation using a nonanatomic phantom looking into how differences in texture analysis features change with variations in CT parameters. The construction of additional phantoms with greater architectural complexity and composed of higher attenuation materials such as iodine, bone, and so forth is the next step in furthering our understanding of how material composition and CT techniques contribute to variations in texture analysis.

There are several limitations to this study. First, it used a nonanatomic phantom constructed of varying internal architecture of relatively low-density material. The texture results of this study would be most applicable to low-density soft-tissue and fat-attenuating structures. Higher attenuation material such as bone was not reflected in the design of this phantom. Future investigations will pursue an anatomic internal architecture, following this initial, pilot study. Second, a discrete range of varying milliampere, kilovolt, section thickness, and pitch was interrogated in this study; however, the ranges investigated may not be broad enough to cover a comprehensive range of all milliampere, kilovolt, section thickness, and pitch values used in radiology practices. Additionally, we chose to manually contour the internal content of our phantom for every study; that introduces a source of standardization. Future work on this subject matter will entail an investigation into automated contouring to reduce any potential variation from manual contouring.

CONCLUSIONS

While texture analysis represents an increasingly popular, post-processing, quantitative evaluation technique that can potentially be used as an adjunct in diagnostic imaging as a possible biomarker, standardization of CT parameters for the use of texture analysis is crucial to prevent features of intrinsic image heterogeneity from being confused with biologic features.

REFERENCES

ing features and texture analysis. AJR Am J Roentgenol 2015;204:1013–23 CrossRef Medline
17. Ganeshan B, Miles KA. Quantifying tumour heterogeneity with CT. Cancer Imaging 2013;13:140–49 CrossRef Medline
ABSTRACT

BACKGROUND AND PURPOSE: Intracranial atherosclerosis is a common cause of ischemic stroke. Intracranial stenosis is most commonly quantified by the Warfarin-Aspirin Symptomatic Intracranial Disease method, which involves calculating a ratio of luminal diameter measurements on conventional angiography. Our purpose was to determine whether a single linear measurement of the narrowest caliber of the intracranial ICA on MRA can accurately predict Warfarin-Aspirin Symptomatic Intracranial Disease stenosis measurements.

MATERIALS AND METHODS: We identified patients from a prospective stroke registry who had undergone head MRAs to quantitatively evaluate the degree of Warfarin-Aspirin Symptomatic Intracranial Disease–derived stenosis in each intracranial ICA. We also made a single linear millimeter measurement at the site of maximal narrowing of the ICA. We calculated a correlation coefficient between the lumen diameter in millimeters and percentage Warfarin-Aspirin Symptomatic Intracranial Disease stenosis. We performed receiver operating characteristic analysis to determine optimal luminal diameter cutoff values.

RESULTS: In 386 unique intracranial ICAs, we found a strong linear relationship between single lumen measurements and Warfarin-Aspirin Symptomatic Intracranial Disease–style stenosis measurements ($R = -0.84, P < .0001$). We found that ICA lumen diameters of ≤2.1 and ≤1.3 mm were optimal cutoffs for identifying patients with ≥50% stenosis and ≥70% stenosis, respectively (area under the curve = 0.96 and 0.99, respectively).

CONCLUSIONS: There is a strong linear relationship between the narrowest lumen diameter of the intracranial ICA and percentage stenosis. Our results suggest that a single lumen diameter measurement on MRA allows accurate estimation of Warfarin-Aspirin Symptomatic Intracranial Disease stenosis, which may affect risk stratification and treatment decisions.

ABBREVIATIONS: CAESAR = Cornell Acute Stroke Academic Registry; WASID = Warfarin-Aspirin Symptomatic Intracranial Disease

Intracranial atherosclerotic disease is a major risk factor for ischemic stroke. As one of the most common causes of stroke worldwide, intracranial atherosclerosis is generally assessed by measuring the degree of luminal stenosis on angiography. Accurate assessment of the degree of luminal stenosis of the cerebral vasculature is important to assign stroke etiology because most stroke classification schemes require a luminal stenosis of ≥50% for a stroke to be attributed to large-vessel atherosclerosis. In addition, given trial data showing that aggressive medical therapy is an effective approach to reduce recurrent stroke risk in patients with ≥70% stenosis of a major intracranial artery, reliable and accurate stenosis measurements are critical for appropriate patient selection for preventative therapy. Similarly, studies have shown that patients with ≥70% stenosis are at an elevated risk of stroke recurrence compared with those with 50%–69% stenosis.

The most commonly used method for calculating intracranial stenosis severity is the Warfarin-Aspirin Symptomatic Intracranial Disease (WASID) method, which has been the basis of several randomized controlled trials for the treatment of major intracranial stenosis. The measurement is performed by calculating a ratio of luminal diameters obtained from angiographic images. Although the WASID technique has high reproducibility with experienced readers and meticulous technique, the practical imple-
For an accurate measurement of the degree of Warfarin-Aspirin Symptomatic Intracranial Disease stenosis, we made a linear measurement of the most stenotic portion of the ICA on either the MIP or axial source image, in this case 5.3 mm (D). If the petrous ICA was also diseased, the most distal, parallel part of the extracranial ICA was measured. The ratio of these 2 measurements was then used to calculate the WASID stenosis, in this case $\frac{1 - 0.9/5.3}{100} = 83\%$ stenosis. Alternatively, we could use the single linear measurement at the site of greatest stenosis, 0.9 mm (C), which correlates with a stenosis of $>70\%$.

![MRA MIP (A) and axial source (B) images demonstrating high-grade stenosis of the right intracranial ICA.](image)

**FIG 1.** MRA MIP (A) and axial source (B) images demonstrating high-grade stenosis of the right intracranial ICA. For an accurate measurement of the degree of Warfarin-Aspirin Symptomatic Intracranial Disease stenosis, we made a linear measurement of the most stenotic portion of the ICA on either the MIP or axial source image, in this case 0.9 mm (C). A second “normal” measurement was made at the widest, nontortuous, normal portion of the petrous ICA that had margins parallel with the site of stenosis, in this case 5.3 mm (D). If the petrous ICA was also diseased, the most distal, parallel part of the extracranial ICA was measured. The ratio of these 2 measurements was then used to calculate the WASID stenosis, in this case $\frac{1 - 0.9/5.3}{100} = 83\%$ stenosis. Alternatively, we could use the single linear measurement at the site of greatest stenosis, 0.9 mm (C), which correlates with a stenosis of $>70\%$.

mentation of WASID measurements can be cumbersome and prone to error, especially in the ICA where multiple vascular segments and an anatomically tortuous course can complicate measurement. For these reasons, strictly defined WASID stenosis measurements are not routinely performed in day-to-day practice, especially with noninvasive vascular imaging modalities. Thus, having a rapid and accurate method of estimating luminal stenosis measurements by using only single-diameter measurements obtained on cross-sectional images would be beneficial, similar to prior work done in the extracranial carotid artery.7 The purpose of our study was therefore to test the hypothesis that a single luminal measurement on MRA images of the ICA can accurately approximate the WASID-determined intracranial ICA stenosis in a cohort of patients with acute ischemic stroke.

**MATERIALS AND METHODS**

Patients admitted to New York Presbyterian Hospital/Weill Cornell Medical Center with acute ischemic stroke in 2013 were included in the prospectively maintained Cornell Acute Stroke Academic Registry (CAESAR). We retrospectively included all patients with MRA of the head because we were interested in characterizing the degree of intracranial arterial stenosis. We excluded any patients with complete occlusion of their intracranial ICAs and any patients with images that were too motion-degraded for accurate interpretation and measurement. The Weill Cornell Medicine institutional review board approved this data collection and waived the need for informed consent, given the retrospective nature of the analysis and minimal risk to subjects.

As part of the CAESAR registry, data on patient demographics, National Institutes of Health Stroke Scale score on admission, and vascular risk factors (including atrial fibrillation, tobacco use, diabetes mellitus, hypertension, dyslipidemia, peripheral vascular disease, and cardiac valvular disease) were prospectively collected by trained hospital personnel.

Retrospectively, 2 neurologists used available medical records to independently identify stroke etiology by using the Trial of Org 10172 in Acute Stroke Treatment classification scheme4 with a third neurologist to independently resolve disagreements.

All included patients underwent noncontrast MRA examinations on either a 1.5T or 3T Signa (GE Healthcare, Milwaukee, Wisconsin) scanner. 3D-TOF acquisitions of the head were performed with an FOV of 20 cm, 1.4-mm section thickness, TR = 25, TE = 3, and matrices of 320 × 192 and 3210 × 224 on 1.5T and 3T scanners, respectively. Maximum intensity projections of each intracranial ICA were created.

We measured the degree of ICA narrowing by using the WASID method6 and single luminal measurements.

**WASID Measurement Technique**

To calculate WASID stenosis, we obtained 2 measurements for each intracranial ICA: 1) a linear measurement at the site of the most severe stenosis on either the MIP or axial source images; and 2) a linear measurement at the widest, nontortuous, normal portion of the petrous ICA parallel to the site of stenosis. Using these measurements, we calculated the degree of WASID stenosis by using the following equation: Percentage Stenosis = $\frac{1 - [D_{\text{stenosis}}/D_{\text{normal}}]}{100}$, where $D_{\text{stenosis}}$ is the diameter of the artery at the site of most severe degree of stenosis and $D_{\text{normal}}$ is the diameter of the proximal artery at its widest, nontortuous, normal segment.6

**Single Luminal Measurement Technique**

For the single luminal measurement calculations, a single luminal measurement was made at the site of the most severe stenosis within the intracranial ICA.

The stenosis measurements for both techniques were obtained in either the cavernous or supraclinoid segments of the ICA. In cases of no measurable stenosis, linear measurements were made in the normal-caliber cavernous segment of the intracranial ICA. Stenosis was evaluated and measurements were made by using a combination of the MIP and axial source MRA images. Linear measurements were obtained on the MIP projection or axial source image showing the greatest degree of stenosis (Fig 1 and On-line Fig 1). Measurements were made from outer lumen to outer lumen in all arteries. A second radiologist independently calculated WASID stenosis measurements of the first 50 consecutive intracranial ICAs ordered by admission date to evaluate interobserver reproducibility.
**Statistical Analysis**

Categoric data are presented as number (percentage); and continuous data, as mean ± SD. Pearson correlation coefficients were used to determine the strength of the linear relationship between millimeter stenosis and percentage WASID stenosis followed by linear regression to determine 95% confidence intervals for the percentage stenosis predicted values at each millimeter stenosis.

**RESULTS**

Of the 335 patients included in the 2013 CAESAR stroke registry, 134 patients were excluded because they did not have an MRA examination of their head; 7 patients, because they had complete occlusion of their intracranial ICA; and 1 patient, because his MRA was too motion-degraded for accurate measurement. Of the 193 patients (386 arteries) who were ultimately included, the mean age was 71.9 years (Table 1).

![Image](https://example.com/fig2.png)

**FIG 2.** Correlation scatterplot demonstrating a linear relationship between millimeter stenosis measurements and percentage WASID stenosis for all included patients with $R = -0.84$ ($P < .0001$).

**DISCUSSION**

In our analysis, we found a strong linear relationship with excellent correlation between millimeter measurement of the narrowest point of the ICA and percentage WASID stenosis. Additionally, we were able to determine optimal millime-
raphy in screening for atherosclerosis. Conversely, MRA is an attractive technique for screening intracranial stenosis for several reasons: First, MR angiographic studies are obtained relatively quickly and are often performed concurrently with brain MR imaging in the evaluation of patients with potential stroke. At many institutions, including our own, cross-sectional imaging is almost exclusively used for the evaluation of intracranial stenosis and to inform treatment decisions, including whether to initiate more intensive medical therapy for stroke prevention. Second, using MIPs from MRA studies simulates the appearance of conventional angiography, thereby allowing measurements to be obtained in the same locations and projections, similar to those for the original WASID measurements. Third, recent studies demonstrate that MRA performs well, with sensitivities and specificities ranging from 80% to 100% and 89% to 95%, respectively, compared with conventional angiography when assessing intracranial stenosis. Furthermore, MRA is not subject to the challenge of separating a high-density contrast-enhanced lumen from adjacent calcification and bone inherent in the evaluation of CTA of the head.

Our study had some limitations. First, time-of-flight MR angiography is sensitive to flow disturbances caused by stenosis and is known to overestimate the degree of stenosis compared with CT and conventional angiography. By evaluating both the MIP and axial source images to measure the degree of stenosis, according to established methods shown to be highly accurate, we attempted to minimize the degree of overestimation of stenosis. Second, obtaining measurements on MR angiograms can be subject to error, given the lack of precise spatial resolution. We used a uniform method for making all WASID measurements, including measuring from outer lumen to outer lumen. Additionally, we had excellent interobserver reliability between our two readers. Last, we had a relatively small number of ICA demonstrating high-grade intracranial stenosis, with an overall prevalence of 12.9% for a stenosis of 50% and 1.6% for a stenosis of 70%. The relatively low prevalence of high-grade stenosis likely contributed to the positive predictive value of 50% for the cutoff for >70% stenosis.

We believe that by focusing on all ICAs in our cohort regardless of stenosis severity, our results are more generalizable, especially given that our data are consistent with the prevalence data of high-grade intracranial ICA stenosis derived from population-based studies. Because we did not limit our study sample exclusively to patients with high-grade stenosis, the precision of our cutoff values to estimate exact degrees of high-grade WASID stenosis values is somewhat limited. We believe that our cutoff values are most useful as a practical and rapid approach to screen for potentially clinically relevant intracranial atherosclerotic stenosis, whose identification may warrant more detailed evaluation and possible treatment. Validating these cutoff values further in a separate prospective cohort would also be valuable.

CONCLUSIONS

We found a strong linear relationship between a simple measurement of the narrowest point of the intracranial ICA and the percentage WASID stenosis. Our results allow a direct estimation of WASID stenosis via a single-diameter measurement on MRA. We found an optimal cutoff measurement of 2.1 mm for identifying patients with 50% stenosis and 1.3 mm for identifying patients with 70% stenosis and that there is a strong linear relationship between the narrowest lumen diameter of the intracranial ICA and WASID-derived percentage stenosis, allowing a single linear measurement to provide accurate WASID stenosis estimates, which are useful in guiding treatment decisions.

Table 2: WASID percentage stenosis estimates from millimeter measurements of the most stenotic portion of the intracranial ICA

<table>
<thead>
<tr>
<th>Stenosis (mm)</th>
<th>% WASID Stenosis (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.0–3.8</td>
<td>14–30</td>
</tr>
<tr>
<td>2.2–2.9</td>
<td>31–48</td>
</tr>
<tr>
<td>1.3–2.2</td>
<td>50–65</td>
</tr>
<tr>
<td>&lt;1.3</td>
<td>≥65</td>
</tr>
</tbody>
</table>

Table 3: Breakdown of the millimeter measurements of the 139 arteries with detectable WASID stenosis

<table>
<thead>
<tr>
<th>Stenosis Measurement</th>
<th>&lt;50%</th>
<th>50%–69%</th>
<th>70%+</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;2.1 mm</td>
<td>75</td>
<td>6</td>
<td>0</td>
<td>81</td>
</tr>
<tr>
<td>1.4–2.1 mm</td>
<td>14</td>
<td>32</td>
<td>0</td>
<td>46</td>
</tr>
<tr>
<td>≤1.3 mm</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>89</td>
<td>44</td>
<td>6</td>
<td>139</td>
</tr>
</tbody>
</table>

FIG 3. A, Receiver operating characteristic curve for predicting >70% stenosis demonstrates an optimal cutoff of 1.3 mm. B, Receiver operating characteristic curve for predicting >50% stenosis demonstrates that a measurement of 2.1 mm is an ideal cutoff.

REFERENCES

ABSTRACT

BACKGROUND AND PURPOSE: Periprocedural thrombus fragmentation is a relevant risk in endovascular stroke treatment. Because factors influencing its occurrence are largely unknown, this study addresses a potential relationship between thrombus histology and clot stability.

MATERIALS AND METHODS: Eighty-five patients with anterior circulation stroke treated with thrombectomy were included in this retrospective study. The number and location of emboli after retrieving the primary thrombus, the number of maneuvers, and TICI scores were evaluated. H&E and neutrophil elastase staining of retrieved clots was performed, and semiquantitative measurements of thrombus components were correlated with procedural parameters.

RESULTS: An inverse correlation between maneuvers required for thrombus retrieval and the number of distal and intermediate emboli was observed (Spearman $r = -0.23; P = .032$). Younger patients were at higher risk for periprocedural thrombus fragmentation (Spearman $r = 0.23; P = .032$). Bridging thrombolysis tended to be associated with fewer maneuvers (2 vs 3, $P = .054$) but more emboli (1 vs 0, $P = .067$). While no consistent correlation between procedural parameters and red/white blood cells and fibrin-/platelet fractions could be found, higher amounts of neutrophil elastase–positive cells within the thrombus were independently associated with the occurrence of multiple emboli (adjusted OR, 4.6; 95% CI, 1.1–19.7; $P = .041$) and lower rates of complete recanalization (adjusted OR, 0.3; 95% CI, 0.1–0.9; $P = .050$).

CONCLUSIONS: Younger age, easy-to-retrieve thrombi, and bridging thrombolysis may be risk factors for periprocedural thrombus fragmentation. Findings from standard histologic stains did not provide insight into thrombectomy-relevant thrombus stability. However, higher neutrophil levels in the thrombus tissue were related to an increased risk of periprocedural thrombus fragmentation. This observation aligns with the proposed thrombolytic capacity of neutrophil elastase and points to its potential clinical relevance in the context of stroke thrombectomy.

ABBREVIATIONS: ACA = anterior cerebral artery; F/P = fibrin-/platelet accumulations; IQR = interquartile range; MT = mechanical thrombectomy, NE = neutrophil elastase; POS = primary occlusion site; PTF = periprocedural thrombus fragmentation; RBC = red blood cells; WBC = white blood cells

Mechanical thrombectomy (MT) of large-vessel occlusion has evolved as a safe and effective procedure that plays an indispensable role in modern therapeutic management of acute ischemic stroke.1-5 In recent randomized trials, high rates of successful recanalization (range, 59%–88%) were considered a key element in achieving excellent rates of good functional outcome (range, 33%–71%).6 However, not all successfully treated patients showed complete (TICI 3) recanalization; this outcome potentially limits therapeutic benefit.

In general, all endovascular MT techniques are accompanied by the risk of periprocedural thrombus fragmentation (PTF) and subsequent downstream embolism,7-9 preventing complete recanalization. Because the neurologic outcome of patients with complete (TICI 3) recanalization is significantly better compared with patients with “almost complete” (TICI 2b)10 or incomplete recanalization (TICI 1–2a),11 understanding the factors contributing to PTF may prove beneficial in achieving maximal therapeutic benefit. Previous reports demonstrated that thrombus stability12 may influence the incidence of PTF, and analyses of cellular thrombus composition have revealed a possible associa-
tion between thrombus histology and thrombus etiology as well as clinical outcome. Discrepant results in previous studies might be primarily explained by low patient numbers and the risk of clot fragmentation. The latter may bias the representative character of the analyzed fragment.

While the main cellular components of a thrombus are known to be fibrin-/platelet accumulations (F/P) as well as red blood cells (RBC), a higher fraction of RBC has been associated with increased rates of successful endovascular recanalization as noninvasively measured by whole-thrombus density (CT) and corresponding blooming artifacts (MR imaging). Besides common thrombus characteristics obtained from H&E staining, new evidence has emerged that the degree of inflammatory cell invasion, particularly by neutrophils, may alter the stability and degradation of a thrombus. This finding is of particular interest because neutrophils exhibit fibrinolytic activity, which may weaken clot stability. Nevertheless, the impact of inflammatory cells on the mechanical properties of a thrombus remains uncertain, especially in the context of stroke thrombectomy. Potential knowledge of the clot composition before MT may be a further valuable tool to aid in the selection of the most appropriate devices and techniques to avoid PTF.

To this end, this is the first study investigating the dependency of procedural thrombectomy characteristics on anatomic and immune-histochemical thrombus histology, to our knowledge.

MATERIALS AND METHODS

Subjects and Outcome

All consecutive patients presenting with a stroke due to large-vessel occlusion in the anterior circulation between July 2010 and September 2012 at a tertiary care center were included in this retrospective single-center study. Parts of this cohort and histologic analyses have been previously published. All patients underwent MT at the Department of Neuroradiology, and thrombus material was preserved. All patients in whom no thrombus material could be retrieved were excluded. Under this premise, 85 patients (45 women; median age, 73 years) formed the final study cohort. This study was approved by the local ethics committee at the Klinikum rechts der Isar of the Technical University of Munich, Germany, in accordance with the ethical standards of the 1964 Declaration of Helsinki and its later amendments. Clinical outcomes were measured as NIHSS at the day of discharge. Substantial neurologic improvement was defined as either NIHSS at day of discharge ≤1 or the difference between NIHSS on admission and NIHSS at the day discharge ≥8. This definition was based on previous studies that have shown this criterion to be a sensitive outcome measure.

Image Analysis

The primary occlusion site (POS) was defined as the location and extent of the initial thrombus as evaluated on a synopsis of admission CT/CTA and initial DSA. After recanalization of the POS, any vessel occlusions distal from the POS were considered emboli due to PTF. This simplified definition is based on previous findings, which have shown that >90% of intracranial vessel occlusions are caused by a single thrombus. Hence, most vessel occlusions after POS recanalization are due to PTF. Emboli following PTF were further categorized into proximal, intermediate, and distal, according to their localization in lateral and corresponding anteroposterior DSA projections (On-line Fig 1). Emboli involving the M1 or M2 segment were always defined as proximal emboli, whereas those located distal to the pericallosal artery (on lateral projections) were considered distal emboli (see examples of emboli in Figs 1 and 2). The remaining vessel occlusions were assigned to the group of intermediate emboli. A schematic classification of emboli locations in cases of initial MCA or carotid-T occlusions is presented in On-line Fig 1. The success of POS recanalization (TICI-POS) as well as after additional rescue maneuvers (TICI-FINAL) was rated according to the original TICI scale, with TICI 2b defined as reperfusion of more than two-thirds of the initial occluded territory. Note that TICI-POS and TICI-FINAL may be the same (ie, if no additional maneuvers were performed). All images were evaluated in consensus by 2 experienced neuroradiologists.

Endovascular Procedure

Patients underwent angiography if groin puncture could be performed within 6 hours after symptom onset, clinical presentation was severe (NIHSS ≥4), and no early infarct signs involving more than one-third of the MCA territory were present on cranial CT on admission. IV rtPA was administered as “bridging therapy” in the absence of contraindication (n = 58, 68.2%). All stent-retriever passages were counted, and the number of maneuvers was registered after POS recanalization and at the end of the procedure. All procedures were performed by using a distal access catheter (MCA/anterior cerebral artery [ACA]) or proximal flow arrest by balloon occlusion (carotid-T) and one of the following stent retrievers (maneuvers with the respective device): Solitaire (n = 45; Covidien, Irvine, California), pREset 4–20 (n = 58; Phenox, Bochum, Germany), Trevo (n = 29; Stryker, Kalamazoo, Michigan), Revive (n = 2; Codman Neurovascular, Raynham, Massachusetts), Pulse (n = 4; Penumbra, Alameda, California), and Separator 3D (n = 6, Penumbra). Aspiration techniques, applied as stand-alone approach, were not performed in this study.

Thrombus Histology

After clot retrieval, thrombus material was fixed in phosphate-buffered 4% formalin. H&E staining and subsequent quantitative analysis of WBC, RBC, and F/P content were performed as described previously. To assess the prevalence of neutrophils within the thrombus, we stained the samples immunohistochemically by using an anti-neutrophil elastase monoclonal mouse antibody (clone NP57, Dako Denmark, Glostrup, Denmark). The number of neutrophil elastase (NE)-positive cells was semiquantitatively evaluated by 2 independent raters (NE index) who were blinded to the clinical data. Discrepancies were rated in consensus in a separate session. Evaluation was performed by using a 5-step grading scale: 0 (none), 1 (scattered), 2 (intermediate), 3 (clustered), 4 (high, >50% of all cells) (Fig 3).

Statistical Analysis and Illustrations

The Shapiro-Wilk test was applied to analyze data for normal distribution. Because all variables, except NIHSS on admission (P = .322 in Shapiro-Wilk test), were non-normally distributed,
bivariate correlation analysis with the 2-sided Spearman correlation was performed. Frequency counts and median/mean comparison were evaluated by using standard statistical measures (Fisher exact test, Mann-Whitney U test). For median values, the interquartile range (IQR) is shown; for mean values, the SD is shown. Because PTF was shown to be a multifactorial process (eg, age, MCA/ACA versus ICA occlusion, bridging therapy; see “Results”), we adjusted the analysis of histologic clot characteristics for these potential confounders by using a multivariate logistic regression. For statistical analysis, SPSS statistics, release 23.0 (IBM, Armonk, New York), was used. Illustrations were prepared by using Adobe Photoshop CS4 (Adobe Systems, Mountain View, California).

RESULTS
Study Population and Outcomes
Inclusion criteria were met by 85 patients (mean age, 70.2 ± 14.6 years; 45 women) (Table). Approximately one-third (n = 26, 30.6%) of patients presented with a carotid-T occlusion, while the rest had isolated occlusions of the MCA/ACA (n = 57, 67.1%, and n = 2, 2.3%, respectively). The median NIHSS score at presentation was 15 (IQR, 10.5–18) and improved to 5.5 (IQR, 2–14) by the day of discharge. The median symptom-onset-to-treatment-time of 225 minutes (IQR, 165–278.75 minutes) could be sufficiently determined for 72 patients. Following recanalization of the POS, 36 patients (42.2%) showed no peripheral emboli, corresponding to an instant TICI 3 recanalization. Rescue therapy, namely retrieving downstream thrombi due to PTF, was successful in 27 cases. A median of 1 (IQR, 1–3.5) additional stent-retriever maneuver was required to perform the rescue therapy. The final rate of successful recanalization was 91.7%.

When a single device type was used, distribution of final reperfusion success and the number of emboli did not differ among different stent-retriever types (P = .442 for final reperfusion success and P = .931 for the number of all emboli; On-line Fig 2). Use
of multiple device types during a solitary thrombectomy was associated with lower rates of successful recanalization \( (P = .042) \). Occurrence of proximal and intermediate emboli after recanalization of the POS was associated with poorer neurologic outcome (NIHSS-proximal emboli: Spearman \( r \), 0.279; \( P = .013 \); NIHSS-intermediate emboli: Spearman \( r \), 0.223; \( P = .049 \)), whereas patients showing a substantial neurologic improvement had fewer proximal \( (P = .029) \) and intermediate emboli \( (P = .018) \).

**Physical Clot Properties**

A consistent, inverse trend could be observed between the ease of the procedure and the risk of PTF: An easier procedure, namely requiring fewer stent-retriever maneuvers and less time to POS recanalization, was associated with a higher risk of distal and intermediate embolization (On-line Table). We also noted an age dependency of emboli occurrence because younger patients were at higher risk for PTF, resulting in higher numbers of intermediate and overall thrombus fragments distal to the POS. No correlation could be found between the time from symptom onset to groin puncture and the risk of embolization or maneuvers required. Administration of preinterventional IV rtPA tended to be associated with fewer endovascular maneuvers \( (2 \text{ versus } 3, P = .054) \), shorter time to POS recanalization \( (30 \text{ minutes versus } 47 \text{ minutes}, P = .051) \), but a higher sum of overall emboli \( (1 \text{ versus } 0, P = .067) \). In carotid-T occlusions as opposed to isolated MCA/ACA occlusions, more device passages were required to retrieve the primary thrombus \( (3 \text{ versus } 2, P = .006) \), resulting in longer times for POS recanalization \( (66 \text{ versus } 33 \text{ minutes}, P < .001) \). However, no difference could be found regarding the location or number of occurring emboli \( (all \ P > .5) \).

**Thrombus Histology**

The mean fraction of RBC, F/P, and WBC was 42% (IQR, 22%–57%), 49% (IQR, 36%–69%), and 7% (IQR, 5%–11%), respectively. The median NE index, as determined for 42 patients, was 2 (IQR, 1–3). Clot histology characteristics did not differ between
patients receiving bridging therapy and patients who did not (all, \( P > .3 \)). Except for a higher NE index in carotid-T thrombi (median 2; IQR, 2–3.5, versus 2; IQR, 1–2; \( P = .008 \)), no difference in histologic clot characteristics could be found when comparing MCA and carotid-T occlusions (all, \( P > .3 \)). We could not observe a correlation between the time from symptom onset to recanalization or age and fraction of RBC, F/P, and WBC (all, \( P > .5 \)). No consistent association between procedural parameters and RBC, F/P, or WBC content within the clot could be found (all, \( P > .1 \); On-line Table). This was also true when restricting analysis to different stent-retriever types (all, \( P > .2 \)). However, the fraction of RBC (median RBC fraction, 50% [IQR, 28%–59%], versus 39% [IQR, 21%–57%; \( P = .183 \)) and the degree of neutrophil invasion (median NE index, 2 [IQR, 2–3] versus 2 [IQR, 1–3; \( P = .192 \)) tended to be higher in patients with multiple embolizations. The effect of neutrophil invasion on the occurrence of multiple emboli remained statistically tangible (without case restriction) when adjusting for age, site of occlusion (carotid-T versus MCA/ACA), and bridging therapy in a multivariate logistic regression model with multiple (≥1) emboli defined as dependent variables (adjusted OR, 4.6; 95% CI, 1.1–19.7 for every NE index grade increase; \( P = .041 \); pseudo-\( R^2 = 0.568 \)). In contrast, the fraction of RBC was not a significant factor associated with the occurrence of multiple emboli by using the same logistic regression model (OR, 1.0; 95% CI, 0.9–1.0; \( P = .390 \)). Pursuant to the first finding, the median NE index was lower in successfully recanalized patients (2 versus 3; \( P = .033 \), On-line Fig 3). Furthermore, the NE index was an independent predictor of unsuccessful recanalization (TICI-FINAL ≤ 2b, \( P < .050 \); pseudo-\( R^2 = 0.322 \)).

**DISCUSSION**

PTF during MT in acute stroke is common and was observed in more than half of the patients. This study shows 3 major findings: 1) In vivo clots differ regarding their MT-relevant mechanical properties, 2) common histologic clot characteristics do not consistently correlate with the mechanical clot properties, and 3) neutrophil invasion of the clot is linked to PTF and might serve as a novel surrogate for clot stability.

Our results show that easily retractable clots are generally prone to small/intermediate fragment dissociation, while those retracted with considerable effort are usually more stable and less

\[\text{FIG 3. Histologic specimens of 2 different clots with H&E staining (magnification } \times 30 \text{) (A and E), segmentation of RBC (red), F/P (purple), WBC (blue), and staining artifacts (brown) for composite quantification (B and F). H&E staining (magnification } \times 200 \text{) (C and G) and NE staining (magnification } \times 200 \text{) (D and H) with accumulation of NE-positive cells marked (asterisks). Clot composition is the following: upper row: 30% RBC, 51% F/P, 19% WBC; NE index 1, lower row: 7% RBC, 84% F/P, 9% WBC, NE index 4.}\]
susceptible to PTF. During endovascular stroke treatment, the use of stent retrievers and aspiration devices has a risk of thrombus fragmentation, which can lead to an “embolic shower.”28-29 Smaller fragments (<200 μm) constitute, by far, the highest percentage of occurring emboli.28 They usually dissolve due to spontaneous thrombolysis or revascularization as achieved by embolus extravasation.30 However, those emboli are usually undetected by conventional DSA, and their true clinical impact remains unclear.31 Larger emboli (>200 μm) cause clinically relevant cerebral occlusion20 and lead to angiographically defined incomplete recanalization (TICI < 3). We have found those emboli to be clinically important because intermediate and proximal emboli were associated with lower NIHSS scores at discharge. Occurrence and “rescue” removal of those emboli are of important clinical relevance because patients with complete recanalization have fewer neurologic deficits.10

The beneficial effects of preinterventional systemic rtPA administration concur with our assessment; this finding suggests a facilitating effect of bridging therapy.33-35 Patients receiving rtPA needed fewer maneuvers and had shorter procedure times. However, the presumed bridging therapy–related thrombus softening came with the risk of PTF, which has been reported previously.36

Thrombus characteristics have been shown to predict technical outcome in endovascular stroke therapy. Particularly, thrombi with higher Hounsfield units are associated with higher rates of successful recanalization.19,21 Given prior histologic analyses, it seems reasonable to assume that those thrombi are RBC rich and have low F/P content.18 In the present study, thrombi prone to dissociate into multiple emboli tended to have a higher RBC fraction. However, this association was inconsistent and statistically irrelevant when correcting for possible confounders. Our data regarding the true success rate of thrombectomy are particularly limited because we only analyzed cases in which thrombus material could be retrieved (usually TICI > 1–2a). Hence, the true impact of RBC on the capability of retrieving the clot and POS recanalization cannot finally be assessed in the presented cohort.

To the best of our knowledge, this is the first study that establishes a possible link between procedurally relevant thrombus stability and the content of neutrophils. Neutrophils invade newly formed thrombi through recruitment due to adherent activated platelets.27 Adhesion triggers activation of neutrophils and subsequent release of neutrophil elastase.28 Initially, proinflammatory mediators promote fibrin cross-linking and the formation of neutrophil extracellular traps, which promote a procoagulatory state.29 However, activated neutrophils in the later phase also seem to restrict growth and promote thrombus degradation by NE-dependent fibrinolysis (“cell-dependent thrombolysis”).72 Komorowicz et al.73 could indeed show that NE triggers the release of thrombus fragments in the soluble phase when incubated under shear stress conditions. Beyond altering the stability properties of thrombi, experimental data further suggest that NE seems to be involved in inflammatory reperfusion damage,40 and a recent study proposed an association between the presence of CD4+ T-cells and CD68+ monocytes/macrophages within the thrombus and stroke severity.74 However, the true impact of “thromboinflammation” has yet to be evaluated in further studies.75 In this context, the presented data underscore the need for considering and correcting for potential effects of inflammatory cells on thrombus stability, when evaluating its effect on clinical outcome in patients with stroke who were endovascularly treated.

Our study has several limitations beyond the common ones of a retrospective study design. Due to our inclusion criteria, the cohort exhibits a selection bias for successfully recanalized patients because histologic analysis was feasible only in cases in which the thrombus could be retrieved. Thus, results regarding the effects of histologic parameters on the true success rates of MT should be interpreted cautiously. Second, assessing the rate and location of emboli by using midprocedural DSA is a simplification and might not be as sensitive as susceptibility-weighted imaging for thrombus detection.63 However, it is the only feasible method of PTF detection during the procedure and has the advantage of assessing the hemodynamic relevance of emboli due to PTF. Third, this technique is not capable of differentiating peri- and preprocedural thrombus fragmentation (distal to the POS). However, because previous studies have found a low frequency of preprocedural fragmentation (<10%),27,43 the rate of false-positive classification is expected to be rather low. Fourth, the grading of the NE of the clot is based on a semiquantitative scale and was performed visually. Subjective analysis may, in the future, be overcome by ensuring comparable staining intensity and following threshold-based software quantification (comparable with the algorithm used for the analysis of H&E staining). Fifth, because we analyzed the thrombus retrieved from the POS, the histologic characteristics of the fragments causing distal emboli remain uncertain and might differ from the clot composition of the retrieved thrombus. One may speculate that dissolvable tissue islands or additional apposition thrombi might accumulate at the clot margins. These areas may be more prone to fragmenting during endovascular manipulation than the main thrombus and might show a different histology. Finally, because different stent retrievers were used, recanalization success and subsequent risk of PTF might be influenced by the choice of the device.

CONCLUSIONS
Younger age, easy-to-retrieve thrombi, and bridging thrombolysis are factors that may have a higher risk of PTF. Conventional histologic characteristics (RBC, WBC, F/P) do not seem to influence MT-relevant thrombus stability. However, higher amounts of clot-infiltrating neutrophils were related to an increased risk of PTF. This observation aligns well with the proposed thrombotic capacity of neutrophilic enzymes and may implicate them as a novel marker for clot stability, which should be addressed in future clot studies.

Disclosures: Claus Zimmer—UNRELATED: Board Membership: Philips Healthcare and Bayer Schering Pharma, Comments: scientific advisory boards*; Payment for Lectures including Service on Speakers Bureau: Bayer Schering Pharma and Philips Healthcare, Comments: speaker honoraria; Other: Biagen ldec, Quintiles, MSD Sharp & Dome, Boehringer Ingelheim, Inventive Health Clinical UK Ltd, advanceCOR, BrainsGate, Pfizer, Bayer Schering Pharma, Novartis, Roche, Servier, Penumbra, WCT GmbH, Syngs, SSS International Clinical Research, PPD Germany GmbH, Worldwide Clinical Trials Ltd, phenox, Covidien, Actelion, Medtronic, Medtronic, Harrison Clinical Research, Concentric, Pharmtrace, Reverse Medical Corp, Premier Research German GmbH, Syngis, SSS International Clinical Research, PPD Germany GmbH, Worldwide Clinical Trials Ltd, phenox, Covidien, Actelion, Medtronic, Medtronic, Harrison Clinical Research, Concentric, Pharmtrace, Reverse Medical Corp, Premier Research Germany Ltd, Surpass Medical Ltd, and GlavoSmithKline, Comments: research support and investigator fees for clinical studies*; Lucas Schirmer—UNRELATED: Patents (Planned, Pending, or Issued): WO201516037A1; Comments: I filed a patent for the detection of antibodies against KIR4.1 in a subpopulation of patients with MS, the
REFERENCES


Selective-versus-Standard Poststent Dilation for Carotid Artery Disease: A Systematic Review and Meta-Analysis

O. Petr, W. Brinjikji, M.H. Murad, B. Glodny, and G. Lanzino

ABSTRACT

BACKGROUND: The safety and efficacy of standard poststent angioplasty in patients undergoing carotid artery stent placement have not been well-established.

PURPOSE: We conducted a systematic review of the literature to evaluate the safety and efficacy of carotid artery stent placement and analyzed outcomes of standard-versus-selective poststent angioplasty.

DATA SOURCES: A systematic search of MEDLINE, EMBASE, Scopus, and the Web of Science was performed for studies published between January 2000 and January 2015.

STUDY SELECTION: We included studies with >30 patients describing standard or selective poststent angioplasty during carotid artery stent placement.

DATA ANALYSIS: A random-effects meta-analysis was used to pool the following outcomes: periprocedural stroke/TIA, procedure-related neurologic/cardiovascular morbidity/mortality, bradycardia/hypotension, long-term stroke at last follow-up, long-term primary patency, and technical success.

DATA SYNTHESIS: We included 87 studies with 19,684 patients with 20,378 carotid artery stenoses. There was no difference in clinical \( P = .49 \) or angiographic outcomes \( P = .93 \) in carotid artery stent placement treatment with selective or standard poststent balloon angioplasty. Both selective and standard poststent angioplasty groups had a very high technical success of >98% and a low procedure-related mortality of 0.9%. There were no significant differences between both groups in the incidence of restenosis \( P = .93 \) or procedure-related complications \( P = .37 \).

LIMITATIONS: No comparison to a patient group without poststent dilation could be performed.

CONCLUSIONS: Our meta-analysis demonstrated no significant difference in angiographic and clinical outcomes among series that performed standard poststent angioplasty and those that performed poststent angioplasty in only select patients.

ABBREVIATION: CAS = carotid artery stent placement

Endovascular therapy of carotid artery disease has advanced during the past decade and is now considered a valuable treatment alternative to surgery in appropriately selected patients. The indications for carotid endarterectomy were initially established in the North American Symptomatic Carotid Endarterectomy Trial in 1991, which expanded treatment indications to patients with symptomatic severe or moderate carotid stenoses. Formerly, patients who were not eligible for surgery were treated with percutaneous transluminal balloon angioplasty, first described by Kerber et al in 1980. Although procedure-related complication rates were similar/comparable for both treatment modalities, some potential drawbacks and specific problems occurred due to the endovascular approach, including luminal compromise from catheters and guidewires crossing the stenotic lesions and/or during balloon inflation (temporary carotid occlusion by a balloon and/or wire catheter), intraprocedural thromboembolic events, elastic vessel recoil, or intimal dissection. 

Received August 29, 2016; accepted after revision December 12.

From the Departments of Neurologic Surgery (O.P., G.L.) and Radiology (W.B., G.L.) and Division of Preventive Medicine (M.H.M.), Mayo Clinic, Rochester, Minnesota; and Departments of Radiology (B.G.) and Neurosurgery (O.P.), Medical University Innsbruck, Innsbruck, Austria.

Please address correspondence to Ondra Petr, MD, Mayo Clinic, 200 1st St SW, Rochester, MN 55905; e-mail: petr.ondra@mayo.edu; ondra.petr@yahoo.com

Indicates article with supplemental on-line tables.

http://dx.doi.org/10.3174/ajnr.A5003

ter the carotid artery stent placement technique was developed, stent-assisted balloon angioplasty showed better results in event-free survival and even lower repeat angioplasty rates. The primarily used balloon-expandable stents were increasingly replaced by self-expanding stents, exhibiting an intrinsic radial expansion force with memory on the stenotic vessel wall. Poststent balloon angioplasty may then be performed to closely appose the stent and intima and, moreover, to expand regions of residual stent narrowing.11

Supporters of standard poststent balloon angioplasty (per protocol) indicated that poststent ballooning decreased the incidence of restenosis by re-establishing the normal luminal diameter. However, numerous studies have suggested that poststent balloon dilation increases the likelihood of postprocedural emboli. Moreover, poststent ballooning can increase the probability of reflex bradycardia and hypotension, which might be associated with higher rates of periprocedural and postprocedural complications.16,19

Some authors claim that poststent dilation should be performed on a selective, case-by-case basis to maximize patient benefits and limit complications. However, to the best of our knowledge, there is no evidence in the recently published literature supporting the superiority of either of these techniques. Standard poststent balloon angioplasty has become the standard of care in many vascular centers, and only some interventionalists prefer performing poststent angioplasty on a selective base. On the basis of the latter studies, standard poststent balloon angioplasty may be associated with additional risks in patients with acceptable angiographic results, without additional post–carotid artery stent placement (CAS) angioplasty.

To evaluate the safety and efficacy of standard poststent angioplasty versus selective poststent angioplasty, we conducted a systematic review and meta-analysis and analyzed outcomes by a series that performed standard poststent balloon angioplasty per protocol on all patients versus those that performed selective poststent balloon angioplasty on only a subset of patients.

MATERIALS AND METHODS

Study Selection

A comprehensive review of the literature was performed by using the keywords “carotid stenosis,” “carotid artery disease,” “revascularization,” “carotid,” “stent,” “angioplasty,” and “endarterectomy” in both “AND” and “OR” combinations to search PubMed, Ovid MEDLINE, Ovid EMBASE, Scopus, and the Web of Science. Inclusion criteria were the following: English language; >30 patients; studies published between January 2000 and January 2015; studies that performed poststent angioplasty regardless of patient selection for this procedure; and studies with adequate data on periprocedural and postprocedural complications, outcome and technical success, and primary patency. The exclusion criteria were the following: case reports; in vitro, cadaveric or animal studies; studies with no poststent angioplasty; review articles, guidelines, and technical notes. In case of any inconsistencies or differences with regard to study inclusion/exclusion into the meta-analysis, the senior author decided on inclusion or exclusion (G.L.).

The electronic search was supplemented by contacting experts in the field and reviewing the bibliographies of included studies for relevant publications. Abstracts, methods, results, figures, and tables of full text for detailed review were searched by 2 independent reviewers (neurosurgeon O.P. and radiologist W.B.) for data on poststent balloon angioplasty selection, technical success, long-term primary patency, procedure-related morbidity and mortality, and possible selection-related complications such as periprocedural hypotension and bradycardia. The reference lists of retrieved articles were also screened for additional studies. Furthermore, in case of multiple publications from the same institution and/or the same authors, only the most recent and updated study was considered to avoid inclusion of overlapping patients.

Definition of Treatment Groups

The objective of this study was to determine whether there was any difference in angiographic and clinical outcomes among series in which standard poststent angioplasty was performed and those in which poststent angioplasty was performed in only select cases. Studies were categorized as either a “standard poststent balloon angioplasty” series or a “selective poststent balloon angioplasty” series. Standard poststent balloon angioplasty series were defined as those in which poststent angioplasty was reportedly performed in all patients, whereas selective poststent balloon angioplasty series were defined as those in which poststent balloon angioplasty was performed in select cases (ie, residual stenosis, poor wall apposition, and so forth).

Data Abstraction

For each study, we extracted the following descriptive clinical and anatomic information: patient demographics, initial clinical status, and the type of patient selection for the poststent balloon angioplasty (standard versus selective). We studied the following outcomes: periprocedural stroke rates with differentiation of minor and major stroke, periprocedural TIA, procedure-related neurologic or cardiovascular morbidity and mortality, periprocedural myocardial infarction, bradycardia and hypotension rates in all patients distinguishing between hemodynamic changes requiring intervention, long-term stroke rate at last follow-up, long-term primary patency, and technical success. Periprocedural complications were defined as those occurring within 30 days of the carotid artery stent placement.

Statistical Analysis

We estimated from each study the cumulative incidence (event rate) and 95% confidence interval for each outcome. Event rates for each intervention were pooled in the meta-analysis across studies by using the random-effects model. Consequently, there was no need for establishing a hierarchy of analyzed outcomes. Subgroup interactions were conducted by using an interaction test as described by Altman and Bland.43 For all outcomes, we quantified between-study heterogeneity by calculating the I² statistics.4,43 Anticipating heterogeneity between studies, we chose this model a priori because it incorporates within-study variance and between-study variance. We were unable to test for publication bias due to the noncomparative nature of these studies.
RESULTS

Literature Review
An initial comprehensive literature search yielded 1585 articles. Eighteen studies were removed as duplicates. On the initial abstract and title review, 1099 were excluded because they were deemed not relevant to the current study. Four hundred sixty-eight studies were reviewed in additional detail; 283 studies were irrelevant because they lacked information about the use of post-stent balloon angioplasty and/or postprocedural outcome/complications of patients. Ninety-eight additional studies were excluded because they were either case reports or had too few patients. In total, 87 studies with 19,684 patients with 20,378 carotid artery stenoses were included. Series reporting standard poststent angioplasty included 16,983 procedures (83.3%), and series reporting selective poststent angioplasty included 3395 procedures (16.7%). Sixty included studies had ≥100 patients. Eleven studies were prospective, and 76 were retrospective. Data are summarized in On-line Table 1.

Methodologic characteristics of included studies are listed in On-line Table 2. A flow diagram describing our literature search process is provided in Fig 1.

Complication Rates by Type of Poststent Angioplasty
There were no statistically significant differences in any of the periprocedural and long-term complication rates by type of poststent angioplasty. Studies reporting selective poststent balloon angioplasty had similar rates of periprocedural stroke (2.3%; 95% CI, 1.8%–3.0%) compared with those reporting standard poststent balloon angioplasty (2.6%; 95% CI, 2.2%–3.1%) (P = .36). The same was true for long-term stroke rates (1.3% versus 1.6%, P = .49). Major stroke rates were similar in the selective poststent angioplasty group compared with the standard poststent angioplasty group (1.2% versus 1.0%, P = .44). There was no difference in minor stroke rates in the selective poststent angioplasty group (1.3% versus 1.7%, P = .19). There was no difference in periprocedural TIA rates either (1.7% versus 2.2%, P = .43). The periprocedural myocardial infarction rate was 0.6% (95% CI, 0.4%–1.1%) in the selective poststent angioplasty group versus 0.7% (95% CI, 0.5%–1.1%) in the standard poststent angioplasty group (P = .66). These data are summarized in On-line Table 3.

Angiographic Long-Term Results by Type of Poststent Angioplasty
There were no statistically significant differences in any of the long-term primary patency rates or technical success by type of
Studies reporting selective poststent angioplasty had similar long-term primary patency rates (94.3%; 95% CI, 90.7%–96.6%) compared with those undergoing standard poststent angioplasty (94.5%; 95% CI, 92.5%–95.9%) (P = .93). Technical success was 98.7% (95% CI, 97.1%–99.5%) in the selective poststent angioplasty group versus 99.0% (95% CI, 98.6%–99.3%) in the standard poststent angioplasty group (P = .61). These data are summarized in On-line Table 3.

**Study Heterogeneity and Characteristics**

Significant heterogeneity (I² value > 50% and P value for the Cochrane Q test < .05) was noted in the analyses of 2 outcomes: bradycardia/hypotension and bradycardia/hypotension requiring intervention. Therefore, confidence in a pooled summary estimate for these 2 outcomes is limited. I² values are summarized in On-line Table 3. Methodologic characteristics of included studies are listed in On-line Table 2.

**DISCUSSION**

This systematic review and meta-analysis of 87 studies with 19,684 patients reporting either selective and standard poststent balloon angioplasty following carotid artery stent placement demonstrated no difference in clinical or angiographic outcomes in the CAS treatment with selective or standard poststent balloon angioplasty. In our study, we also found that both selective and standard poststent angioplasty groups had very high technical success rates of >98% and very low procedure-related mortality rates of 0.9%. Furthermore, there were no statistically significant differences between both groups in the incidence of restenosis or in procedure-related complication rates. Last, despite the selective poststent balloon angioplasty being associated with higher rates of periprocedural bradycardia/hypotension events, there was a very slight trend toward lower rates of periprocedural TIA’s in these patients. These findings are important because they suggest that standard poststent angioplasty is not required during carotid stent placement. The similar rates of primary patency between groups suggest that the addition of standard poststent angioplasty does not provide any definite benefit.

Comparisons of clinical and angiographic outcomes between standard and selective poststent dilations in the literature are limited largely due to the small sizes of most case series. Numerous transcranial Doppler studies demonstrate the presence of emboli with each passage across a stenosis with a guidewire, embolic protection device, balloon, or stent, with the highest potential for embolization occurring during poststent dilation when the balloon pushes the stent struts against the atheromatous plaque. Ackerstaff et al reported in a series of 550 patients that multiple microemboli (>5 showers) at poststent angioplasty were independently associated with neurologic deficits. While our study found no statistically significant differences between the selective and the standard poststent angioplasty groups in terms of risk of stroke and TIA, the standard poststent angioplasty producing no benefit in primary patency suggests that this additional procedure may not always be necessary.

There has been considerable debate in the literature with regard to hemodynamic depression as a possible predictor of adverse events. Some studies suggested that hypotension may result in a greater incidence of periprocedural complications, and even death. To our knowledge, in the largest study to date of 103 patients evaluating the effect of poststent ballooning on hemodynamic stability during and after carotid stent placement, Qazi et al demonstrated that poststent balloon angioplasty was a significant predictor of hemodynamic depression (OR, 3.8; 95% CI, 1.3–11; P < .01) with increased risk of major adverse cardiovascular events. Gupta et al showed that patients with persistent hypotension are at a higher risk of developing an adverse clinical event such as stroke or death after CAS. On the other hand, our study found higher rates of brady/hypotension in patients treated with selective poststent angioplasty compared with standard poststent angioplasty; however, there was no impact on procedure-related complications. The reasons behind this surprising finding in our study are not known. Presumably, patients who undergo selective post-CAS dilation are more likely to have rigid stenoses or insufficient alignment of the stent with the vessel wall compared with those undergoing standard post-CAS angioplasty. Consequently, selective post-CAS dilation is then performed in these cases. During this procedure, relevant hemodynamic events such as bradycardia and/or hypotension can occur, reflecting forced opening of the rigid stenosis.

**Limitations**

We acknowledge that our meta-analysis has several limitations. Studies reporting only carotid artery stent placement without subsequent angioplasty were excluded. With this design, no comparison with a group without poststent dilation could be performed. There was a paucity of studies comparing outcomes of patients who received no poststent angioplasty and those who did. Our results should not be interpreted as saying that poststent angioplasty is ineffective because it is likely useful on a case-by-case basis. Rather, our results should be interpreted to say that there is no difference in angiographic and clinical outcomes among patients undergoing standard poststent angioplasty and those undergoing selective poststent angioplasty.

Ecologic bias (eg, comparisons are made across studies and not within studies), the possibility of publication bias, and statistical heterogeneity are important limitations that affect inferences derived from this study. None of the included studies were randomized or included control groups. There was no detailed information regarding the indications for selective poststent angioplasty. Similarly, there was no information about residual stent narrowing before performing balloon angioplasty. However, given the contemporary and widely accepted literature evidence for treatment of carotid artery disease, all included studies most likely had similar strict indications for carotid artery stent placement.

**CONCLUSIONS**

This meta-analysis of >87 series reporting selective and standard poststent balloon angioplasty following carotid artery stent placement demonstrated that both standard and selective approaches were associated with low rates of procedure-related neurologic or cardiovascular morbidity and high rates of long-term primary patency of >94%. There were no statistically
significant differences in clinical or angiographic outcomes between series reporting standard and selective poststenotic angioplasty. Comparative prospective studies are needed to confirm our findings.

Disclosures: Giuseppe Lanzino—UNRELATED: Consultancy: Covidien/Medtronic.* *Money paid to the institution.

REFERENCES

2. Ederle J, Bonati LH, Dobson J, et al; CAVATAS Investigators. Endo-

vascular treatment with angioplasty or stenting versus endarter-

3. Ederle J, Dobson J, Featherstone RL, et al; International Carotid Stenting Study investigators. Carotid artery stenting compared with endarterectomy in patients with symptomatic carotid steno-

4. Yadav JS, Wholey MH, Kuntz RE, et al; Stenting and Angioplasty with Protection in Patients at High Risk for Endarterectomy Investi-

6. North American Symptomatic Carotid Endarterectomy Trial Col-

aborators. Beneficial effect of carotid endarterectomy in symp-


9. Kerber CW, Cromwell LD, Loehden OL. Catheter dilatation of proximal carotid stenosis during distal bifurcation endarterec-


cedure systolic blood pressure changes with neurological events in high-risk carotid stent patients. *J Endovasc Ther* 2002;9:810–16 CrossRef Medline
18. Park BD, Divinagracia T, Madej O, et al. Predictors of clinically significant postprocedural hypotension after carotid endarterec-

19. Qazi U, Obeid TE, Enwerem N, et al. The effect of balloon fol-


plasty and stenting of cervicalcerebral arteries in Iran. *Asian Car-

otid Thorac Ann* 2010;18:49–53 CrossRef Medline

tected carotid artery stenting in symptomatic and asymptomatic patients. *J Endovasc Ther* 2010;17:298–307 CrossRef Medline

terv* 2010;76:9–15 CrossRef Medline
23. Montorsi P, Galli S, Ravagnani P, et al. Randomized trial of predi-

lation versus direct stenting for treatment of carotid artery steno-

sis. *Int J Cardiol* 2010;138:233–38 CrossRef Medline
25. Almekhlafi MA, Couillard PL, Pandya A, et al. Outcomes after ca-

rotid angioplasty and stenting in symptomatic octogenarians. *Can J Neurol Sci* 2011;38:446–51 CrossRef Medline

tions after carotid artery stenting: a contemporary experience from the University at Buffalo neuroendovascular surgery team. *Neurosurvory* 2013;73:689–93; discussion 693–94 CrossRef Medline
31. Lago A, Parkhutik V, Tembl JI, et al. Diabetes does not affect out-


son between proximal versus distal protection devices in 287 cases of carotid revascularization using angioplasty and stenting. *Periprocedure complications, morbidity, and mortality. Cardio-

vasc Intervent Radiol* 2014;37:639–45 CrossRef Medline
cerebral protection in 100 consecutive patients: immediate and two-year follow-up results. Ital Heart J 2003;4:695–700 Medline
stroke, myocardial infarction, and death. Radiology 2006;240:145–51 CrossRef Medline
Presurgical Brain Mapping of the Ventral Somatomotor Network in Patients with Brain Tumors Using Resting-State fMRI


ABSTRACT

BACKGROUND AND PURPOSE: Resting-state fMRI readily identifies the dorsal but less consistently the ventral somatomotor network. Our aim was to assess the relative utility of resting-state fMRI in the identification of the ventral somatomotor network via comparison with task-based fMRI in patients with brain tumor.

MATERIALS AND METHODS: We identified 26 surgically naïve patients referred for presurgical fMRI brain mapping who had undergone both satisfactory ventral motor activation tasks and resting-state fMRI. Following standard preprocessing for task-based fMRI and resting-state fMRI, general linear model analysis of the ventral motor tasks and independent component analysis of resting-state fMRI were performed with the number of components set to 20, 30, 40, and 50. Visual overlap of task-based fMRI and resting-state fMRI at different component levels was assessed and categorized as full match, partial match, or no match. Rest-versus-task-fMRI concordance was calculated with Dice coefficients across varying fMRI thresholds before and after noise removal. Multithresholded Dice coefficient volume under the surface was calculated.

RESULTS: The ventral somatomotor network was identified in 81% of patients. At the subject level, better matches between resting-state fMRI and task-based fMRI were seen with an increasing order of components (53% of cases for 20 components versus 73% for 50 components). Noise-removed group-mean volume under the surface improved as component numbers increased from 20 to 50, though ANOVA demonstrated no statistically significant difference among the 4 groups.

CONCLUSIONS: In most patients, the ventral somatomotor network can be identified with an increase in the probability of a better match at a higher component number. There is variable concordance of the ventral somatomotor network at the single-subject level between resting-state and task-based fMRI.

ABBREVIATIONS: BOLD = blood oxygen level–dependent; ICA = independent component analysis; rs-fMRI = resting-state fMRI; tb-fMRI = task-based fMRI; VSMN = ventral somatomotor network; VUS = volume under the surface

Functional MRI is widely used as a noninvasive tool for presurgical localization of the eloquent cortex, typically involving somatomotor and language mapping. Mapping of these eloquent brain areas with fMRI correlates well with invasive methods such as intraoperative electrocortical stimulation and can result in reduced surgical time, increased extent of resection, and decreased craniotomy size. In addition, postoperative morbidity correlates with the distance of the resection margin from fMRI-identified eloquent cortex. In current clinical practice, changes in blood oxygen level–dependent (BOLD) signal are measured across time as the patient performs a specific task (ie, task-based fMRI [tb-fMRI]). These BOLD signal changes reflect characteristic hemodynamic responses to neural activity. However, tb-fMRI has several limitations. Primarily, accurate localization of function is dependent on the patient’s cooperation and ability to...
adequately perform the task, which can be an important limitation in those with physical or mental debilitation or those who cannot understand the instructions such as children. The patient must be awake and cooperative during the task; therefore, sedation cannot be used; this feature is a limitation in the pediatric population. Furthermore, due to the inherently low signal-to-noise ratio of this technique, long and repeated acquisitions are often required to ensure adequate data sampling for analysis.\textsuperscript{5,6}

Resting state fMRI (rs-fMRI) is a promising method of assessing brain function that can overcome some of the limitations of tb-fMRI. In rs-fMRI, spontaneous fluctuations in BOLD signal are measured across time while no specific task is performed. Temporal correlations of these spontaneous fluctuations can be organized into spatially distinct intrinsic networks.\textsuperscript{5} These networks are now described as resting-state networks, many of which have topography similar to that of networks engaged in sensory, motor, and cognitive tasks.\textsuperscript{7} These intrinsic resting-state networks persist, though somewhat modified, in states of decreased awareness such as sleep\textsuperscript{8} or sedation.\textsuperscript{9} Rs-fMRI has thus gained interest as a potential viable alternative to tb-fMRI, especially in pediatric or cognitively/neurologically impaired patients. Multiple studies have investigated the potential role of rs-fMRI in presurgical mapping of somatomotor\textsuperscript{5,10-12} and language\textsuperscript{2,6,13} networks. In a limited number of subjects, moderate overlap between rs-fMRI and tb-fMRI was found for mapping of the motor cortex.\textsuperscript{5,11,12} Furthermore, there is good qualitative concordance between intraoperative cortical stimulation and rs-fMRI in the localization of the eloquent motor cortex.\textsuperscript{10} With quantitative analysis, rs-fMRI and tb-fMRI perform comparably, but the shortest distance to stimulation points is observed for tb-fMRI.\textsuperscript{11} Rs-fMRI has been shown to identify a larger pattern of the motor network compared with tb-fMRI.\textsuperscript{11,14} In addition, localization by rs-fMRI and tb-fMRI may include different parts of the sensorimotor network.\textsuperscript{11,15} While the motor system at large has been the target of most prior investigations, distinct subnetworks of the motor network exist. We narrowed our focus to the ventral motor area, which largely reflects oral somatomotor function. Three recent studies evaluated the concordance of rs-fMRI and tb-fMRI in the localization of the face representation area of the primary motor cortex in addition to limb motor areas in a limited number of patients with a variety of brain lesions in different locations.\textsuperscript{11,12,16}

To our knowledge, no study has specifically investigated the concordance between tb-fMRI activation and the rs-fMRI-derived ventral somatomotor network (VSMN) maps in a large cohort of patients with lesions close to or involving the VSMN. We hypothesized that the ventral motor network can be identified in patients with brain tumors and that there is good concordance with tb-fMRI with a tongue motor paradigm. In addition, we hypothesized that a higher number of components in independent component analysis (ICA) may yield a better concordance between rs-fMRI and tb-fMRI.

**MATERIALS AND METHODS**

**Study Subjects**

The institutional review board approved this retrospective study. Searching the Radiology Information Systems, we identified patients who underwent fMRI for presurgical brain mapping between January 1, 2009, and July 31, 2014. Fifty-eight patients underwent rs-fMRI in addition to tb-fMRI during the same imaging session. Seventeen patients had a prior history of a brain operation (including biopsy) and were excluded. One patient had imaging features characteristic of an arteriovenous malformation and was excluded. At the Johns Hopkins Hospital, tb-fMRI for presurgical mapping is tailored for each patient, predominantly based on the location of their tumor and the relation of eloquent regions along the expected surgical trajectory and in the immediate vicinity of the lesion. Among the remaining 40 patients, 26 with brain tumors involving or in close proximity to the ventral somatomotor network who had undergone both tb-fMRI for localization of the VSMN and rs-fMRI were included for final analysis (age range, 21–69 years; mean age, 43.6 years; 15 men and 11 women).

**Brain Tumor Characterization**

Tumor location, volume (in cubic millimeters), pathology, and World Health Organization histologic grade (when available) were recorded by a subspecialty board-certified neuroradiologist for each patient. Lesion volume was measured by manually drawing the ROI on FLAIR images in the Medical Image Processing, Analysis, and Visualization application software (MIPAV; National Institutes of Health, Bethesda, Maryland; http://mipav.cit.nih.gov). For high-grade brain tumors, we assessed the entire region of T2 signal abnormality, which represents a combination of infiltrative neoplasm and vasogenic edema.

**MR Imaging**

A 3T Tim Trio system (Siemens, Erlangen, Germany) with a 12-channel head matrix coil was used. Structural images included a 3D T1 sequence (TR = 2300 ms, TI = 900 ms, TE = 3.5 ms, flip angle = 9°, FOV = 24 cm, acquisition matrix = 256 × 256 × 176, section thickness = 1 mm) and a 2D T2 FLAIR sequence (TR = 9310 ms, TI = 2500 ms, TE = 116 ms, flip angle = 141°, FOV = 24 cm, acquisition matrix = 320 × 240 × 50, section thickness = 3 mm). Functional T2*-weighted BOLD images for both tb-fMRI and rs-fMRI were acquired by using 2D gradient-echo echo-planar imaging (TR = 2000 ms, TE = 30 ms, flip angle = 90°, FOV = 24 cm, acquisition matrix = 64 × 64 × 33, section thickness = 4 mm, section gap = 1 mm, interleaved acquisition). Instructions for rs-fMRI were the following: try to keep still, keep your eyes closed, and do not fall asleep. For rs-fMRI, 180 volumes were acquired (6 minutes).

**Tongue Motor Task**

The face representation area of the primary motor cortex was mapped by using the vertical tongue movement task as a robust and spatially extensive representative of the ventral face motor network.\textsuperscript{17,18} The duration of the task is 3 minutes and consists of 3 cycles of 30-second blocks of rest alternating with 30-second blocks of repetitive vertical tongue movement. For each patient, a board-certified neuroradiologist with experience in fMRI provided instructions and practice sessions on the tongue motor task outside the scanner. The quality of the tb-fMRI maps was monitored in real-time by the neuroradiologist monitoring the session. Per protocol, if real-time maps dem-
onstrated suboptimal activation or excessive noise, the task was repeated. Tasks were presented with the environment implemented in the Prism Acquire software (Prism Clinical Imaging, Elm Grove, Wisconsin). Only runs that met quantitative quality control criteria (<2-mm net head displacement along any axis) were considered.

Image Processing
fMRI data were processed by using Statistical Parametric Mapping, Version 8 (SPM8 software; http://www.fil.ion.ucl.ac.uk/spm/software/spm12) and custom Matlab (MathWorks, Natick Massachusetts) scripts. Processing of tb-fMRI included section-timing correction, motion correction, normalization to a Montreal Neurological Institute 152 template, and spatial smoothing included a 6-mm full width at half maximum Gaussian kernel. For processing the rs-fMRI data, section-timing correction and motion correction were performed. The motion-correction step included registration of tb-fMRI and rs-fMRI to each other. The ArtRepair toolbox (http://cibs.stanford.edu/tools/human-brain-project/artrepair-software.html)19 was used to detect volumes with large shifts in global average signal intensity, which include contributions from scan-to-scan motion. The outlier volumes and additional volumes recommended for deweighting in ArtRepair were tagged for subsequent removal from analysis (ie, scrubbing). The rs-fMRI data were then linearly detrended. Rs-fMRI and T1-weighted images were coregistered and normalized to perform physiologic nuisance regression of rs-fMRI with the component-based noise-correction method.20 The same transformation matrix was used for normalization between tb-fMRI and rs-fMRI to ensure that spatial comparison between these 2 was valid for each subject. Bandpass filtering from 0.01 to 0.1 Hz and smoothing were performed with a 6-mm full width at half maximum Gaussian kernel. At the end, previously tagged images by ArtRepair were scrubbed.

Statistical Analysis
Tb-fMRI Analysis. We used a general linear model analysis for tb-fMRI implemented in SPM8 with the canonical hemodynamic response function convolved with the boxcar function for each task with standard parameters previously described.6 High-pass filtering was performed at the default setting to remove drift. The hemodynamic response function-convolved task vectors were input into a design matrix, and a contrast was created to target the associated parameter comparing the hemodynamic response function-weighted time engaged in the task with rest. The SPM t-contrast maps were generated without clustering because these activation maps were subsequently thresholded across multiple levels for comparison as previously described.6 Activation maps were reviewed to ensure that ventral motor activation was present.

Rs-fMRI Analysis. Rs-fMRI was analyzed with the Group ICA of the fMRI Toolbox Software (GIFT; http://mialab.mrn.org/software/gift/). Independent component analysis was performed separately for each subject by using the InfoMax algorithm with ICASSO (http://research.ics.aalto.fi/ica/icasso/) set at 5 repeats,21 including selection of the “best run” to ensure consistent estimates. ICA maps were generated for 20, 30, 40, and 50 components, designated here as ICA20, ICA30, ICA40, and ICA50, respectively. Following scrubbing, 1 subject had only slightly >50 volumes left; therefore, the maximum number of components was limited to 50. For each ICA group, the component that best represented the VSMN based on overlap with tb-fMRI was selected visually (Fig 1). Any component whose spatial map was specific and limited to the ventral periorbital cortices corresponding to the tb-fMRI activation maps was considered a “full match.” If the component included additional networks outside these regions, it was considered a “partial match,” indicating that the VSMN was present but not separated from other networks. If no VSMN was identified, a “no match” designation was assigned to that ICA group for that subject. A mixed-effects logistic regression analysis was performed to see whether there was a significant increase in the probability of getting a full match as a function of the number of ICA components.

Comparison of tb-fMRI and rs-fMRI. The level of tb-fMRI T-map thresholding may affect the degree of rs-tb fMRI concordance. At low thresholds, there will be artificially high concordance due to the introduction of a higher percentage of voxels, many of which may not represent true activation but rather statistical noise that.

FIG1. Task-based fMRI map demonstrating activation of the VSMN with the vertical tongue movement task in a single subject (A). Rs-fMRI maps with the best visual correlates to tb-fMRI at different ICA groups (B–E) were selected.
exceeds this threshold. Although there are strategies to determine the level of optimum fMRI thresholding, there is no consensus for reproducible results at a subject level.\textsuperscript{18} To minimize this issue, we used a multithresholding technique to compare concordance across a wide range of thresholds as previously described (Fig 2).\textsuperscript{6,22} We calculated Dice coefficients as a quantitative measure of overlap across each threshold to generate a matrix of rs-versus-tb overlap (left map, Fig 3).\textsuperscript{6} Dice coefficients vary between 0 and 1 and give an objective evaluation of similarity or concordance between 2 sets of data. At very low thresholds, high Dice coefficients result from overlap of random noise as can be seen in Fig 2. We used a previously described noise-removal method\textsuperscript{6} briefly described here.

For calculation of noise, an ICA component representing the anterior ventricular signal was selected for each subject. After removing negative values, we normalized image maps in value from 0 to 1, and we subsequently used multiple thresholds of this normalized map to calculate Dice coefficients between ventricular "noise" and task maps (middle map, Fig 3). The resultant noise-versus-task Dice map was subtracted from the resting-versus-task Dice map (right map, Fig 3).\textsuperscript{6} Noise-corrected resting state-versus-task-based Dice map volume under the surface (VUS) was calculated across different ICA orders. The VUS measurement collapses this multithresholded dice map into a single variable that can be used as a metric; the VUS is equivalent to the area under the curve in a 2D graph (such as a receiver operating characteristic curve); however, because the multithresholded dice map has 3 dimensions (the rs-fMRI threshold, the task-threshold, and value), a volume under the surface is computed. One-way ANOVA was performed to determine significant differences in VUS across the 4 ICA orders. Maximum Dice coefficient values at the group level and subject level were calculated.

RESULTS

Tumors

The On-line Table summarizes the patients’ demographic data, location and volume of lesions, pathology, World Health Organization grade (if applicable), and the distance from the edge of the lesion to the edge of the ventral somatomotor activation cluster based on the clinical task-fMRI maps. Brain lesions were mainly centered in the left cerebral hemisphere (19 patients), and 4 lesions were primarily located in the right cerebral hemisphere. Tumor size ranged from 0.84 to 159.05 cm\textsuperscript{3} (mean, 41.82 cm\textsuperscript{3}).

FMRI Comparison

Twenty-six patients met the inclusion criteria and were included for analysis. In 21 patients (81\%), rs-fMRI successfully identified a VSMN (Fig 4). Among these, 2 patients were considered a partial match because the VSMN was not distinct from the dorsal somatomotor network. There was an increase in the number of full matches with an increase in the ICA order (14 patients at ICA 20 and 19 patients at ICA 50). In 14 patients, a separate VSMN was identified at all ICA groups. Mixed-effect logistic regression demonstrated a significant increase in the probability of getting a full match as a function of the number of components (P < .00001). This probability is shown in Table 1.

In 5 patients, rs-fMRI failed to identify a VSMN at any of the ICA levels. In 3 of these 5 patients, only the dorsal somatomotor

FIG 2. Rs-fMRI ICA maps (yellow) and Tb-fMRI T-maps (red) were thresholded, and 100 threshold maps were generated for each in a single subject. At a low threshold, there is artificially higher overlap (orange) between the maps due to noise. At a very high threshold, the overlap is smaller.

FIG 3. The Dice coefficient matrix at different thresholds at the subject level between the rs-fMRI (x-axis) at ICA 20 and tb-fMRI (y-axis). An artificially high Dice coefficient is seen in the top left corner of the left map due to overlap of noise. A noise matrix was generated (middle map) and was subtracted (right map). Noise-removed Dice coefficient maps were generated for all the subjects across 4 different ICA orders.
network was identified. In the remaining 2 patients, no somatomotor network was identified.

Group mean Dice maps for each ICA order are shown in Fig 5. Group mean Dice VUS overall increased with the ICA order (Fig 6); however, 1-way ANOVA demonstrated no significant differences among the 4 ICA groups ($P$ value = .4). The range of Dice coefficient values at the subject level is shown in Table 2. Negative Dice values in these noise-subtracted maps may occur when the concordance between the noise maps and tb-fMRI at low thresholds is greater than the concordance between rs-fMRI and tb-fMRI due to the randomness of noise at low thresholds.

**DISCUSSION**

A growing number of studies have explored the feasibility of rs-fMRI as a substitute or complement to tb-fMRI for presurgical mapping of the eloquent somatomotor cortex in patients with brain tumor. Previous studies have demonstrated the reliability of rs-fMRI for localization of the hand motor area in comparison with tb-fMRI in a limited number of patients, with seed-based or ICA analysis. Concordance of rs-fMRI with intraoperative cortical stimulation was assessed in a few patients. In addition, investigators have assessed the entire somatomotor cortex by using hand, foot, and face paradigms, with evaluation of the face motor area in a small subset of patients. A novel data-driven algorithm for mapping the functional cortex in preoperative planning of motor and language networks in a small group of patients with tumor and epilepsy has been reported. The existing data predominantly evaluate the usefulness of rs-fMRI in localization of the hand representation area of the somatomotor cortex, but there has been very limited evaluation of the ventral somatomotor cortex to date. In addition, there is wide variation in analysis methods, and most studies have only qualitatively compared tb-fMRI and rs-fMRI. To our knowledge, our study represents a relatively large cohort of patients with brain tumors involving or in close spatial proximity to the VSMN. In addition, we evaluated the overlap quantitatively by using multithresholding at different ICA levels.

![FIG 4](image-url) | ![FIG 5](image-url) | ![FIG 6](image-url) | Table 1: Probability of getting a full match between rs-fMRI and tb-fMRI maps as a function of number of ICA components

<table>
<thead>
<tr>
<th>ICA 20</th>
<th>ICA 30</th>
<th>ICA 40</th>
<th>ICA 50</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.53</td>
<td>0.65</td>
<td>0.69</td>
<td>0.73</td>
<td></td>
</tr>
</tbody>
</table>

![Table 2: Dice coefficient values range, median, and mean at the subject level across different ICA orders](image-url)
Our results show a strong concordance between the rs-fMRI and tb-fMRI in some of the subjects, demonstrating the potential utility of rs-fMRI as a viable preoperative mapping tool (Fig 7); however, there is significant variability across subjects. In particular, if rs-fMRI is to be used without a tb-fMRI acquisition, one must ensure that a reliability estimate of rs-fMRI can be calculated from the data itself. Improvement in technique and data analysis may overcome some of these limitations in the future. Increasing the scan time may improve the quality of rs-fMRI and thus potentially decrease such intersubject variability. Alternative methods of analysis such as seed-based analysis that have shown promise in localizing the motor network in patients with brain tumors may also be considered. However, in the current study population of patients with brain tumors, accurate placement of seeds may be limited due to anatomic gyral distortion or compression. Therefore, an unbiased method such as ICA may be preferable. In addition, ICA remains an attractive choice due to availability of easy-to-use software such as the MELODIC tool in FSL (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/MELODIC) or the GIFT toolbox.

Several limitations in this study should be addressed. rs-fMRI was performed after tb-fMRI in this clinical cohort of patients to ensure that the patients tolerated the lengthy scan and performed well on their tb-fMRI, which was critical for clinical presurgical mapping. Thus, there may have been an inadvertent task effect on the observed functional connectivity. Another limitation is the inability to ensure that patients did not fall asleep during rs-fMRI. We instructed the patients to stay awake during the rs-fMRI and confirmed that they stayed awake during a postscan interview; however, the accuracy of their statements could not be verified by physiologic measures indicating sleep during the acquisition. While this inability poses a potential limitation because changes in functional connectivity have been reported in sleep or altered consciousness, functional connectivity in the somatomotor network has been shown to be preserved during different states of arousal. Nevertheless, the effect of sleep or altered arousal on rs-fMRI concordance with tb-fMRI could be further investigated. In addition, we used tb-fMRI T-maps to find the best candidate ICA map to represent the VSMN. However, brain tumors may cause alteration of somatomotor network organization, and brain tumor–induced neurovascular uncoupling may further compromise our ability to accurately detect the VSMN. Therefore, using only rs-fMRI data to identify the best somatomotor map may be challenging with brain tumors. The use of a data-driven neural network algorithm to identify the eloquent cortex in 7 patients with brain tumor with distorted anatomy has shown some promise.

CONCLUSIONS

We demonstrate variable concordance of rs- and tb-fMRI at the single-subject level for detection of the VSMN in patients with brain tumor. We demonstrate improved reliability of rs-fMRI VSMN maps with higher ICA orders. Failure of rs-fMRI to identify the VSMN in about one-fifth of patients in our study limits the ability of rs-fMRI to completely substitute for tb-fMRI for presurgical assessment, though it may be considered as a supplement for tb-fMRI, for example, when task activation is suboptimal. Further improvements in technique, processing, and analysis methods may enhance the potential of using rs-fMRI as a substitute or complement to tb-fMRI in mapping of the VSMN.
REFERENCES


23. Li YO, Adali T, Calhoun VD. Estimating the number of independent components for functional magnetic resonance imaging data. Hum Brain Mapp 2007;28:1251–66 CrossRef Medline


Prognostic Predictors of Visual Outcome in Open Globe Injury: Emphasis on Facial CT Findings


ABSTRACT

BACKGROUND AND PURPOSE: The present prognostic models for open globe injuries have a limited ability to predict visual outcome before a comprehensive ophthalmologic examination or operation because they depend on the data derived from the ophthalmologic examination and intraoperative findings. The purpose of our study was to determine the specific CT and preoperative clinical data that can predict the prognosis of open globe injury.

MATERIALS AND METHODS: We analyzed the relationship of 29 variables derived from clinical and CT data from 97 globe injuries with visual acuity at 1 month. A prediction model was derived from 49 globe injuries by regression analysis, followed by receiver operating characteristic curve analysis of the best CT predictor.

RESULTS: Four variables with significance on a regression model were the following: posterior segment hemorrhage ($\beta = -0.93, P < 0.0001$), presenting visual acuity ($\beta = 0.28, P = .042$), orbital emphysema ($\beta = 0.46, P = .0018$), and complex facial fracture ($\beta = -0.43, P = .09$). Receiver operating characteristic analysis of the posterior segment hemorrhage predicted profound vision loss (light perception or no light perception) with an area under the curve of 0.97. The receiver operating characteristic table indicated that grade III posterior segment hemorrhage has a strong positive predictive value of 100% for profound vision loss. On the other hand, the absence of posterior segment hemorrhage has a strong positive predictive value of 93% for mild-to-severe vision loss (visual acuity better than light perception).

CONCLUSIONS: Radiologists, with the help of CT and preoperative clinical data, can predict visual acuity after open globe injury.

ABBREVIATIONS: IOP = intraocular pressure; LP = light perception; logMAR = logarithm of the minimum angle of resolution; NLP = no light perception; PPV = positive predictive value; RAPD = relative afferent pupillary defect; VA = visual acuity

Open globe injury is described as a full-thickness injury of the eye wall resulting from either blunt or penetrating trauma. It is a vision-threatening injury, and primary surgical repair is the standard practice to restore the structural and physiologic integrity of the globe, regardless of the extent of the injury and the presenting visual acuity (VA). Prompt diagnosis and surgical repair of the injury are crucial to optimize visual outcome. After open globe injury, the Ocular Trauma Score system and the Classification and Regression Tree model, that were developed to guide both clinicians and patients in clinical decision-making. The Ocular Trauma Score is derived from presenting VA, globe rupture, endophthalmitis, perforating injury, retinal detachment, and relative afferent pupillary defect (RAPD) as predictors of poor visual outcome. The Classification and Regression Tree model identifies RAPD, presenting VA, lid laceration, and posterior wound location at surgery as predictors of poor outcome. Both models depend on data derived from a comprehensive ophthalmologic examination and intraoperative findings. This requirement limits the ability of the models to consistently provide prognostic information before ophthalmologic examination or an operation. The time between the injury and surgery is crucial because that is when clinicians need the most prognostic information to reduce the patient’s anxiety and assist in informed decision-making regarding treatment choices. A comprehensive ophthalmologic assessment is challenging and is often delayed in the acute trauma setting due to periorbital soft-tissue swelling, poor patient co-
operation, and altered mental status due to concomitant head trauma or the use of mind-altering medications. Subjecting patients with globe rupture to an aggressive ophthalmologic examination can even worsen the initial injury.

In contemporary trauma care, CT has evolved as the imaging technique of choice for evaluating orbital trauma, especially in patients with difficult ophthalmologic examinations. Various studies have shown the sensitivities and specificities of CT, ranging from 56% to 76% and 79% to 100%, respectively, in detecting open globe injuries. CT offers the additional benefit of detecting concomitant intraorbital soft-tissue injuries and orbitofacial fractures. Objective preoperative prognostic data that can be derived from CT findings and limited clinical data obtainable in acute trauma settings would be of great clinical value. Such information could facilitate effective communication and counseling of patients.

Our purpose was to determine the specific CT and preoperative clinical data that can predict the prognosis of open globe injury. VA at 1 month after surgery was used as the reference standard for visual outcome in this study.

MATERIALS AND METHODS
The study was compliant with the Health Insurance Portability and Accountability Act, and permission was obtained from the University of Maryland, Baltimore (UMB) institutional review board. The study was conducted at a level 1 trauma center. The inclusion criteria for this retrospective study were the following: 1) a history of open globe injury and subsequent surgical repair between July 2005 and January 2014, 2) CT of the face performed before surgical repair, 3) subject at least 18 years of age, and 4) clinical follow-up with determination of VA at 1 month after globe repair.

Subjects
A search of the trauma data base of our institution from July 2005 to January 2014 yielded 132 patients with 138 globe repairs; 6 patients sustained bilateral injuries. Only 1 globe was randomly selected from patients with bilateral injuries. One-month VA was available in 97 globes and constituted the study group. CT of the face was performed in 85 of the 97 patients. Figure 1 shows the patient selection flowchart. In the study group, there were 72 men and 25 women (mean age, 41.4 years; range, 18–94 years). The mechanism of injury was blunt trauma in 44% (43 of 97), gunshot wound in 25% (24 of 97), and stab wounds/penetrating injuries by sharp objects in 15% (15 of 97); and in 15% (15 of 97) of patients, the mechanism could not be determined.

Clinical information regarding the complete ophthalmic/pupil examination, if performed, was obtained. The data included presenting VA, RAPD, and intraocular pressure (IOP) measurements. Clinical information was obtained from the initial consultation note, progress notes, and operative reports. A senior ophthalmology resident (J.A.M.) reviewed the medical records.

Imaging Technique
Protocols for multidetector row CT are shown in On-line Table 1. Admission multidetector row CT was performed with a 16-, 40-, or 64 section CT system (Brilliance 16-, 40-, and 64-channel system; Philips Healthcare, Best, the Netherlands). Facial CT was performed from the frontal sinuses through the mental symphysis. The CT images included 2-mm axial sections and reformatted sagittal and coronal images (2-mm thickness, obtained at 1-mm intervals through the face).

Variable Construction
We analyzed the relationship of 29 study variables derived from demographic, clinical, and CT imaging data (Table 1 and On-line Table 2). Predictors of visual outcome identified in the literature were also included, in addition to the CT variables most commonly related to intraorbital soft-tissue injuries and craniofacial fractures.
**Image Analysis and Definitions**

CT studies were loaded onto the thin-client server of our institution (IntelliSpace Portal; Philips Healthcare) to facilitate postprocessing of multiplanar reconstruction images in additional planes. Three trauma attending radiologists (reviewer 1, 5 years of experience; reviewer 2, 8 years of experience; reviewer 3, 10 years of experience), blinded to the clinical data, performed independent reviews of each CT study on the PACS of our institution with additional use of the thin-client software. For all the CT variables, discrepancies between the assessments of 2 reviewers were resolved by adjudication by a third reviewer.

The reviewers assessed and recorded the presence of the CT variables of globe and orbitofacial trauma. All the variables were given nominal scores based on the presence (score of 1) versus the absence (score of 0) of each variable. Apart from the qualitative analysis, a quantitative analysis was performed for posterior segment hemorrhage, fractional decrease in globe volume, and fractional increase in globe volume. Reviewer 3 measured globe volumes on the thin-client server by using semiautomated 3D segmentation. ROIs were drawn on axial sections with the use of sagittal sections to exclude unwanted surrounding bone and soft tissues from the ROIs (Fig 2). Volumes of both globes were obtained in all patients. In patients with lens destruction or severe intraocular hemorrhage too dense to allow visualization of the lens on CT, evaluation for lens subluxation and dislocation and assessment of the anterior chamber depth were not performed. It is quite difficult to distinguish hemorrhage into the vitreous chamber and subretinal or suprachoroidal subtypes in globes with severe injury. Hence, we considered it more appropriate to refer to them collectively as posterior segment hemorrhage (Fig 3).

**Study Term Definitions**

**Grade of Posterior Segment Hemorrhage.** Posterior segment hemorrhage was graded from 0 to III, based on the amount of blood in the posterior segment: grade 0 (no posterior hemorrhage), grade I (<25% of posterior segment filled with blood), grade II (25%–75% filled), and grade III (>75% filled) (Fig 3).

**Fractional Change in Globe Volume.** Fractional increase or fractional decrease in the globe volume was obtained by using the contralateral normal globe volume as a reference (Fig 2). In patients with bilateral globe injuries, the mean volume of all the normal globes included in the study was used as reference.

**Intraorbital Hemorrhage and Emphysema.** Intraorbital hemorrhage and emphysema were further divided into extraconal and intraconal components.

**Facial and Orbital Fractures.** Simple facial fractures were defined as fractures without involvement of facial buttresses. Complex facial fractures, defined as buttress fractures, included naso-orbitoethmoid fractures, zygomaticomaxillary complex fractures, and Le Fort fractures. Orbital fractures included isolated wall fractures and also those commonly associated with posterior propagation of complex midfacial fractures.

**Perforating Injury.** Injuries caused by stab wounds and gunshot wounds were divided into penetrating injury (only entry wound) and perforating injury (both entry and exit wounds present).1

**Profound Vision Loss.** Profound vision loss was defined as VA of light perception (LP) or no light perception (NLP). Mild-to-severe vision loss was further defined as VA better than LP.

**Statistical Analysis**

Statistical analysis was performed by K.S., by using statistical software (R statistical and computing software, Version 3.3.1; http://www.r-project.org/; and JMP 12 software; SAS Institute, Cary, North Carolina). Univariate analysis was performed by the Spearman rank correlation for continuous variables. Categoric vari-
ables were compared by 1-way ANOVA. After testing for homogeneous variance (Levene test), a post hoc analysis was performed by using the Welch and Wilcoxon tests. All significant predictor variables in univariate analysis were incorporated into a multivariate logistic regression model. A backward selection model was used. Variables thought to be of infrequent incidence were excluded to result in a model with less overfitting and greater generalizability. Missing data were dealt with by pair-wise deletion for univariate analysis and by a list-wise deletion for multivariate analysis. The regression equation was derived from a continuous dependent variable (ie, VA). The variance inflation factor and adjusted $R^2$ were also obtained from regression analysis. A $P$ value of .05 was significant. To examine the degree of overfitting of the prediction model, we performed a k-fold $\times 10$ cross-validation. First, the sample was partitioned into 10 equal-sized subsamples. Of the 10 subsamples, 9 were used to develop the regression model and the resulting prediction equation was applied to the left-out sample. This procedure was repeated 10 times, each time rotating the cross-validation subset to derive the root mean square error and standard error for the root mean square error.

VA was converted into a logarithm of the minimum angle of resolution (logMAR) units (On-line Table 3) to provide a numeric scale for statistical analysis. Diagnostic performance of the strongest CT predictor derived from the regression analysis was further analyzed by receiver operating characteristic curves after converting the visual outcome into a binary variable (VA $\leq$ LP and VA $\geq$ LP). The data excluded from regression analysis were used as a test set to validate the results because all the patients had admission facial CT scans. Contingency tables were used to obtain sensitivities, specificities, positive predictive values (PPVs), and negative predictive values. The $k$ statistic was used to test interobserver reliability in assessing CT variables by the radiologists.

### RESULTS

Of the 97 globe injuries that constituted the study group, IOP could be measured in 29 globe injuries, only 41 eyes could be evaluated for the presence of RAPD, and preoperative VA could be measured in 59 eyes. Patients who had IOP, RAPD, and presenting VA measured on the initial examination were compared with those who did not. Patients with measured IOP had a mean logMAR of $-3.38$ (difference $=0.2$, $P =.5$), the RAPD evaluated had a mean logMAR of $-3.1$ (difference $=0.68$, $P =.06$), and those with presenting VA had a logMAR of $-3.1$ (difference $=1.15$, $P =.0002$). The data indicate that patients with poor outcomes were less likely to have had a complete ophthalmologic examination, mainly a VA test.

### Univariate Analysis

Univariate analysis showed that penetrating injuries with stabs/sharp objects predicted a favorable VA at 1 month ($P = .03$), while penetrating injuries caused by gunshot wounds predicted unfavorable VA ($P < .0001$), with no significant difference in outcome between penetrating and perforating injuries ($P = .4$) (On-line Table 2).

The ophthalmologic examination data associated with poor VA include higher IOP (Spearman $\rho = -0.48$, $P = .008$), the presence of RAPD ($P = .008$), and poor presenting VA (Spearman $\rho = 0.84$, $P < .0001$) (Table 1 and On-line Table 2).

The CT findings that predicted poor 1-month VA included the presence of posterior segment hemorrhage ($P < .0001$) (Fig 3), fractional decrease in globe volume ($P = .007$) (Fig 2), fractional increase in globe volume ($P = .026$) (Fig 3), hyphema ($P < .0001$), intraorbital hemorrhage ($P = .0001$), extraocular hemorrhage ($P < .0001$), intraconal hemorrhage ($P < .0001$), intraconal emphysema ($P < .0001$), extraocular emphysema ($P < .0001$), extraconal emphysema ($P < .0001$), intraocular gas ($P < .0001$), lens subluxation ($P = .0018$), orbital fractures ($P < .0001$), complex facial fractures ($P < .0001$), and a shallow anterior chamber ($P = .04$) (Fig 3). A visible lens ($P = .0054$) on CT predicted a favorable VA at 1 month (Table 1 and On-line Table 2).

There was fair-to-very good agreement among the reviewers (On-line Table 2), except for traumatic cataract and lens subluxation, which had poor agreement.

### Multivariate Analysis

Statistically significant variables such as IOP, RAPD, lens subluxation, and shallow anterior chamber were excluded from multivariate analysis as part of a list-wise deletion and due to their infrequent availability; this elimination resulted in a model with less overfitting and greater generalizability. Forty-nine globe injuries were selected to derive the regression model. The remaining 36 globe injuries, which were excluded from regression analysis due lack of adequate clinical data, were retained as test samples to validate the value of the strongest CT variable. The regression model identified 4 predictor variables with significance (Table 2): posterior segment hemorrhage, presenting VA, orbital emphysema, and complex facial fractures. The 4 variables resulted in an adjusted $R^2$ of 0.827 and $P < .0001$ (root mean square error $= 0.71$). The variance inflation factor was <4 for the variables. On internal k-fold cross-validation with 10-fold analysis, the results were consistent in all the folds and the average root mean square error was 0.75 (standard error, 0.02).

Receiver operating characteristic analysis of posterior segment hemorrhage in predicting profound vision loss yielded an area under the curve value of 0.97 in the derivation sample of 49 globes. The result was further validated on the test sample of 36 globes, which yielded a comparable area under the curve of 0.98. The addition of presenting VA to the model already containing posterior segment hemorrhage increased the area under the curve from 0.97 to 0.98 in the derivation sample.

The sensitivities, specificities, positive predictive values, and negative predictive values of various grades of posterior segment hemorrhage are given in Table 2.
hemorrhage in predicting profound vision loss (LP or NLP) are shown in Table 3. The absence of posterior segment hemorrhage predicted a 1-month VA better than LP (mild-to-severe vision loss) with sensitivity, specificity, PPV, and negative predictive values of 64%, 98%, 93%, and 89%, respectively. The results indicate that grade III posterior segment hemorrhage has a strong PPV for profound vision loss (LP or NLP) at 1-month follow-up, with all 49 injured globes with grade III hemorrhage resulting in profound vision loss. On the other hand, the absence of posterior segment hemorrhage has a strong PPV for VA better than LP (mild-to-severe vision loss), with 14 of 15 patients without hemorrhage having a VA of better than LP on follow-up.

**DISCUSSION**

In this study, we systematically analyzed the CT findings of open globe injuries, intraorbital soft-tissue injuries, orbitofacial skeletal injuries, and limited clinical data that are possible to obtain in the acute trauma setting to determine the predictors of VA at 1 month. According to our results, posterior segment hemorrhage, presenting VA, orbital emphysema, and complex facial fractures are the independent predictors of poor VA. The developed model has a good predictive performance, and the strengths of the model are the following: 1) Most of the predictors can be assessed on admission CT, 2) it requires a minimum amount of clinical data, and 3) the prediction model does not depend on surgical findings or postsurgical outcomes. Posterior segment hemorrhage was found to be the strongest predictor of VA at 1 month. Moreover, grade III posterior segment hemorrhage had a strong PPV (100%) for profound vision loss (LP or NLP) at 1 month, while absence of hemorrhage had a strong PPV (93%) for mild-to-severe vision loss (VA better than LP).

The Ocular Trauma Score model predicted profound vision loss with a sensitivity of 100% and specificity of 91%, respectively. On the other hand, Classification and Regression Tree analysis has shown a sensitivity of 82% and specificity of 86%. Our results showed that grade III posterior segment hemorrhage, which can be easily derived from facial CT, alone has a sensitivity of 78% and specificity of 100% in predicting profound vision loss. The absence of posterior segment hemorrhage, on the other hand, had a sensitivity of 64% and a specificity of 98% in predicting mild-to-severe vision loss. The best predictor from Classification and Regression Tree analysis was RAPD. In contrast, RAPD was not incorporated into the logistic regression model due to listwise deletion, though RAPD was significant on univariate analysis. The other disadvantage of depending on RAPD as a predictor was that only 42% of our injuries could be evaluated for the presence of RAPD in the acute trauma setting.

Joseph et al have reported that nearly 90% of globes with moderate-to-severe deformity had poor visual outcomes and 64% underwent enucleation, while no globes without scleral deformity underwent enucleation. Yuan et al showed that scleral deformities are associated with a decompressed globe in 36% of their studied globe ruptures and 15% showed enlarged globes. Decompression of the globe results from vitreous or uveal prolapse through the defect. To increase the quantitative precision of the degree of scleral deformity, we measured the globe volumes and calculated the fractional decrease or fractional increase in the volume of the ruptured globe. Our results showed that neither fractional decrease nor fractional increase in the globe volume was significant on regression analysis, though both variables were significant on univariate analysis. Margo et al, in a recent study, showed a correlation between high IOP and poor visual prognosis. Increased IOP, which had collinearity with posterior segment hemorrhage (Spearman p = 0.015), was associated with poor visual outcome. Joseph et al showed that 60% of patients with vitreous hemorrhage end up with enucleation and 75% of those with blood occupying >50% of the vitreous chamber end up with enucleation. Although our study did not evaluate the rates of enucleation, the results confirmed high sensitivity (98%; 95% CI, 91–100) of posterior segment hemorrhage in predicting profound vision loss.

**Limitations**

Our study has several limitations. First, it is a retrospective design with inherent biases. Second, some significant CT variables may have been excluded from analysis because CT variable selection was performed on the basis of the previously identified risk factors and some of the variables were selected by the study team members on the basis of their clinical experience. Simultaneous review of all signs may have resulted in some degree of bias in determining their individual performance. Finally, a direct comparison of our model with either the Ocular Trauma Score or the Classification and Regression Tree model was not performed because we did not record the surgical data that were the major components of those models, in keeping with the purpose of our study (ie, to derive a predictive model before surgery).

**CONCLUSIONS**

Previously known prognostic prediction models after open globe injury depend on data derived from ophthalmologic examination and surgical findings. However, ophthalmologic examination or even testing for RAPD, the best predictor in the Classification and Regression Tree model, may not be possible or reliable in some patients as seen from our results, due to nonreactive pupils frequently encountered in patients with associated traumatic brain injury, coma, or raised intracranial pressures. In this context, our
results may have clinical implications when radiologists, with the help of CT and limited preoperative clinical data that is realistically possible to obtain in acute trauma settings, can predict VA after open globe injury. Such information reduces anxiety in patients and helps informed decision-making regarding treatment choices.


REFERENCES

Magnetic Resonance Imaging (MRI)-Based Texture Analysis to Differentiate Sinonasal Squamous Cell Carcinoma from Inverted Papilloma


ABSTRACT

BACKGROUND AND PURPOSE: Because sinonasal inverted papilloma can harbor squamous cell carcinoma, differentiating these tumors is relevant. The objectives of this study were to determine whether MRI imaging–based texture analysis can accurately classify cases of noncoexistent squamous cell carcinoma and inverted papilloma and to compare this classification performance with neuroradiologists’ review.

MATERIALS AND METHODS: Adult patients who had inverted papilloma or squamous cell carcinoma resected were eligible (coexistent inverted papilloma and squamous cell carcinoma were excluded). Inclusion required tumor size of >1.5 cm and preoperative MR imaging with axial T1, axial T2, and axial T1 postcontrast sequences. Five well-established texture analysis algorithms were applied to an ROI from the largest tumor cross-section. For a training dataset, machine-learning algorithms were used to identify the most accurate model, and performance was also evaluated in a validation dataset. On the basis of 3 separate blinded reviews of the ROI, isolated tumor, and entire images, 2 neuroradiologists predicted tumor type in consensus.

RESULTS: The inverted papilloma (n = 24) and squamous cell carcinoma (n = 22) cohorts were matched for age and sex, while squamous cell carcinoma tumor volume was larger (P = .001). The best classification model achieved similar accuracies for training (17 squamous cell carcinomas, 16 inverted papillomas) and validation (7 squamous cell carcinomas, 6 inverted papillomas) datasets of 90.9% and 84.6%, respectively (P = .537). For the combined training and validation cohorts, the machine-learning accuracy (89.1%) was better than that of the neuroradiologists’ ROI review (56.5%, P = .0004) but not significantly different from the neuroradiologists’ review of the tumors (73.9%, P = .060) or entire images (87.0%, P = .748).

CONCLUSIONS: MR imaging–based texture analysis has the potential to differentiate squamous cell carcinoma from inverted papilloma and may, in the future, provide incremental information to the neuroradiologist.

ABBREVIATIONS: DOST = Discrete Orthonormal Stockwell Transform; GFB = Gabor Filter Banks; GLCM = Gray-Level Co-occurrence Matrix; IP = inverted papilloma; LBP = local binary patterns; LoGHist = Laplacian of Gaussian histogram; PC = principal component; SCC = squamous cell carcinoma; T1 = axial T1-weighted MRI pulse sequence; T2 = axial T2-weighted MRI pulse sequence with frequency-selective fat suppression; TIC = axial T1-weighted postcontrast MRI pulse sequence with frequency-selective fat suppression

Inverted papilloma (IP) is an uncommon sinonasal tumor of ectodermal origin that most commonly arises from the lateral nasal wall. In addition to its pattern of locally aggressive behavior and a propensity for postoperative recurrence, there is an association with malignancy, mostly squamous cell carcinoma (SCC). Reports vary widely in frequency, but the rate of carcinoma is on the order of 10%–15%, and approximately 60%–70% of these are synchronous. Although office-based endoscopic incisional biopsy is safe, the sensitivity for the diagnosis of malignancy has been called into question due to sampling errors.

It can be useful to preoperatively identify SCC when coexistent with IP to guide biopsy, expedite surgery, and plan an oncologically sound resection. Although bone thinning and remodeling without large areas of erosion on CT are more characteristic of IP than SCC, this finding is imperfect because IP may also aggressively destroy...
bone, particularly when contacting the walls of the sphenoid sinuses and floor of the anterior cranial fossa. Not surprising, FDG-PET has shown a higher mean standard uptake value for SCC compared with IP, but overlap limits the clinical utility of PET. To date, MR imaging has the most promise in differentiating SCC from IP, although early work initially found no signature appearance on MR imaging for IP, alternating hypointense and hyperintense bands on T2-weighted and contrast-enhanced T1-weighted sequences have become recognized as a distinctive feature of IP and have been described as a convoluted cerebriform pattern. As with any subjective finding, interpretive errors can occur, especially when tumors are small or incompletely express the convoluted cerebriform pattern. Hence, a more objective form of image analysis is desirable.

Texture analysis is a form of image processing that seeks to characterize complex visual patterns by quantitatively identifying simpler but characteristic subpatterns. Within the field of head and neck radiology, texture analysis has shown applicability in predicting the p53 status of SCC, classifying SCC as human papilloma virus–related, predicting treatment response in head and neck cancer, differentiating benign from malignant thyroid nodules, and characterizing parotid tumors and structural changes after radiation therapy. Because a large volume of data is generated when multiple texture analysis algorithms are applied to MR imaging sequences, the statistical comparison of individual texture features is of limited practical value; instead, a multifactorial data-driven analysis is necessary (ie, “radiomics”). Therefore, the objective of this study was to determine whether MR imaging–based texture analysis can differentiate sinonasal SCC from IP by using a multiparametric machine-learning model. Model performance was additionally compared against qualitative neuroradiologists’ interpretation to determine its potential for added clinical value.

MATERIALS AND METHODS

Subject Enrollment

The Mayo Clinic Institutional Review Board approved this retrospective study, and the need for informed consent was waived. The pathology data base was queried to identify adult patients (18 years of age or older) who underwent resection of sinonasal IP or SCC. Subjects enrolled from January 1, 2009, to December 31, 2014, were included in the training dataset for model development, while those enrolled between January 1, 2015, and July 1, 2016, composed the validation dataset. To ensure that only a single histologic tumor type would be used for texture analysis, we excluded cases of coexistent IP and SCC. Potential subjects were screened to determine which of them had preoperative face MR imaging available for review. The MRIs, which were performed on numerous scanners within the authors’ institution and at external facilities, had to be of diagnostic image quality. At a minimum, the imaging had to include an axial T1-weighted MRI pulse sequence (T1), an axial T2-weighted pulse sequence with frequency-selective fat-suppression (T2), and an axial T1-weighted postcontrast MRI pulse sequence with frequency-selective fat suppression (T1C) for texture analysis.

Image Preparation and Texture Analysis

DICOM files containing the T1, T2, and T1C pulse sequences (also referred to as “contrasts” for the purpose of texture analysis) were anonymized and encoded so that all subsequent image analysis was blinded. To ensure uniformity for texture analysis, we performed resampling and/or zero-padding to generate images with an 18-cm FOV and a 256 × 256 pixel array and normalized image intensities to a dynamic range of 0–255. The studies were then reviewed by a board-certified neuroradiologist with OsiriX (Version 6.5; http://www.osirix-viewer.com). The borders of the tumor were manually traced on all T1C images on which tumor was visible to generate an ROI-based cross-sectional area for each image and an estimated tumor size by using the ROI Volume function in OsiriX. On the axial image with the greatest tumor cross-sectional area, the neuroradiologist inserted the largest possible rectangular ROI that would fit within the tumor for all 3 sequences (Fig 1). To prevent 2D texture analysis from being biased by tumor size, a computer script determined the maximal square ROI that could fit within all manually drawn rectangular ROIs across all subjects and automatically positioned this smallest common square ROI at the isocenter of each of the rectangular ROIs. The contents of this square ROI, with 16 × 16 pixels, served as the input for texture analysis.

Texture analysis of each ROI consisted of 3 first-order intensity-based features (mean, SD, and range of gray-level intensities) and features computed by using 5 widely available texture algorithms (all implemented in Python 2.7 programming language [https://www.python.org/downloads/], by using either custom-written code based on publications or open-source libraries as noted):

1) Gray-Level Co-occurrence Matrix (GLCM) is a widely applied method that uses second-order statistics to assess the ar-

FIG 1. ROI placement. A 51-year-old man with an IP involving the right maxillary sinus. Axial T2-weighted fat-suppressed MR imaging pulse sequence demonstrates the manual placement of the largest rectangular ROI that would fit within the tumor margins on the axial image with the greatest tumor cross-sectional area. The inset image in the lower right corner is representative of the final 16 × 16 matrix that was derived from the ROI isocenter and served as the input for texture analysis.
rangement of similar gray-scale intensities within an ROI. \(^\text{20}\) GLCM evaluates how frequently a pair of intensity levels is identified in an orientation based on a specified angle and radius. In the current study, the co-occurrence matrix was determined for a distance of 1 pixel over 4 angular directions (0°, 45°, 90°, and 135°). The mean and range for 13 rotationally invariant features (including measures of homogeneity, entropy, angular second moment, correlation, and dissimilarity) were computed at each ROI for each MR imaging contrast. \(^\text{20}\)

2) Local binary patterns (LBP) evaluates the set of points within a fixed radius of a specified voxel to determine in a binary fashion whether they are higher or lower in intensity than neighboring voxels. \(^\text{21}\) Depending on the number of bitwise transitions across this interrogated region, the LBP can be classified as uniform or nonuniform, and histograms of these data provide a measure of ROI uniformity. A 3-voxel radius was selected to complement the smaller scale patterns already assessed by GLCM. A 12-bin histogram was used, resulting in 12 LBP texture features being calculated at each ROI for each MR imaging contrast. \(^\text{20}\)

3) Discrete Orthonormal Stockwell Transform (DOST) provides a rotationally invariant multiresolution spatial-frequency representation of an image based on dyadic sampling of the Fourier representation of the image. \(^\text{22}\) Ten DOST features were calculated at each ROI for each MR imaging contrast.

4) Laplacian of Gaussian Histogram (LoGHist) is a convolution-based method to capture the spectral composition of an image in intermediate scales not achievable with first- and second-order statistics. Through the use of varying sizes of bandpass filters, different scales of texture ranging from fine to coarse are highlighted. \(^\text{23}\) Gaussians with 3 different values of \(\sigma\) (2.0, 4.0, and 6.0) were used to cover the range of fine-to-medium-scale textures, and 18 LoGHist features were generated at each ROI for each MR imaging contrast.

5) The Gabor Filter Banks (GFB) technique uses localized and linear filters to capture details in various frequency resolutions. \(^\text{24}\) Four different Gabor filters were rendered by using 2 \(\sigma\) levels (1.0 and 3.0) and 2 frequency levels (0.6 and 1.0). By calculating the mean and SD of the filtered ROI, we computed 8 GFB features at each ROI for each MR imaging contrast.

**Neuroradiologists’ Review**

Using OsiriX, 2 neuroradiologists with 25 and 28 years of experience, respectively, performed a blinded review to reach a consensus diagnosis of IP or SCC for each case. This was performed during 3 separate rounds of image review, each of which was randomized and completed in the following order:

1) ROI: For the T1, T2, and T1C series, the neuroradiologists exclusively reviewed the 16 \(\times\) 16 square ROIs that had been used for texture analysis.

2) Tumor: On all images in the T1, T2, and T1C series, the data outside the tumor margins were zero-filled so that the neuroradiologists could only base their assessment on the intrinsic appearance of the tumor without information regarding tumor location and invasive behavior.

3) Image: The neuroradiologists were able to review the unaltered T1, T2, and T1C imaging datasets in their entirety.

**Machine Learning and Statistical Analysis**

Open-source R statistical and computing software (http://www.r-project.org) was used to perform the analyses and classification. Hypothesis tests were 2-sided, and statistical significance was defined as \(P<.05\). The comparison of subject demographics and tumor size between IP and SCC was performed by using a 2-sample \(t\) test for subject age and tumor volume and a Fisher exact test for sex. The 2-sample \(t\) test was used for a univariate comparison of texture features between IP and SCC before the application of machine-learning methodology, and \(P\) values were corrected for multiple comparisons by using the false discovery rate. \(^\text{25}\)

A total of 231 texture features were calculated for each case (77 texture features per MR imaging contrast \(\times\) 3 contrasts). To reduce the dimensionality of the texture features and increase the generalizability of the predictive model for the training dataset, we used principal component (PC) analysis. \(^\text{26-28}\) PCs, which are linear combinations of features, were identified separately for each texture algorithm and MR imaging contrast. Those PCs that sufficiently accounted for 90% of the texture feature variability were selected for further processing. Three commonly described classification algorithms, Diagonal Linear Discriminate Analysis, Support Vector Machines, and Diagonal Quadratic Discriminate Analysis, were conducted on the basis of the selected PCs in an attempt to differentiate SCC from IP. \(^\text{29-31}\) Sequential forward-feature selection identified the image-based PCs that yielded the greatest accuracy. \(^\text{26-27}\) In developing the classification model, we initially selected the PC with the largest discriminatory power and incorporated additional PCs that improved model accuracy in an iterative fashion until incremental gains in accuracy were \(<1\%\).

Classification accuracy was determined by using leave-one-out cross-validation, in which all samples except for 1 were used, while the left-out sample served as the test case with which to assess classification accuracy. \(^\text{32}\) This process was repeated until all samples in the training dataset had served as the test case, and the overall cross-validation accuracy was the averaged accuracy. The most accurate classification model was applied in a blinded fashion to the validation dataset, and the diagnostic performance of the model was assessed. Model performance accuracies between
the training and validation datasets and between the best classification model and neuroradiologists’ review were compared by using a 2-tailed test of population proportion.

**RESULTS**

Thirty-three subjects were included in the training set, 16 IPs and 17 SCCs, while the validation set consisted of 6 IPs and 7 SCCs (Table 1). The patients were similarly matched for age and male-predominant sex. Mean tumor volume was larger for SCC than IP ($P = .001$).

Individual features derived from the 5 different texture analyses across all 3 MR imaging contrasts (T1, T2, T1C) were initially evaluated in a univariate fashion to look for significant differences between the IP and SCC groups (Fig 2). The greatest number of texture features showing statistically significant differences were derived from the DOST and GFB texture analyses.

Model performance for the training and validation datasets is presented in Table 2. Following PC analysis and machine-learning classification, the predictive classifier with the best classification result was Support Vector Machines, yielding 90.9% accuracy for the training dataset. The 84.6% accuracy of the validation dataset did not significantly differ from that achieved in the training dataset ($P = .537$). When we combined the training and validation cohorts ($n = 46$), the accuracy achieved by texture analysis (89.1%) was significantly better than that of the ROI-based neuroradiologists’ review (Table 3, 56.5%, $P = .004$) and showed a trend toward improved accuracy over neuroradiologists’ review of the entire tumor (73.9%, $P = .060$). Texture-analysis accuracy was not significantly different from that of the neuroradiologists reviewing the entire unaltered images (87.0%, $P = .748$).

**FIG 2.** Heat map showing MR imaging texture feature significance in distinguishing tumor type. Univariate analysis compared the pathology status (SCC versus IP) with MR imaging–texture features. Color maps show the false discovery rate–adjusted $P$ values of a 2-sample $t$-test. MR imaging contrasts (pulse sequences) are listed above the columns, and MR imaging–based texture features are listed in rows. DOST features 0–9 correspond with low-to-high frequency patterns. LBP 0–11 are the normalized bin counts in the LBP histogram. The reader is referred to the “Materials and Methods” section for additional details about the features.

**Table 2. Diagnostic performance of machine-learning classification in training and validation datasets**

<table>
<thead>
<tr>
<th>Tumor Type (Pathologic Diagnosis)</th>
<th>SCC</th>
<th>IP</th>
<th>Diagnostic Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model prediction for training dataset</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCC</td>
<td>16</td>
<td>2</td>
<td>Accuracy 90.9% $^*$</td>
</tr>
<tr>
<td>IP</td>
<td>1</td>
<td>14</td>
<td>Specificity 87.5%</td>
</tr>
<tr>
<td>Total</td>
<td>17</td>
<td>16</td>
<td>NPV 93.3%</td>
</tr>
<tr>
<td>Model prediction for validation dataset</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCC</td>
<td>6</td>
<td>1</td>
<td>Accuracy 84.6% $^*$</td>
</tr>
<tr>
<td>IP</td>
<td>1</td>
<td>5</td>
<td>Specificity 83.3%</td>
</tr>
<tr>
<td>Total</td>
<td>7</td>
<td>6</td>
<td>NPV 83.3%</td>
</tr>
<tr>
<td>Model prediction for entire cohort</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCC</td>
<td>22</td>
<td>3</td>
<td>Accuracy 89.1%</td>
</tr>
<tr>
<td>IP</td>
<td>2</td>
<td>19</td>
<td>Specificity 86.4%</td>
</tr>
<tr>
<td>Total</td>
<td>24</td>
<td>22</td>
<td>NPV 90.5%</td>
</tr>
</tbody>
</table>

Note: —NPV indicates negative predictive value; PPV, positive predictive value.

* With a 2-tailed test of population proportion, the accuracies for the training and validation datasets were not significantly different ($P = .537$).
The prior work focused on a qualitative imaging appearance technique with which to distinguish sinonasal SCC from IP. Most of MR imaging has long been recognized as the most useful technique, but this pattern has a high level of sensitivity for IP, it is not entirely specific. As an example, Jeon et al\(^9\) evaluated the performance of the convoluted cerebriform pattern on T2-weighted images for tumors of \(<2\) cm. While tumors smaller than \(1.5 \times 1.5\) cm were excluded from enrollment in the current study, the final processed ROIs were only \(1.125 \times 1.125\) cm.

Because small noninvasive sinonasal tumors are not universally imaged with MR imaging, the greatest potential benefit for texture analysis might be in detecting a small focus of SCC within a larger IP to expedite patient management. Accurately assessing small regions would be a prerequisite for the detection of such tumor heterogeneity. The potential for interpretive error is greatest when a small focus of SCC exists within a much larger IP and goes unrecognized because a convoluted cerebriform pattern is still present. Indeed, this pattern of a “partial” convoluted cerebriform pattern has been described.\(^6,9,35\) Likewise, necrosis, recognized as non-enhancing tissue on contrast-enhanced MR imaging, is associated with SCC but may not be apparent when a small focus of SCC coexists with an IP.\(^9,10,35\) A future goal for texture analysis of a mixed tumor containing both IP and SCC is to assist with interpretation by highlighting areas that are most suspicious for SCC.

Texture analysis integrated into a machine-learning model was able to classify SCC and IP with an accuracy on par with the previously published results based on the convoluted cerebriform pattern.\(^9\) It is also similar to the best consensus neuroradiologists’ interpretation in the current study. However, this technology is meant to supplement a neuroradiologist’s interpretive skills rather than compete with them. In clinical practice, a diagnosis is rendered by synthesizing all available data that include not only intrinsic tumor appearance but also other imaging features such as site of origin, tumor size, extrasinonasal extension, and tumor margins. Indeed, the current results support a neuroradiologist’s accuracy improving for differentiating sinonasal IP and SCC as more imaging information is made available. On the basis of a \(16 \times 16\) ROI, the texture-based machine-learning model outperformed the accuracy of the neuroradiologists \((P = .0004)\).

In terms of assessing the intrinsic tumor appearance, texture analysis stands to provide incremental benefit when human pattern recognition becomes most limited, and this can occur with a small tumor. For example, Maroldi et al\(^6\) found it more challenging to recognize the convoluted cerebriform pattern on T2-weighted images for tumors of \(<2\) cm. While tumors smaller than \(1.5 \times 1.5\) cm were excluded from enrollment in the current study, the final processed ROIs were only \(1.125 \times 1.125\) cm.

Table 3: Diagnostic performance of texture analysis with machine learning compared with neuroradiologists’ review for the differentiation of SCC from IP\(^a\)

<table>
<thead>
<tr>
<th>Analysis Method</th>
<th>Accuracy(^b)</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Texture analysis with machine learning</td>
<td>89.1%</td>
<td>91.7%</td>
<td>86.4%</td>
<td>88.0%</td>
<td>90.5%</td>
</tr>
<tr>
<td>Neuroradiologists’ review, ROI</td>
<td>56.5% ((P = .0004))</td>
<td>54.2%</td>
<td>59.1%</td>
<td>59.1%</td>
<td>54.2%</td>
</tr>
<tr>
<td>Neuroradiologists’ review, tumor</td>
<td>73.9% ((P = .060))</td>
<td>75.0%</td>
<td>72.7%</td>
<td>75.0%</td>
<td>72.7%</td>
</tr>
<tr>
<td>Neuroradiologists’ review, image</td>
<td>87.0% ((P = .748))</td>
<td>91.7%</td>
<td>81.8%</td>
<td>84.6%</td>
<td>90.0%</td>
</tr>
</tbody>
</table>

Note: NPV indicates negative predictive value; PPV, positive predictive value.

\(^a\) Results are shown for the entire cohort (22 IPs, 24 SCCs) and reflect the best classification model. The labels for the neuroradiologists’ assessment indicate whether they reviewed the \(16 \times 16\) ROI (ROI), tumor alone (tumor), or entire images (image).

\(^b\) \(P\) values represent comparison of texture analysis with machine learning against each neuroradiologist’s review using a 2-tailed test of population proportion.

**DISCUSSION**

MR imaging has long been recognized as the most useful technique with which to distinguish sinonasal SCC from IP. Most of the prior work focused on a qualitative imaging appearance known as the “convoluted cerebriform pattern.”\(^6,9,11,33,34\) Although this pattern has a high level of sensitivity for IP, it is not entirely specific. As an example, Jeon et al\(^9\) evaluated the performance of the convoluted cerebriform pattern in 30 patients with IP relative to 128 patients with sinonasal malignancies and reported a sensitivity of 100%, specificity of 87%, positive predictive value of 64%, negative predictive value of 100%, and accuracy of 89%.

Relative contributions to model accuracy from each texture analysis algorithm and MR imaging contrast are presented (Fig 3). The most significant texture features were derived from T1C-GFB, T1-GLCM, and T1-DOST (Fig 4).

![Contributions of Selected PCs](chart1.png)

**FIG 3.** Relative contributions to model accuracy. Of the 90.9% overall model accuracy for the training dataset, the bar graph demonstrates the accuracy attributable to PCs derived from T1C-GFB, T1-DOST, and T1-GLCM (upper panel). Across all texture algorithms, the contribution to total model accuracy was derived predominantly from T1C, with minor contributions from T1 and no input from T2 (lower panel).

![Contributions of Selected Contrasts](chart2.png)

**DISCUSSION**

MR imaging has long been recognized as the most useful technique with which to distinguish sinonasal SCC from IP. Most of the prior work focused on a qualitative imaging appearance known as the “convoluted cerebriform pattern.”\(^6,9,11,33,34\) Although this pattern has a high level of sensitivity for IP, it is not entirely specific. As an example, Jeon et al\(^9\) evaluated the performance of the convoluted cerebriform pattern in 30 patients with IP relative to 128 patients with sinonasal malignancies and reported a sensitivity of 100%, specificity of 87%, positive predictive value of 64%, negative predictive value of 100%, and accuracy of 89%.

Texture analysis integrated into a machine-learning model was able to classify SCC and IP with an accuracy on par with the previously published results based on the convoluted cerebriform pattern.\(^9\) It is also similar to the best consensus neuroradiologists’ interpretation in the current study. However, this technology is meant to supplement a neuroradiologist’s interpretive skills rather than compete with them. In clinical practice, a diagnosis is rendered by synthesizing all available data that include not only intrinsic tumor appearance but also other imaging features such as site of origin, tumor size, extrasinonasal extension, and tumor margins. Indeed, the current results support a neuroradiologist’s accuracy improving for differentiating sinonasal IP and SCC as more imaging information is made available. On the basis of a \(16 \times 16\) ROI, the texture-based machine-learning model outperformed the accuracy of the neuroradiologists \((P = .0004)\).

In terms of assessing the intrinsic tumor appearance, texture analysis stands to provide incremental benefit when human pattern recognition becomes most limited, and this can occur with a small tumor. For example, Maroldi et al\(^6\) found it more challenging to recognize the convoluted cerebriform pattern on T2-weighted images for tumors of \(<2\) cm. While tumors smaller than \(1.5 \times 1.5\) cm were excluded from enrollment in the current study, the final processed ROIs were only \(1.125 \times 1.125\) cm.

Because small noninvasive sinonasal tumors are not universally imaged with MR imaging, the greatest potential benefit for texture analysis might be in detecting a small focus of SCC within a larger IP to expedite patient management. Accurately assessing small regions would be a prerequisite for the detection of such tumor heterogeneity. The potential for interpretive error is greatest when a small focus of SCC exists within a much larger IP and goes unrecognized because a convoluted cerebriform pattern is still present. Indeed, this pattern of a “partial” convoluted cerebriform pattern has been described.\(^6,9,35\) Likewise, necrosis, recognized as non-enhancing tissue on contrast-enhanced MR imaging, is associated with SCC but may not be apparent when a small focus of SCC coexists with an IP.\(^9,10,35\) A future goal for texture analysis of a mixed tumor containing both IP and SCC is to assist with interpretation by highlighting areas that are most suspicious for SCC.

Texture analysis can also extract useful features from images that have been traditionally neglected by the human eye. The convoluted cerebriform pattern has been historically described on T1-weighted postcontrast and T2-weighted sequences.\(^6,9,11,33,35\) However, noncontrast T1-weighted MR imaging has received no attention to date, to our knowledge. The texture analysis in the current study found more significant features for T1 than T2 on a univariate basis (Fig 2). Although T1-DOST and T1-GLCM made a minority contribution to the final model, no T2 features contributed to final model accuracy (Fig 3).
protocols and scanner platforms. Certainly, this possibility would make such results clinically meaningless because a new model would have to be created for each scanner running a unique protocol. Fruehwald-Pallamar et al concluded that texture analysis is not practical for differentiating malignant and benign tumors of the head and neck when using different protocols on different MR imaging scanners. However, their cohort was very heterogeneous, containing numerous types of benign and malignant lesions. The subjects for the current study were accrued during a long period and were not imaged with a common scanner and protocol. Nevertheless, an accurate texture-based model was achieved that performed similarly in the training and validation datasets. At least for the context of sinonasal IP and SCC, this outcome holds promise for reproducibility across scanner platforms.

The current study is limited, given its retrospective nature and small sample size. Hence, the high accuracy for the differentiation of SCC from IP with texture analysis is not meant to represent the performance of an established diagnostic imaging test. Instead, these results merely confirm the feasibility of this technique for distinguishing these 2 tumor types. In showing proof of concept, a 2D ROI-based analysis was used to confirm discriminatory ability with a limited data sample. Moreover, because SCC tends to be a larger tumor than IP on average, this approach eliminated the potential for falsely finding texture differences on the basis of relative oversampling of a larger tumor. Future directions will include the refinement of the texture analysis pipeline into a volumetric tool with the objective of highlighting foci of SCC when it is coexistent with IP. This will need to be studied prospectively to ensure that the histopathologic analysis can be accurately coregistered to MR imaging.

**CONCLUSIONS**

With MR imaging–based texture analysis, a machine-learning model for the differentiation of sinonasal SCC and IP achieved accuracy comparable with both neuroradiologists’ interpretation and previously published reports on the convoluted cerebriform pattern. Because the classification model was significantly more accurate than the neuroradiologists’ interpretation for a small ROI, texture analysis has the potential to provide incremental benefit to the neuroradiologists’ interpretation, particularly in cases of small or heterogeneous tumors.

**Disclosures:** Shreya Ramkumar—RELATED: Grant: Arizona State University/Mayo Clinic Seed Grant, Comments: A $40,000 intramural research grant funding was awarded to foster collaborative research between Arizona State University and Mayo Clinic.* Sara Ranjbar—RELATED: Grant: Arizona State University/Mayo Clinic Seed Grant, Comments: A $40,000 intramural research grant funding was awarded to foster collaborative research between Arizona State University and Mayo Clinic.* Devyani Lal—UNRELATED: Board Membership: American Rhinologic Society, Comments: not for profit; Employment: Mayo Clinic; Grants/Grants Pending: Flinn Foundation, Sanofi, Joseph Ross Mitchell—UNRELATED: Employment: Mayo Clinic; Grants/Grants Pending: National Institutes of Health*, Patients (Planned, Pending, or Issued): Mayo Clinic, Comments: Provisional patents were filed relating to texture analysis of medical imaging data.* Joseph Hoxworth—RELATED: Grant: National Institutes of Health; Comments: not for profit; Patent: Pending.

**REFERENCES**


Trochlear Groove and Trochlear Cistern: Useful Anatomic Landmarks for Identifying the Tentorial Segment of Cranial Nerve IV on MRI

P.M. Bunch, H.R. Kelly, D.A. Zander, and H.D. Curtin

ABSTRACT

BACKGROUND AND PURPOSE: The trochlear groove and trochlear cistern are anatomic landmarks closely associated with the tentorial segment of cranial nerve IV. The purposes of this study were to describe the MR imaging appearances of the trochlear groove and trochlear cistern and to test our hypothesis that knowledge of these anatomic landmarks facilitates identification of cranial nerve IV in routine clinical practice.

MATERIALS AND METHODS: For this retrospective study, consecutive MR imaging examinations of the sinuses performed in 25 patients (50 sides) at our institution were reviewed. Patient characteristics and study indications were recorded. Three readers performed independent assessments of trochlear groove, cistern, and nerve visibility on coronal images obtained by using a T2-weighted driven equilibrium radiofrequency reset pulse sequence.

RESULTS: Interobserver agreement was 78% for visibility of the trochlear groove, 56% for the trochlear cistern, and 68% for cranial nerve IV. Following consensus review, the trochlear groove was present in 44/50 sides (88%), the trochlear cistern was present in 25/50 sides (50%), and cranial nerve IV was identified in 36/50 sides (72%). When the trochlear groove was present, cranial nerve IV was identified in 35/44 sides (80%), in contrast to 1/6 sides (17%) with no groove (P = .0013). When the trochlear cistern was present, cranial nerve IV was identified in 23/25 sides (92%), in contrast to 13/25 sides (52%) with no cistern (P = .0016).

CONCLUSIONS: The trochlear groove and trochlear cistern are anatomic landmarks that facilitate identification of cranial nerve IV in routine clinical practice.

ABBREVIATION: coronal T2 DRIVE = coronal 2D T2-weighted driven equilibrium radiofrequency reset pulse

Historically, imaging of cranial nerve IV has been difficult.¹,² These difficulties likely relate to the small size of the nerve and the presence of numerous adjacent blood vessels in the quadrigeminal and ambient cisterns.³,⁴ These difficulties are compounded by the fact that MR images are often acquired in the axial plane, nearly parallel to the cisternal course of the nerve.

In research settings, high-resolution MR images have reliably demonstrated cranial nerve IV in healthy volunteers; however, long acquisition times ranging from 7 to 26 minutes make these sequences impractical for routine clinical use.³,⁵,⁶ Acknowledging the perceived impracticality of consistently identifying cranial nerve IV in routine practice, more recently published work has proposed using volumetric analyses of the superior oblique muscles as a correlate for the presence or absence of cranial nerve IV in patients with congenital superior oblique palsy.⁷ However, the ability to reliably identify the nerve itself can be important in identifying underlying pathology and can aid in treatment and surgical planning.³,⁸

Various segments and subsegments have been proposed to describe the course of cranial nerve IV distal to its exit from the dorsal midbrain and proximal to its entrance into the cavernous sinus.³,⁴,⁶,⁹,¹⁰ Some authors consider the tentorial segment to be the distal subdivision of the cisternal segment,³,⁶,⁹,¹⁰ and others describe the tentorial segment as distinct from the cisternal segment of cranial nerve IV.⁴,¹¹ However, there is generally agreement that the tentorial segment of cranial nerve IV begins when the nerve becomes closely associated with the inferior aspect of the cerebellar tentorium and ends at the entrance of the nerve to the cavernous sinus.
We hypothesized that knowledge of these anatomic landmarks will facilitate identification of the tentorial segment of cranial nerve IV in routine clinical practice.

The purposes of this study were to describe the MR imaging appearances of the trochlear groove and trochlear cistern and to test our hypothesis regarding the utility of these landmarks for identifying cranial nerve IV.

MATERIALS AND METHODS

Subjects

For this Health Insurance Portability and Accountability Act–compliant, institutional review board–approved study, all sinus MR imaging examinations performed at Massachusetts Eye and Ear between December 1, 2015, and April 1, 2016, were retrospectively reviewed. Inclusion criteria were the following: 1) older than 18 years of age, 2) coronal T2 DRIVE sequence performed, and 3) no superior oblique palsy documented in the patient’s electronic medical record. Twenty-five MR imaging examinations in 25 patients met the inclusion criteria and were included in the study. Patient characteristics and study indications were recorded.

To maximize generalizability and to best simulate routine clinical practice, we did not exclude examinations containing motion, pulsation, and other MR imaging artifacts from our study.

Image Acquisition

All examinations were performed on a 3T scanner (Achieva; Philips Healthcare, Best, the Netherlands) by using an 8-channel head coil (SENSE Flex M coil; Philips Healthcare).

The coronal T2 DRIVE sequence acquisition parameters for the study group were as follows: TR, 3369–6355 ms; TE, 100 ms; echo–train length, 11; FOV, 120–160 mm; matrix, 324 × 300 to 456 × 454; voxel size, from 0.3 × 0.35 × 3.0 to 0.4 × 0.44 × 3.0; gap, 0.3 (n = 12) or 1 mm (n = 13); NEX, 1 (n = 15) or 2 (n = 10). These acquisition parameters resulted in a mean sequence acquisition time of 3 minutes 54 seconds ± 37 seconds (range, 2 minutes 54 seconds–5 minutes).

Reader Assessment

Anatomic Structures. After a brief self-guided training session based on 10 sinus MR imaging examinations not included in the study group, 3 radiologists independently reviewed the coronal T2-DRIVE sequences for the 25 patients in the study group with the Massachusetts Eye and Ear PACS. The radiologists included 2 fellowship-trained neuroradiologists (6 and 3 years’ subspecialty experience) and 1 neuroradiology fellow. They were asked to determine whether the trochlear groove, trochlear cistern, and cranial nerve IV were identifiable on each side in all patients, and discrepancies were resolved by consensus review. To be considered “identifiable,” the tentorial segment of cranial nerve IV had to be identified by the reader.

The purposes of this study were to describe the MR imaging appearances of the trochlear groove and trochlear cistern and to test our hypothesis regarding the utility of these landmarks for identifying cranial nerve IV.

Materials and Methods

Subjects

For this Health Insurance Portability and Accountability Act–compliant, institutional review board–approved study, all sinus MR imaging examinations performed at Massachusetts Eye and Ear between December 1, 2015, and April 1, 2016, were retrospectively reviewed. Inclusion criteria were the following: 1) older than 18 years of age, 2) coronal T2 DRIVE sequence performed, and 3) no superior oblique palsy documented in the patient’s electronic medical record. Twenty-five MR imaging examinations in 25 patients met the inclusion criteria and were included in the study. Patient characteristics and study indications were recorded.

To maximize generalizability and to best simulate routine clinical practice, we did not exclude examinations containing motion, pulsation, and other MR imaging artifacts from our study.

Image Acquisition

All examinations were performed on a 3T scanner (Achieva; Philips Healthcare, Best, the Netherlands) by using an 8-channel head coil (SENSE Flex M coil; Philips Healthcare).

The coronal T2 DRIVE sequence acquisition parameters for the study group were as follows: TR, 3369–6355 ms; TE, 100 ms; echo–train length, 11; FOV, 120–160 mm; matrix, 324 × 300 to 456 × 454; voxel size, from 0.3 × 0.35 × 3.0 to 0.4 × 0.44 × 3.0; gap, 0.3 (n = 12) or 1 mm (n = 13); NEX, 1 (n = 15) or 2 (n = 10). These acquisition parameters resulted in a mean sequence acquisition time of 3 minutes 54 seconds ± 37 seconds (range, 2 minutes 54 seconds–5 minutes).

Reader Assessment

Anatomic Structures. After a brief self-guided training session based on 10 sinus MR imaging examinations not included in the study group, 3 radiologists independently reviewed the coronal T2-DRIVE sequences for the 25 patients in the study group with the Massachusetts Eye and Ear PACS. The radiologists included 2 fellowship-trained neuroradiologists (6 and 3 years’ subspecialty experience) and 1 neuroradiology fellow. They were asked to determine whether the trochlear groove, trochlear cistern, and cranial nerve IV were identifiable on each side in all patients, and discrepancies were resolved by consensus review. To be considered “identifiable,” the tentorial segment of cranial nerve IV had to be identified by the reader.

The purposes of this study were to describe the MR imaging appearances of the trochlear groove and trochlear cistern and to test our hypothesis regarding the utility of these landmarks for identifying cranial nerve IV.

Materials and Methods

Subjects

For this Health Insurance Portability and Accountability Act–compliant, institutional review board–approved study, all sinus MR imaging examinations performed at Massachusetts Eye and Ear between December 1, 2015, and April 1, 2016, were retrospectively reviewed. Inclusion criteria were the following: 1) older than 18 years of age, 2) coronal T2 DRIVE sequence performed, and 3) no superior oblique palsy documented in the patient’s electronic medical record. Twenty-five MR imaging examinations in 25 patients met the inclusion criteria and were included in the study. Patient characteristics and study indications were recorded.

To maximize generalizability and to best simulate routine clinical practice, we did not exclude examinations containing motion, pulsation, and other MR imaging artifacts from our study.

Image Acquisition

All examinations were performed on a 3T scanner (Achieva; Philips Healthcare, Best, the Netherlands) by using an 8-channel head coil (SENSE Flex M coil; Philips Healthcare).

The coronal T2 DRIVE sequence acquisition parameters for the study group were as follows: TR, 3369–6355 ms; TE, 100 ms; echo–train length, 11; FOV, 120–160 mm; matrix, 324 × 300 to 456 × 454; voxel size, from 0.3 × 0.35 × 3.0 to 0.4 × 0.44 × 3.0; gap, 0.3 (n = 12) or 1 mm (n = 13); NEX, 1 (n = 15) or 2 (n = 10). These acquisition parameters resulted in a mean sequence acquisition time of 3 minutes 54 seconds ± 37 seconds (range, 2 minutes 54 seconds–5 minutes).

Reader Assessment

Anatomic Structures. After a brief self-guided training session based on 10 sinus MR imaging examinations not included in the study group, 3 radiologists independently reviewed the coronal T2-DRIVE sequences for the 25 patients in the study group with the Massachusetts Eye and Ear PACS. The radiologists included 2 fellowship-trained neuroradiologists (6 and 3 years’ subspecialty experience) and 1 neuroradiology fellow. They were asked to determine whether the trochlear groove, trochlear cistern, and cranial nerve IV were identifiable on each side in all patients, and discrepancies were resolved by consensus review. To be considered “identifiable,” the tentorial segment of cranial nerve IV had to be identified by the reader.

The purposes of this study were to describe the MR imaging appearances of the trochlear groove and trochlear cistern and to test our hypothesis regarding the utility of these landmarks for identifying cranial nerve IV.

Materials and Methods

Subjects

For this Health Insurance Portability and Accountability Act–compliant, institutional review board–approved study, all sinus MR imaging examinations performed at Massachusetts Eye and Ear between December 1, 2015, and April 1, 2016, were retrospectively reviewed. Inclusion criteria were the following: 1) older than 18 years of age, 2) coronal T2 DRIVE sequence performed, and 3) no superior oblique palsy documented in the patient’s electronic medical record. Twenty-five MR imaging examinations in 25 patients met the inclusion criteria and were included in the study. Patient characteristics and study indications were recorded.

To maximize generalizability and to best simulate routine clinical practice, we did not exclude examinations containing motion, pulsation, and other MR imaging artifacts from our study.

Image Acquisition

All examinations were performed on a 3T scanner (Achieva; Philips Healthcare, Best, the Netherlands) by using an 8-channel head coil (SENSE Flex M coil; Philips Healthcare).

The coronal T2 DRIVE sequence acquisition parameters for the study group were as follows: TR, 3369–6355 ms; TE, 100 ms; echo–train length, 11; FOV, 120–160 mm; matrix, 324 × 300 to 456 × 454; voxel size, from 0.3 × 0.35 × 3.0 to 0.4 × 0.44 × 3.0; gap, 0.3 (n = 12) or 1 mm (n = 13); NEX, 1 (n = 15) or 2 (n = 10). These acquisition parameters resulted in a mean sequence acquisition time of 3 minutes 54 seconds ± 37 seconds (range, 2 minutes 54 seconds–5 minutes).

Reader Assessment

Anatomic Structures. After a brief self-guided training session based on 10 sinus MR imaging examinations not included in the study group, 3 radiologists independently reviewed the coronal T2-DRIVE sequences for the 25 patients in the study group with the Massachusetts Eye and Ear PACS. The radiologists included 2 fellowship-trained neuroradiologists (6 and 3 years’ subspecialty experience) and 1 neuroradiology fellow. They were asked to determine whether the trochlear groove, trochlear cistern, and cranial nerve IV were identifiable on each side in all patients, and discrepancies were resolved by consensus review. To be considered “identifiable,” the tentorial segment of cranial nerve IV had to be identified by the reader.
radiologist on at least 2 images for a given side, though the images did not have to be contiguous.

Artifacts. Potential confounding effects of MR imaging artifacts (motion, pulsation, and so forth) on cranial nerve IV visibility were assessed by consensus. Sequences were designated as containing artifacts that would “not preclude nerve identification,” “possibly preclude nerve identification,” or “likely preclude nerve identification.” This consensus review was performed 3 weeks following the anatomic structure identification consensus review to minimize potential for recall bias.

Measurements

When the anatomic structures of interest were identified, related measurements were obtained. For the trochlear groove, the maximum groove depth and the shortest distance between the free tentorial edge and the medial border of the groove were measured (Fig 3A). For the trochlear cistern, the vertical distance between the inferior margin of the trochlear cistern and the superior margin of the Meckel cave was measured (Fig 4A). For the tentorial segment of cranial nerve IV, the cross-sectional nerve diameter was measured.

Statistical Analysis

Mean, SD, and range are reported for continuous variables. Absolute and relative frequencies are reported for categorical variables. The Pearson $\chi^2$ test was used to compare proportions. For statistical analysis, consensus opinion regarding the identification of the trochlear groove and trochlear cistern was taken to indicate the “presence” or “absence” of these landmarks. Consensus opinion regarding identification of cranial nerve IV indicated whether the nerve was “identifiable” or “not identifiable.” JMP Pro, Version 12 (SAS Institute, Cary, North Carolina) was used for all analyses, and a $P$ value $<.05$ indicated a statistically significant difference.

RESULTS

Subjects

To reach a study group of 25 patients and 50 sides, we reviewed 32 examinations. Seven examinations were excluded for not containing the coronal T2 DRIVE sequence of interest. Patient age and sex and study indication data are provided in the Table.

Reader Assessment

Anatomic Structures. Interobserver agreement was 78% for identification of the trochlear groove, 56% for the trochlear cistern, and 68% for cranial nerve IV.

Following consensus review, the trochlear groove was present in 44/50 sides (88%), and the trochlear cistern was present in 25/50 sides (50%). In 46/50 sides (92%), either the trochlear groove or trochlear cistern was present. The tentorial segment of

**FIG 3.** Labeled (A) and unlabeled (B) coronal T2 DRIVE images demonstrate cranial nerve IV (black arrowhead on the patient’s right) in cross-section traveling along the bilateral trochlear grooves (straight white arrow on the patient’s right). The right tentorial free edge (curved black arrow), right cranial nerve V (V), right parahippocampal gyrus (PHG), belly of the pons (P), and basilar artery (B) are labeled for orientation. The dashed lines on the patient’s left illustrate the measurement of the maximum trochlear groove depth, and the solid lines on the patient’s left illustrate the measurement of the shortest distance between the free tentorial edge and the medial border of the trochlear groove.

**FIG 4.** Labeled (A) and unlabeled (B) coronal T2 DRIVE images demonstrate cranial nerve IV (white arrow) in cross-section within the right trochlear cistern. The right cranial nerve III (asterisk), right Meckel cave (M), and right parahippocampal gyrus (PHG) are labeled for orientation. The dashed lines illustrate the measurement of the vertical distance between the inferior margin of the trochlear cistern and the superior margin of the Meckel cave.

<table>
<thead>
<tr>
<th>Characteristics of the study group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td><strong>Age (yr)</strong></td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>SD</td>
</tr>
<tr>
<td>Max</td>
</tr>
<tr>
<td>Min</td>
</tr>
<tr>
<td><strong>Indication</strong></td>
</tr>
<tr>
<td>Sinonasal carcinoma</td>
</tr>
<tr>
<td>Inverted papilloma</td>
</tr>
<tr>
<td>Adenoid cystic carcinoma</td>
</tr>
<tr>
<td>Other ($n = 1$ each)</td>
</tr>
</tbody>
</table>

Note:—Max indicates maximum; Min, minimum.
craniocerebral nerve IV was identified in 36/50 sides (72%). When the trochlear groove was present, craniocerebral nerve IV was identified in 35/44 sides (80%), in contrast to 1/6 sides (17%) with no groove. This observed difference in proportions was statistically significant (P = .0013). When the trochlear cistern was present, craniocerebral nerve IV was identified in 23/25 sides (92%), in contrast to 13/25 sides (52%) with no cistern. This observed difference in proportions was also statistically significant (P = .0016).

When neither the trochlear groove nor trochlear cistern was present, craniocerebral nerve IV was identified in 0/4 sides (0%), in contrast to 36/46 sides (78%) with at least 1 landmark present (P = .0008).

Artifacts. None of the coronal T2 DRIVE sequences were considered to contain artifacts that would “likely preclude nerve identification,” and 4 examinations (8 sides) were thought to contain artifacts that would “possibly preclude nerve identification.” Among these 4 examinations with artifacts that would possibly preclude nerve identification, craniocerebral nerve IV was identified in 5/8 sides (63%). If these 8 sides had been initially excluded, the observed differences in proportions between craniocerebral nerve IV identification in the presence and absence of the trochlear groove and trochlear cistern would have remained statistically significant (P = .0196 and .0082, respectively).

Measurements
When present, the trochlear groove was located a mean distance of $1.4 \pm 0.44$ mm inferolateral to the free tentorial edge (range, 0.7–2.4 mm), and the mean maximum groove depth was $0.5 \pm 0.15$ mm (range, 0.2–0.8 mm). When present, the trochlear cistern was located a mean distance of $3.7 \pm 1.8$ mm superior to the Meckel cave (range, 1.2–7.0 mm). When identified, the mean diameter of the tentorial segment of craniocerebral nerve IV was $0.4 \pm 0.07$ mm (range, 0.2–0.6 mm).

DISCUSSION
This study describes the MR imaging appearances of the trochlear groove and trochlear cistern on a fluid-sensitive sequence used in our routine clinical practice and demonstrates that successfully identifying these anatomic landmarks facilitates identification of the tentorial segment of craniocerebral nerve IV.

Recent neurosurgical studies report that the trochlear groove and trochlear cistern are present in up to 100% of cadaver specimens and represent reliable anatomic landmarks for intraoperative nerve identification.6,11 We identified the trochlear groove in 88% of sides and the trochlear cistern in 50% of sides, with at least 1 of these landmarks seen in 92% of sides. The discrepancy in rates of landmark identification between the current MR imaging–based study and previous cadaveric studies likely reflects limitations of the coronal T2 DRIVE sequence for resolving these small structures, with the increased length of the trochlear groove relative to the trochlear cistern likely contributing to its relatively increased visibility.

We identified the tentorial segment of craniocerebral nerve IV in 72% of total sides, 80% of sides in which the trochlear groove was present, and 92% of sides in which the trochlear cistern was present. Although these results for the tentorial segment of craniocerebral nerve IV do not match those previously reported by Kanoto et al,6 who used a high-resolution motion-sensitized driven equilibrium sequence with a 26-minute acquisition time in a population of healthy volunteers, the coronal T2 DRIVE sequence investigated in our study with an average acquisition time of < 4 minutes is more likely to be of practical benefit when daily MR imaging volume imposes sequence acquisition time constraints. Additionally, patients being evaluated for cranial nerve deficits may be less likely to remain motionless during longer sequence acquisitions than healthy volunteers; this scenario may result in image-quality degradation. Image-acquisition parameters were not the focus of this study, and it is possible that the coronal T2 DRIVE sequence or a similar fluid-sensitive sequence could be modified to improve visualization of the trochlear groove, trochlear cistern, and craniocerebral nerve IV without meaningfully prolonging the acquisition.

Regardless of the particular sequence used, there are multiple clinical scenarios in which it is important for the radiologist to provide an accurate imaging assessment of craniocerebral nerve IV, in-
cluding suspected congenital absence of the nerve, involvement of the nerve by schwannoma or perineural spread of malignancy, trauma to the nerve related to penetrating injury or shear forces, and extrinsic compression of the nerve by a skull base mass, aneurysm, or crossing vessel.

Operative planning for skull base mass resection likely represents the scenario in which the radiologist has the greatest opportunity to positively impact patient care, because cranial nerve IV palsy is the most frequent complication following meningioma surgery in the region of the tentorial incisura. The frequency of this complication has been hypothesized to reflect difficulty in identifying and preserving the nerve at the time of the operation related to the small diameter of the nerve and displacement and distortion of the nerve along its lengthy intracranial course.

The relatively low interobserver agreement for the presence of the trochlear groove, trochlear cistern, and cranial nerve IV likely reflects both the difficulty of confidently identifying these small structures noninvasively and the relatively limited self-guided training of the readers before completing the study interpretations. Interobserver agreement may have been improved with a more extensive group training session before performing the study, and given that consensus was easily reached in all discrepant cases, we expect interobserver agreement to improve with continued experience.

A limitation of this study is that we were not able to confirm the presence of cranial nerve IV along its entire cisternal length, though the nerve had to be present on at least 2 images to be considered “identifiable.” In the evaluation of patients with trochlear nerve palsy at our institution, we have observed cases in which the nerve could be reliably identified along most of its course on the normal side but could not be seen within the trochlear groove or trochlear cistern on the contralateral, symptomatic side (Fig 5). Although anecdotal, the “empty” trochlear groove and “empty” trochlear cistern in such patients may represent important imaging findings.

An additional weakness of this study is the exclusion of pediatric patients, who represent a substantial proportion of patients being evaluated for suspected cranial nerve IV pathology and most patients with suspected congenital palsy.

Additional work could optimize sequence acquisition parameters for the visualization of the trochlear groove, trochlear cistern, and tentorial segment of cranial nerve IV; determine whether the trochlear groove and trochlear cistern are present in patients with congenital absence of cranial nerve IV; and assess whether the trochlear groove and trochlear cistern in combination with the coronal T2 DRIVE sequence or a sequence similar to it could be used for the following: 1) to identify cranial nerve IV in pediatric patients, 2) to reliably determine the presence or absence of cranial nerve IV, 3) to facilitate preoperative identification of cranial nerve IV in patients with skull base masses, or 4) to have a measurable effect on the historically high cranial nerve IV palsy complication rate following resection of tentorial meningiomas.

CONCLUSIONS
The trochlear groove and trochlear cistern are anatomic landmarks with which the tentorial segment of cranial nerve IV is closely associated en route to the cavernous sinus. These landmarks can be reliably identified by using a fluid-sensitive MR sequence and may improve the chances of identifying cranial nerve IV in routine clinical practice. These findings are expected to improve diagnostic confidence and diagnostic accuracy in the MR imaging evaluation of patients with suspected cranial nerve IV pathology.

REFERENCES
1. Fischbach F, Müller M, Bruhn H. High-resolution depiction of the cranial nerves in the posterior fossa (N III–N XII) with 2D fast spin echo and 3D gradient echo sequences at 3.0 T. Clin Imaging 2009;33:169–74 CrossRef Medline
ABSTRACT

BACKGROUND AND PURPOSE: As the practice of in utero repair of myelomeningoceles becomes more prevalent, knowledge of the expected MR imaging findings has become increasingly important. Our aim was to examine neuroimaging findings with a focus on hindbrain herniation and ventricular size in fetuses with open spinal dysraphism and to compare them with postnatal imaging features in groups undergoing prenatal-versus-postnatal repair.

MATERIALS AND METHODS: Single-center retrospective analysis was performed on MRIs of fetuses with open spinal dysraphism from January 2004 through July 2015 with available postnatal imaging. One hundred two fetuses were included. Reports from available fetal ultrasound were also examined. Images were reviewed by 2 board-certified fellowship-trained pediatric neuroradiologists. Descriptive analyses were performed to demonstrate the distribution of the imaging findings.

RESULTS: Thirty-two of 102 (31.3%) fetuses underwent in utero repair of open spinal dysraphism; 68.6% (70/102) underwent postnatal repair. Ninety-four of 102 (92.2%) fetuses had cerebellar ectopia. Of those who underwent prenatal repair (26 grade 3, 6 grade 2), 81.3% (26/32) had resolved cerebellar ectopia postnatally. Of those who had severe cerebellar ectopia (grade 3) that underwent postnatal repair, 65.5% (36/55) remained grade 3, while the remaining 34.5% (19/55) improved to grade 2. The degree of postnatal lateral ventriculomegaly in those that underwent prenatal repair (20.3 ± 5.6 mm) was not significantly different from that in those that underwent postnatal repair (21.5 ± 10.2 mm, \( P = .53 \)). Increased Chiari grade was significantly correlated with decreased head size for gestational age on fetal sonography (\( P = .0054 \)).

CONCLUSIONS: In fetuses with open spinal dysraphism and severe Chiari II malformation that do not undergo prenatal repair, most have no change in the severity of cerebellar ectopia/Chiari grade. However, in fetuses that undergo in utero repair, most have resolved cerebellar ectopia postnatally.

ABBREVIATION: OSD = open spinal dysraphism

Since the Management of Myelomeningocele (MOMS) trial revealed that prenatal surgery for myelomeningocele reduces the need for ventricular shunting and improves motor outcomes in infants with Chiari II malformation, the practice has become more widely available and is being offered and performed with increasing frequency.\(^1\) Fetal MR imaging continues to play an essential role in evaluating fetuses prenatally not only to determine which candidates may potentially benefit from the procedure but also to evaluate associated anomalies to shed some light on prognostic information that will aid in counseling.\(^2\)

Ample literature describes postnatal MR imaging findings of Chiari II malformation after postnatal repair of myelomeningoceles.\(^3,5\) However, the literature describing imaging findings in Chiari II malformation on fetal MR imaging is somewhat limited, mostly comprised of review articles with a few small studies.\(^6-10\)

There are even fewer articles in the literature comparing neuroimaging findings between pre- and postnatal brain MRIs in patients with Chiari II malformation.\(^2,11,12\) In order to improve our understanding of the disease and improving outcomes, studies evaluating the neuroimaging findings in these patients both pre- and postnatally are becoming increasingly important, particularly studies that describe imaging parameters that can be easily implemented in clinical practice.
Our aim was to evaluate neuroimaging findings in fetuses with open spinal dysraphism (OSD) and to compare them with the postnatal imaging findings, with a focus on hindbrain herniation and ventricular size.

MATERIALS AND METHODS

Study Design
This study was a single-center, retrospective review. The case list was compiled from all of the fetal MRIs performed between 2004 and 2015. Inclusion criteria encompassed fetuses with diagnostic-quality fetal MRIs for OSD (either myelomeningocele or myelocle). Patients underwent either immediate postnatal repair or open in utero repair of OSD. Fetuses without cerebellar ectopia (grade 1) underwent postnatal repair only. Only fetuses with adequate available postnatal neuroimaging and clinical/neurosurgical follow-up were included. Criteria for adequate postnatal neuroimaging for this study included a diagnostic-quality MR imaging of the brain within the first 3 months of life. Determination of diagnostic-quality imaging was made at the neuroradiologists’ discretion. A chart review was performed to obtain relevant clinical data, including reports from level II obstetric ultrasound when available. This study was compliant with the Health Insurance Portability and Accountability Act and approved by the institutional review board. The requirement for informed consent was waived.

Scanning Parameters
All fetuses included in our study were scanned prenatally on a 1.5T magnet at Cincinnati Children’s Hospital Medical Center by using a Ingenia 1.5T (Philips Healthcare, Best, the Netherlands) or a 1.5T Signa HDxt (GE Healthcare, Milwaukee, Wisconsin) system. T2 single-shot FSE images of the brain were obtained in 3 planes: axial, sagittal, and coronal. Three-millimeter-section thickness no-gap interleaved images at ≤24 weeks gestational age and 4-mm no-gap interleaved images at >24 weeks were used. Although this imaging protocol did not change during the study period, the TRs and TEs varied between each scanner and were changed at times of scanner upgrades to optimize image quality. At least 2 stacks in each plane were obtained to the radiologist’s satisfaction. The smallest FOV possible was used. Axial DWI and T1 spoiled gradient-recalled images of the fetal brain were inconsistently implemented at the radiologist’s discretion at the time of imaging. The postnatal imaging parameters varied because 1 of 6 available clinical magnets was used, and the routine use of volume T1 3D spoiled gradient-recalled imaging was not implemented until 2009.

Image Interpretation
All images were reviewed by 2 board-certified radiologists (U.D.N., B.M.K.-F.), both with added qualifications in pediatric radiology and fellowship training in pediatric neuroradiology, one (B.M.K.-F.) with >10 years of postfellowship attending experience in pediatric neuroradiology in a large academic center, the other (U.D.N.) with 2 years of experience. The images were viewed on a PACS workstation. Only diagnostic-quality MRIs for the assessment of the fetal brain were included in our study; image quality was a subjective assessment made by the neuroradiologists. All patients included in our analysis also had diagnostic-quality postnatal brain MRIs available for interpretation, as determined by the neuroradiologists (6 patients were excluded for this reason). The readers were blinded to the pre- and postnatal reported imaging findings at the time of interpretation and were also blinded to therapy (prenatal-versus-postnatal repair). Differences were resolved by consensus.

The degree of Chiari II malformation was determined by a grading scale in the fetus so that grade 1 was either normal or had a downward sloping tentorium and a normal patent fourth ventricle and cisterna magna without cerebellar ectopia; grade 2 had cerebellar ectopia with effacement of the fourth ventricle but a patent cisterna magna; and grade 3 had cerebellar ectopia and effacement of both the cisterna magna and fourth ventricle (Fig 1). This grading system was slightly modified from a previously reported system by Sutton et al in that we did not differentiate normal (grade 0) from grade 1 because the described difference was difficult to objectify. The same grading system was applied postnatally except that cerebellar ectopia with patency of the fourth ventricle or cisterna magna was considered grade 2. The grading system was modified postnatally because unlike fetal MR imaging in which we did not see cerebellar ectopia when the fourth ventricle was patent, we frequently observed varying degrees of fourth ventricle and cisterna magna patency with cerebellar ectopia, so to maintain 3 groups, we modified the grading scale.
Lateral ventricular size was measured in transverse dimensions in the axial plane at the level of the frontal horns or in the coronal plane at the level of the glomus of the choroid plexus, analogous to previously established fetal sonography guidelines. Fetal ventricular size was considered normal at <10 mm, mild-moderate ventriculomegaly at 10–15 mm, and severe ventriculomegaly at >15 mm. Postnatal ventricular size was considered normal at <10 mm, mild-moderate at 10–15 mm, severe at 16–25 mm, and extreme at >25 mm. Third ventricle size was measured in the transverse dimension in the coronal plane and was considered enlarged in the fetus if it was >3 mm.

Extra-axial CSF spaces over the cerebral hemispheres were described as either completely patent or effaced (including both partial and complete effacement). The presence or absence of ventricular rupture/dehiscence was determined, with rupture being defined as the presence of direct communication between the lateral ventricle and the extra-axial CSF space over the cerebral convexity.

**Statistical Analysis**
Descriptive analyses were performed to demonstrate the distribution of the imaging findings. Continuous variables were presented as mean ± SD, and categoric variables were presented as number (percentage). A 2-sample t test or 1-way ANOVA was used to detect the differences of continuous variables between different groups. The correlation between categoric variables was assessed by the χ² or Fisher exact test when appropriate. All analyses were performed by using SAS, Version 9.4 (SAS Institute, Cary, North Carolina). A P value of <.05 was considered statistically significant.

**RESULTS**
**Description of the Cohort**
A total of 102 fetuses (45 male, 57 female) met the criteria and were included in this analysis. Average gestational age at fetal MR imaging was 23.9 ± 3.7 weeks in the cohort as a whole, 24.6 ± 4.2 weeks in the postnatal repair group, and 22.4 ± 1.5 weeks in the prenatal repair group (P = .004). Thirty-two of 102 (31.3%) fetuses underwent open in utero repair of OSD, while the remaining 68.6% (70/102) underwent postnatal repair. The average gestational age at delivery was significantly lower in the prenatal repair group (20.3 ± 3.7 weeks) versus the postnatal (23.9 ± 4.2 weeks) repair group (P < .0001). Eighty-five of 102 (83.3%) had a level II obstetric sonogram performed within a week of the fetal MR imaging, with images and/or a detailed report available for review. The average age at postnatal brain MR imaging was 19.7 ± 21.9 days.

**Posterior Fossa**
Imaging findings in the cohort as a whole are summarized in Tables 1 and 2. Seventy of 102 (68.6%) fetuses underwent postnatal repair. Of these, 78.6% (75/97) had grade 3 and 10% (7/70) had grade 2 Chiari II malformation on fetal MR imaging. Of the 55 that were grade 3, 65.5% (36/55) remained grade 3 (Fig 2), while the remaining 34.5% (19/55) improved to grade 2 postnatally. Of the 12.7% (7/55) that were fetal grade 2, 2 progressed to grade 3, 3 remained at grade 2, and 2 had resolved ectopia (grade 1) postnatally. Fetuses without cerebellar ectopia (grade 1) underwent postnatal repair only. Of the 11.4% (8/70) that did not have fetal cerebellar ectopia (grade 1), most 87.5% (7/8) did not progress to a higher grade postnatally (Figs 1A and 3).

Thirty-two of 102 (31.4%) fetuses underwent open in utero repair of OSD, 81.3% (26/32) with grade 3 and 18.8% (6/32) with grade 2 Chiari II malformation. Of those with grade 3, 76.9% (20/26) had resolved (grade 1) cerebellar ectopia (Fig 4), 11.5% (3/26) had improved ectopia (grade 2), and 11.5% (3/26) had persistent grade 3 Chiari II malformation postnatally. Of those with grade 2, 100% (6/6) had resolved cerebellar ectopia postnatally after prenatal repair.

**Ventricle Size and Extra-Axial CSF Spaces**
Imaging findings, including the degree of ventriculomegaly, the presence of at least partial effacement of the extra-axial CSF spaces over the cerebral convexities, and lateral ventricular rupture/dehiscence in the cohort as a whole, are summarized in Tables 1 and 2.

The degree of postnatal lateral ventriculomegaly in those that underwent prenatal repair (20.3 ± 5.6 mm) was not significantly different from that in those that underwent postnatal repair (21.5 ± 10.2 mm, P = .53). The difference between the mean postnatal third ventricle size for the prenatal repair group (4.7 ± 2.8 mm) and the postnatal repair group (4.2 ± 1.5 mm) was also not statistically significant (P = .29). Postnatal lateral ventricular size in the 8 patients with grade 1 Chiari II malformation on fetal

| Table 1: Summary of imaging findings on fetal MR in the cohort as a whole |
|-----------------------------|-----------------|
| Imaging Findings | % of Cohort |
| Cerebellar ectopia | 92.2% (94/102) |
| Fetal Chiari grade | |
| 1 | 7.8% (8/102) |
| 2 | 12.7% (13/102) |
| 3 | 79.4% (81/102) |
| Effacement of preoptine cistern | 92.2% (94/102) |
| Lateral ventricular size | |
| Normal | 20.6% (21/102) |
| Mild-moderate ventriculomegaly | 57.8% (59/102) |
| Severe ventriculomegaly | 21.6% (22/102) |
| Third ventriculomegaly | 9.8% (10/102) |
| Lateral ventricular rupture | 1% (1/102) |
| Extra-axial CSF effacement | 92.1% (94/102) |

| Table 2: Summary of imaging findings on postnatal MR imaging in the cohort as a whole |
|-----------------------------|-----------------|
| Imaging Finding | % of Cohort |
| Cerebellar ectopia | 63.7% (65/102) |
| Postnatal Chiari grade | |
| 1 | 36.3% (37/102) |
| 2 | 26.5% (27/102) |
| 3 | 37.3% (38/102) |
| Effacement of preoptine cistern | 14.7% (15/102) |
| Lateral ventricular size | |
| Normal | 4.9% (5/102) |
| Mild-moderate ventriculomegaly | 21.6% (22/102) |
| Severe ventriculomegaly | 51% (52/102) |
| Extreme ventriculomegaly | 22.5% (23/102) |
| Lateral ventricular rupture | 2.9% (3/102) |
| Extra-axial CSF effacement | 22.5% (23/102) |
natal repair group. Of the 22 fetuses, 1 in the prenatal repair and (33/70) of the postnatal repair group and 3.1% (1/32) of the pre-catheter/shunt inplace at the time of postnatal MR imaging, 47.1% hindbrain herniation in either the pre- or postnatal repair groups. There was also no statistically significant relation-ship between lateral ventricular size and the presence of improved ventricular size. There was also no statistically significant relation-
ship between prenatal Chiari grade and prenatal postnatal third ventricle size, with a significant correlation between prenatal Chiari grade and (for Chiari grade) and 1-way ANOVA (for ventricle size) showed a significant correlation between the prenatal Chiari grade and prenatal lateral and third ventricle size was signifi-
cantly smaller than the postnatal ventricular size in both the pre-
natal and postnatal repair groups (P < .001). The Fisher exact test (for Chiari grade) and 1-way ANOVA (for ventricle size) showed a significant correlation between the prenatal Chiari grade and postnatal third ventricle size, with a P value of .0443. There was no significant correlation between prenatal Chiari grade and prenatal ventricular size (lateral and third ventricle) or postnatal lateral ventricular size. There was also no statistically significant relationship between lateral ventricular size and the presence of improved hindbrain herniation in either the pre- or postnatal repair groups.

Thirty-four of 102 (33.3%) fetuses had a ventricular drainage catheter/shunt in place at the time of postnatal MR imaging, 47.1% (33/70) of the postnatal repair group and 3.1% (1/32) of the pre-natal repair group. Of the 22 fetuses, 1 in the prenatal repair and 21 in the postnatal repair group, with se vere lateral ventriculomegaly (>15 mm), 54.5% (12/22) had a shunt at the time of postnatal MR imaging, com-pared with 27.5% (22/80) with a fetal lateral ventricular size of ≤15 mm. All except 2 (94.1%, 32/34) of the patients with a shunt on postnatal MR imaging had a grade 3 Chiari II malformation on fetal MR imaging. Only 12.5% (1/8) of patients with grade 1 Chiari II malformation on fetal MR imaging had a shunt on postnatal MR imaging. The Fisher exact test showed a significant correla-
tion between increased Chiari grade and the presence of prenatal extra-axial CSF space effacement, with a P value of < .0001. There was no significant correla-
tion between Chiari grade and postnatal extra-axial CSF space effacement.

**Sonographic Findings**

Reports from fetal ultrasound performed within 1 week of fetal MR imaging were available in 84 of the fetuses, 81 of which included amniotic fluid measurements. Of these, 96.3% (78/81) had reported normal amniotic fluid, 2.5% (2/81) had decreased amniotic fluid (deepest vertical pocket, ≤2 cm), and 1.2% (1/81) had increased amniotic fluid (deepest vertical pocket, ≥8 cm). Head circumference measurements were available in all 84 fetuses. Of these, 35.7% (30/84) had normal head circumference, 61.9% (52/84) had decreased head circumference (<10th percentile), and 2.4% (2/84) of fetuses had increased head circumference (>90th percentile) for gestational age. The Fisher exact test showed that an increased Chiari grade was significantly corre-
lated with decreased head size, with a P value of .0054.

**DISCUSSION**

We describe pre- and postnatal MR imaging brain findings in pa-
patients with OSD with a focus on hindbrain herniation. We found that 92.2% (94/102) of fetuses had, prenatally and presurgically, cerebel-
lar ectopia (grade 2 or 3), while 7.8% (8/102) did not. In those with severe/grade 3 cerebellar ectopia who underwent postnatal repair, 65.5% (36/55) remained grade 3, while the remaining 34.5% (19/55) improved to grade 2. Of those who underwent prenatal repair, 81.3% (26/32) had resolved cerebellar ectopia postnatally (Table 3). We found a significant correlation between increased fetal Chiari grade and the presence of prenatal extra-axial CSF space effacement. We also found that the degree of postnatal lateral ventriculomegaly in those that underwent prenatal repair was not significantly different from that in those that underwent postnatal repair. Finally, we ob-
served that an increased fetal Chiari grade was significantly correlated with decreased head size on fetal sonography.

The MOMS trial, in addition to describing the reduced need for shunting and improved motor outcomes in fetuses that un-
dergo prenatal repair of a myelomeningocele, also reported im-
proved hindbrain herniation.1 Smaller clinical studies in other centers have had similar results.19-21 Our findings are in keeping
with these studies because we found that 81% of fetuses that underwent prenatal repair had resolved cerebellar ectopia postnatally, a result that was not observed in any of the fetuses that underwent postnatal repair. Our study adds to the clinical literature by describing in more detail the imaging findings on both pre- and postnatal MR imaging by adapting the grading system used by Sutton et al.4 on fetal MR imaging, which can be applied to clinical practice. It seems clear that prenatal repair, by preventing leakage of CSF through the open neural tube defect, can progressively remodel the posterior fossa and revert the Chiari malformation to different degrees as we saw in the MR imaging studies of our patients. We also describe the evolution of the Chiari grade in the patients with postnatal repair in the cohort, including the 35% of fetuses with grade 3 Chiari who underwent postnatal repair and improved to grade 2; this improvement has not been previously described, to our knowledge.

The association between OSD and Chiari II malformation in infants has been long established and is accepted as a constant by many.4,3,13 The widely accepted unified theory of McLone and Dias22 and McLone and Knepper23 describes how defective occlusion caused by an open neural tube defect precludes CSF accumulation and pressure within the developing cerebral ventricles, causing the findings of Chiari II malformation, including a small posterior fossa with hindbrain herniation. However, in our study, 8% of fetuses with postnatally confirmed OSD did not have cerebellar ectopia on fetal MR imaging. Although there is little in the literature describing this phenomenon, one other study described an incidence as high as 23%.10 Our findings support the idea that not all fetuses with OSD have cerebellar ectopia. In addition, our study adds the finding that nearly all of these fetuses (7/8) did not develop cerebellar ectopia in the postnatal period after repair, which suggests that hindbrain malformation in Chiari is unlikely to worsen. This suggestion may be helpful for prenatal counseling.

There is very little in the literature analyzing ventricle size in fetuses with OSD. One study found that on fetal MR imaging ventriculomegaly (>10 mm) was associated with the presence of cerebellar ectopia in fetuses with OSD.18 However, we demonstrate not only that there was no significant association between Chiari grade and ventricle size on fetal MR imaging but also that there was no statistically significant difference in postnatal ventricle size between the pre- and postnatal repair groups. This finding questions the role of ventricular size in shunt placement, because the MOMS trial and our own data illustrate a decreased need for shunting in patients who undergo prenatal repair. The decision to place a shunt does not rely on ventricle size alone, and other clinical factors are taken under consideration.1 In addition, our study only examines the initial postnatal brain MR imaging, so it is possible that differences in ventricular size may develop later and that the patients with prenatal repair may have stabilization of ventricular dilation that was previously obstructive. Evaluation of ventricle size may be clinically relevant because an update to the MOMS trial suggests that larger lateral ventricle size at initial screening is associated with an increased need for shunting and prenatal repair may not improve outcomes in fetuses with ventricles of >15 mm.24 In our study, only 1 fetus of 32 that underwent prenatal repair had severe ventriculomegaly (>15 mm). However, we observed that of the 22 fetuses with severe lateral ventriculomegaly in both groups combined, 54.5% (12/22) had a shunt at the time of postnatal MR imaging, compared with 27.5% (22/80) with a fetal lateral ventricular size of ≤15 mm. This finding may suggest that in fetuses with a ventricle size of >15 mm, there may be a degree of obstruction that is more prone to progressively worsen.

Also very little in the literature exists describing the extra-axial CSF spaces over the cerebral convexities in fetuses with Chiari II, with only 1 study describing the depth of this space being lower in patients with Chiari II than in controls.19 Our study adds to the literature by not only describing a high association between CSF effacement and higher Chiari grade in fetuses but also showing that this association does not persist in the postnatal period after repair. When we examined head size, 62% of the fetuses in our study had decreased head circumference.

![Figure 4](image-url)  
**Figure 4.** Sagittal T2 SSFSE from fetal MR imaging at 23 weeks’ and 3 days’ gestational age (A) with severe or grade 3 Chiari II malformation (arrow). Sagittal T2 FSE image from postnatal MR imaging in the same patient at 2 weeks of age after in utero repair of OSD (B) demonstrates resolved cerebellar ectopia (arrow).

**Table 3: Benefits of prenatal-versus-postnatal repair by imaging**

<table>
<thead>
<tr>
<th></th>
<th>Prenatal Repair Group (n = 32)</th>
<th>Postnatal Repair Group (n = 70)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of fetal grade 3 Chiari II malformation that remained grade 3 postnatally</td>
<td>11.5% (3/26)</td>
<td>65.5% (36/55)</td>
</tr>
<tr>
<td>% of fetal grade 3 Chiari II malformation that improved or resolved postnatally</td>
<td>88.5% (23/26)</td>
<td>34.5% (19/55)</td>
</tr>
<tr>
<td>% of fetal grade 2 Chiari II malformation that resolved postnatally (grade 1)</td>
<td>100% (6/6)</td>
<td>28.6% (2/7)</td>
</tr>
<tr>
<td>% that had a ventricular shunt catheter on postnatal MRI</td>
<td>3.1% (1/32)</td>
<td>47.1% (33/70)</td>
</tr>
</tbody>
</table>
on fetal sonography, which is consistent with previously published literature. We also found that an increased Chiari grade on fetal MR imaging is associated with decreased head size on fetal ultrasound, which has not been previously described. This finding may suggest that decreased head size is a reflection of effaced extra-axial CSF spaces and that fetuses with decreased head circumference may have more benefit from in utero repair than those with normal or increased head circumference. It may be interesting to perform further studies examining head size in utero after prenatal repair to see whether head size has the potential to normalize with this procedure.

Our study does have some limitations. First, the retrospective nature of this study limits its internal validity. Second, given that this is a single-institution study performed within a certain timeframe, the external validity may be limited as well. Along those same lines, our study likely has some degree of selection bias, given that these data are from one of the largest referral centers for fetal repair of OSD in the country, which we started performing at our institution in 2011. This may explain why the gestational age at fetal MR imaging is significantly less in the prenatal repair group compared with the postnatal repair group, as prenatal repair is typically performed before 26 weeks gestational age. However, differences in gestational age are unlikely to substantially affect our ventricular size data because it has been documented in the literature that lateral ventricular size remains relatively constant throughout gestation. Given that imaging studies were acquired during a 10-year period on multiple different clinical magnets with periodic upgrades, the heterogeneity of scanning parameters may affect our results as well.

CONCLUSIONS
We describe imaging parameters on pre- and postnatal brain MR imaging in patients with OSD and find that in fetuses with severe Chiari II malformation that do not undergo prenatal repair, most (65.5%) have no change in the severity of cerebellar ectopia/Chiari grade postnatally. In fetuses that undergo in utero repair, most (81.3%) have resolved cerebellar ectopia postnatally. We also describe how cerebellar ectopia is not always present in fetuses with OSD on fetal MR imaging. We found no significant difference in the degree of postnatal lateral ventriculomegaly between those that underwent prenatal repair and those that underwent postnatal repair (P = .53). Finally, we found a relationship between the degree of cerebellar ectopia with head size on fetal ultrasound and effacement of extra-axial CSF spaces on fetal MR imaging. The exact clinical implications of these findings are yet to be determined, and long-term clinical follow-up of these patients will be essential.

REFERENCES
Brain Development in Fetuses of Mothers with Diabetes: A Case-Control MR Imaging Study


ABSTRACT

BACKGROUND AND PURPOSE: Offspring exposed to maternal diabetes are at increased risk of neurocognitive impairment, but its origins are unknown. With MR imaging, we investigated the feasibility of comprehensive assessment of brain metabolism ([1H-MRS]), microstructure (DWI), and macrostructure (structural MRI) in third-trimester fetuses in women with diabetes and determined normal ranges for the MR imaging parameters measured.

MATERIALS AND METHODS: Women with singleton pregnancies with diabetes (n = 26) and healthy controls (n = 26) were recruited prospectively for MR imaging studies between 34 and 38 weeks’ gestation.

RESULTS: Data suitable for postprocessing were obtained from 79%, 71%, and 46% of women for [1H-MRS, DWI, and structural MRI, respectively. There was no difference in the NAA/Cho and NAA/Cr ratios (mean [SD]) in the fetal brain in women with diabetes compared with controls (1.74 [0.79] versus 1.79 [0.64], \( P = .81 \); and 0.78 [0.28] versus 0.94 [0.36], \( P = .12 \), respectively), but the Cho/Cr ratio was marginally lower (0.46 [0.11] versus 0.53 [0.10], \( P = .04 \)). There was no difference in mean [SD] anterior white, posterior white, and deep gray matter ADC between patients and controls (1.16 [0.12] versus 1.16 [0.08], \( P = .96 \); 1.54 [0.16] versus 1.59 [0.20], \( P = .56 \); and 1.49 [0.23] versus 1.52 [0.23], \( P = .89 \), respectively) or volume of the cerebrum (243.0 mL [22.7 mL] versus 253.8 mL [31.6 mL], \( P = .38 \)).

CONCLUSIONS: Acquiring multimodal MR imaging of the fetal brain at 3T from pregnant women with diabetes is feasible. Further study of fetal brain metabolism in maternal diabetes is warranted.

ABBREVIATIONS: T1DM = type 1 diabetes mellitus; T2DM = type 2 diabetes mellitus; GDM = gestational diabetes; sMRI = structural MRI

Diabetes is the most common medical disorder of pregnancy with the prevalence of type 1 (T1DM), type 2 (T2DM), and gestational diabetes (GDM) all increasing among women of child-bearing age in resource-rich settings. The perinatal complications of maternal diabetes, which reflect altered metabolic function in utero, include major congenital malformations, macrosomia, and stillbirth. At increased risk for cognitive impairment, impaired working memory, and altered language development. These adverse outcomes are not fully explained by postnatal events; this question focuses research attention on the vulnerability of the developing brain during fetal life. Identification of the nature and timing of alterations to brain structure and function that underlie neurocognitive impairment could help the development of strategies designed to improve the long-term outcome of children of diabetic mothers.

During fetal life, the predominant source of brain energy is glucose, which crosses the placenta by facilitated diffusion. While severe perturbations in glucose homeostasis after birth are associated with neonatal brain injury, the effect of chronic fluctuant glucose concentration experienced by fetuses of women with diabetes on in utero brain development has not been investigated, to our knowledge. Maternal diabetes is also associated with disturbances in fatty acid metabolism: Umbilical venous blood docosahexaenoic acid concentration is reduced; this reduction reflects lower docosahexaenoic acid transfer to the fetus. Docosahexaenoic acid accumulates in the brain in abundance from the...
third trimester and is essential for neurogenesis, neurotransmission, and protection from oxidative stress. Reduced bioavailability of this key metabolite has been suggested as a putative mechanism for programming altered neurodevelopment.8,9

Advances in proton MR spectroscopy (1H-MRS) and diffusion-weighted and structural MR imaging (sMRI) have led to the development of objective and sensitive measures of fetal brain structure and metabolism. Use of these technologies has revealed alterations in the cerebral NAA:choline ratio and gyriﬁcation in fetuses with congenital heart disease,10 temporal lobe volumes in fetuses with congenital cytomegalovirus infection,11 and ADC values and parenchymal volume in antenatal ventriculomegaly.12,13 Historically, most fetal imaging studies have been undertaken at 1.5T. However, although an increasing number of studies have been performed at 3T field strength,14–20 which has beneﬁts over 1.5T due to improved signal-to-noise and is likely to be advantageous for depicting fetal anatomy,21 to date, there have been no studies assessing the feasibility of recruiting women with diabetes for fetal neuroimaging.

Early-life metrics derived from 1H-MRS, DWI, and sMRI are associated with function in childhood. After preterm birth, NAA/Cho and Cho/Cr ratios are associated with neurodevelopmental outcome at 2 years of age,22 lactate/NAA predicts outcome following hypoxic-ischemic encephalopathy,23 and abnormalities in the NAA/Cr and Cho/Cr ratios in neonates24 and older children25 predict developmental delay. Increased ADC values in white matter are associated with diffuse white matter injury following preterm birth26 and with poor outcome after hypoxic-ischemic encephalopathy in term infants.27,28 Finally, reduced regional and whole-brain volumes are associated with speciﬁc preterm comorbidities,29,30 and structural alteration predicts long-term impairment after preterm birth.31,32

On the basis of disturbances to fetal glucose and fatty acid metabolism associated with maternal diabetes and the neurocognitive proﬁle of offspring, we aimed to investigate the feasibility of comprehensive fetal brain assessment by acquiring measurements of NAA/Cho, NAA/Cr, and Cho/Cr ratios; regional apparent diffusion coefﬁcient measurements; and volume of the cerebrum during the third trimester of pregnancy from women with diabetes and from healthy controls by using 3T MR imaging. The secondary aim was to determine normal values for these measures for future studies designed to investigate the effect of maternal disease on fetal brain development and in utero origins of neurodevelopmental impairment.

MATERIALS AND METHODS

Study Population

Ethics approval was obtained from the National Research Ethics Committee (South East Scotland Research Ethics Committee), and written informed consent was obtained. Women with a pregnancy complicated by diabetes (n = 26) and healthy controls (n = 26) were recruited prospectively from antenatal diabetes clinics at the Simpson Centre for Reproductive Health at the Royal Infirmary, Edinburgh, UK. The inclusion criteria were a singleton pregnancy and a fetal anomaly scan with normal ﬁndings at 20 weeks’ gestation. Women with diabetes were eligible to participate if they had gestational diabetes, diagnosed by using the Scottish Intercollegiate Guideline Network diagnostic criteria33 as a fasting venous plasma glucose of ≥5.1 mmol/l or 2-hour glucose of ≥8.5 mmol/l after a 75-g oral glucose tolerance test or pregestational type 1 or type 2 diabetes. Exclusion criteria were serious coexisting maternal systemic disease other than maternal diabetes and women with any contraindications to MR imaging, including metal implants and pacemakers.

MR Image Acquisition

MR imaging studies were performed at the Clinical Research Imaging Centre in the Queen’s Medical Research Institute, University of Edinburgh, UK, by using a Magnetom Verio 3T MR imaging clinical scanner (Siemens, Erlangen, Germany). To avoid vena-cava compression, we placed women in a left-lateral tilt, with blood pressure being constantly monitored by using a Veris MR imaging vital signs monitor (Medrad, Indiana, Pennsylvania). No fetal sedation was used, women were limited to spending 45 minutes in the scanner, and data were acquired with women free-breathing throughout. MR images were obtained between 34 and 38 weeks’ gestation. A radiologist with experience in MR imaging reported all images.

T2-weighted half-Fourier-acquisition single-shot turbo spin-echo images were acquired of the fetal brain in sagittal, coronal, and transverse orientations (HASTE: TR/TE = 1800/86 ms, FOV = 400 × 400 mm, matrix = 192 [phase] × 256 [frequency], section thickness = 8 mm, acquisition time = 18 seconds). These images were used to plan the position of the single 20-mm3 spectroscopy voxel within the fetal brain. The scanner bed was moved to ensure that the fetal brain was positioned at the isocenter, and the voxel was positioned within 1 hemisphere of the fetal brain, avoiding ventricles and contaminant signal from surrounding tissue. An optimized semiautomated shimming protocol was systematically applied until the full width at half maximum of the water peak was <20 Hz. A single-voxel point-resolved spectroscopy technique was applied with TR/TE = 1500/30 ms, 96 signal averages, bandwidth = 2000 Hz, and a water suppression bandwidth = 50 Hz. The spectral acquisition took 2 minutes 30 seconds. Signal was received from selected elements of the spine matrix coil and body matrix surface coils positioned to allow adequate coverage of the fetal brain. A postspectroscopy 3-plane HASTE acquisition was then compared with the prespectroscopy HASTE images to allow visual assessment of fetal movement during the spectral acquisition. If the expert operator observed evidence of movement between HASTE acquisitions, then the spectroscopy voxel was repositioned and the spectral acquisition was repeated. No additional filtering or quality-control limiting of data was applied during the processing stage. We therefore processed all the MR spectroscopy data that were acquired. An example of voxel positioning for the MR spectroscopy acquisition is shown in Fig 1A.

Transverse DWIs of the whole fetal brain (TR/TE = 7300/106 ms, FOV = 400 × 400 mm, matrix = 128 × 128, section thickness = 3 mm, b-values = 0, 300, and 1000 s/mm2) were acquired. DWI was checked at the point of acquisition for obvious signs of fetal motion and was repeated if required. ADC maps were generated automatically from the diffusion-weighted images.

Finally, additional transverse HASTE images were acquired with identical coverage to the DWIs to aid subsequent ROI anal-
Data Analysis: \( ^1H \text{-MRS} \)
Spectral analysis was performed by using the QUEST algorithm available in jMRUI (www.mrui.uab.es/mrui/mrui_download/). This technique estimates metabolite amplitudes by using a non-linear least-squares fit of simulated metabolite signals to the acquired spectrum. A metabolite basis set was generated by using the NMR-Scope function available in jMRUI and included contributions from NAA (2.01, 2.49, and 2.70 ppm), Cho (3.20, 3.53, and 4.08 ppm), and Cr (3.04 and 3.93 ppm). We then calculated the following ratios: NAA/Cho, NAA/Cr, and Cho/Cr. The QUEST algorithm calculates errors associated with the estimated metabolite amplitudes by using an extended version of the Cramer–Rao lower bounds calculation. The errors for each of the calculated metabolite ratios were derived through error propagation of the jMRUI output.

Data Analysis: Diffusion and sMRI
Apparent Diffusion Coefficients.
ROI analysis was performed on ADC maps by using standard software on the 3T Magnetom Verio MR imaging system (Siemens). First, ROIs within white matter and gray matter were identified from the HASTE images acquired in the same plane and with the same coverage as the diffusion-
weighted images. A section above the ventricles was identified as white matter, and a section at the level of the thalami was identified as deep gray matter by using landmarks described in Boardman et al. The identical sections were then identified on the corresponding ADC map; 4 ROIs were positioned in the white matter (2 posterior and 2 anterior) and 2 were positioned in the gray matter. Due to differences in fetal brain volume, we used an anatomically appropriate ROI size for each individual brain, taking care to avoid partial volume effects from adjacent structures and artifacts. The mean (SD) ADC value for each ROI was recorded. The mean white matter ROI size was 0.30 ± 0.12, and the mean gray matter ROI size was 0.32 ± 0.13. Sample ROI placements for white and gray matter are shown in Fig 1B. Interrater agreement was checked by 2 independent investigators (D.A., G.M.).

**sMRI.** For each participant, a single 3D motion-corrected brain volume was reconstructed by using a section-to-volume registration method (Fig 1C). The fetal brain was extracted from surrounding fetal and maternal tissue by using an atlas-based approach. All reconstructed images were nonlinearly aligned to the closest age-matched template from a publicly available 4D fetal brain atlas. Then, an automatic method based on an expectation-maximization framework for brain tissue segmentation was used, in which the priors of brain tissues were propagated by using prior probabilities provided by the 4D atlas. Finally, binary masks of the cerebral (intracranial contents excluding intraventricular CSF, extra-axial CSF, and the choroid plexus, brain stem, cerebellum, and pons structures) and the intracranial volume (GM, WM, and CSF) were deformed to the subject’s native space, and volumes were calculated.

**Statistical Analysis**

This was a feasibility study, so a formal power calculation for sample size was not required. For normally distributed data, the mean and SD are reported, and for non-normally distributed data, the median and interquartile range are reported. For groupwise comparisons of normally distributed variables, an independent-sample *t* test was used, and for skewed data, the Mann-Whitney *U* test was used. To analyze regional ADC values, we first tested for evidence of laterality in the anterior and posterior white matter and deep gray matter values with a paired-samples *t* test, and if there were no significant differences between left and right, the values were averaged to compute mean anterior white matter ADC, mean posterior white matter ADC, and mean deep gray matter ADC per individual. The distributions were assessed for normality, and an independent-samples *t* test was used for groupwise comparisons of regional ADC. Interobserver agreement in ADC measurements was assessed for each region in a randomly selected subset of 20 participants by using Bland-Altman statistics. For group-wise analysis of NAA/Cho, NAA/Cr, and Cho/Cr ratios and cerebral and intracranial volumes, an independent-samples *t* test was used after assessing the equality of variance between groups. Statistical analyses were performed by using SPSS 21 (IBM, Armonk, New York) with statistical significance defined as *P* < .05.

**RESULTS**

**Participants**

The maternal demographics and delivery outcomes of the study population are shown in Table 1. All women tolerated the MR imaging well, and no scan had to be abandoned due to maternal discomfort or claustrophobia. Of the women with diabetes, 13 were diagnosed with GDM during pregnancy, 12 had T1DM, and 1 had T2DM. In women with GDM, the median gestation at diagnosis and diagnosis-to-scan interval were 27.1 weeks (interquartile range, 12.0–31.0 weeks) and 8.9 weeks (interquartile range, 4.4–23.6 weeks), respectively. Only 1 woman with GDM was treated with diet alone. The other 12 were treated with metformin (*n* = 9) or metformin and insulin (*n* = 3) to achieve glycemic control. All women with T1DM were insulin-treated, and the 1 woman with T2DM was treated with insulin and metformin. The hemoglobin A1c (glycolated hemoglobin) at booking (11–13 weeks’ gestation) for women with T1DM and T2DM was 51.9 mmol/mol (16.6 mmol/mol). Two women with GDM, 4 women with T1DM, and 1 control had antenatal steroids for fetal lung maturation before MR imaging. Three neonates of women with T1DM were admitted to the neonatal unit for other medical reasons (72 hours).

---

**Table 1: Demographics, MRI details, and delivery outcomes**

<table>
<thead>
<tr>
<th>Maternal demographics</th>
<th>Control (<em>n</em> = 26)</th>
<th>All (<em>n</em> = 26)</th>
<th>GDM (<em>n</em> = 13)</th>
<th>T1DM (<em>n</em> = 12)</th>
<th>T2DM (<em>n</em> = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age (mean) (SD) (yr)</td>
<td>31 (5)</td>
<td>31 (5)</td>
<td>32 (5)</td>
<td>30 (6)</td>
<td>34</td>
</tr>
<tr>
<td>Parity (median) (range)</td>
<td>0 (0–3)</td>
<td>0 (0–3)</td>
<td>1 (0–2)</td>
<td>0 (0–3)</td>
<td>0</td>
</tr>
<tr>
<td>Current smoker (No.) (%)</td>
<td>1 (4)</td>
<td>3 (12)</td>
<td>1 (8)</td>
<td>2 (17)</td>
<td></td>
</tr>
<tr>
<td>Deprivation (No.) (%)</td>
<td>SIMD 1–3</td>
<td>SIMD 4–5</td>
<td>SIMD 1–3</td>
<td>SIMD 4–5</td>
<td>SIMD 1–3</td>
</tr>
<tr>
<td>MRI details</td>
<td>Gestation at MRI (mean) (SD) (wk)</td>
<td>36.1 (0.9)</td>
<td>36.0 (0.8)</td>
<td>36.0 (0.8)</td>
<td>36.0 (0.9)</td>
</tr>
<tr>
<td></td>
<td>MRI-to-delivery interval (mean) (SD) (wk)</td>
<td>3.6 (1.6)</td>
<td>2.1 (1.2)</td>
<td>2.6 (1.2)</td>
<td>1.6 (1.1)</td>
</tr>
<tr>
<td>Neonatal outcome</td>
<td>Gestation delivery (mean) (SD) (wk)</td>
<td>39.7 (1.5)</td>
<td>38.1 (1.4)</td>
<td>38.6 (1.1)</td>
<td>37.6 (1.5)</td>
</tr>
<tr>
<td></td>
<td>Birthweight (mean) (SD) (g)</td>
<td>3372 (467)</td>
<td>3551 (627)</td>
<td>3629 (483)</td>
<td>3508 (780)</td>
</tr>
<tr>
<td></td>
<td>Sex (male/female)</td>
<td>13/13</td>
<td>9/17</td>
<td>6/7</td>
<td>2/10</td>
</tr>
<tr>
<td></td>
<td>Occipitofrontal circumference (mean) (SD) (cm)</td>
<td>34.4 (1.4)</td>
<td>34.8 (1.8)</td>
<td>35 (1.6)</td>
<td>35 (2.2)</td>
</tr>
</tbody>
</table>

**Note:** SIMD indicates Scottish Index of Multiple Deprivation; SIMD 1, most deprived; SIMD 5, most affluent.
for bacteria) and transient low blood glucose (n = 1), a fractured clavicle sustained during a forceps delivery with shoulder dystocia, and a duplication cyst that was not diagnosed antenatally. No neonates born to healthy controls required admission. All neonates were discharged home alive and well.

There was no difference in the gestation in weeks at MR imaging between women with diabetes and healthy controls (mean [SD]) (36.0 weeks [0.8 weeks] versus 36.1 weeks [0.9 weeks], P = .69). No adjustment was therefore made for gestational age in the statistical analysis. No congenital anomalies, acquired brain injuries, or incidental findings were detected by MR imaging.

**MR Spectroscopy**

In utero ¹H-MRS of the fetal brain of suitable quality for analysis was obtained in 41/52 (79%) women in the study population, 22/26 (85%) with diabetes and 19/26 (73%) healthy controls. There was no difference in the clinical characteristics of women in whom interpretable data were acquired compared with those in whom they were not (data not shown). There was no difference in the NAA/Cho and NAA/Cr ratios in the fetal brain in women with diabetes compared with controls (1.74 [0.70] versus 1.79 [0.64], P = .81; and 0.78 [0.28] versus 0.94 [0.36], P = .12, respectively). The Cho/Cr ratio was marginally lower in the fetal brain in women with diabetes compared with controls (0.46 [0.11] versus 0.53 [0.10], P = .04) (Fig 2).

**Diffusion-Weighted Imaging: ADC**

DWIs amenable to ADC computation were available for 37/52 (71%) women in the study population, 18/26 (69%) with diabetes and 19/26 (73%) healthy controls. Fetal motion or maternal size prevented interpretable data from being obtained from 9/52 (17%) of the study population. There was no difference in the clinical characteristics of women in whom interpretable data were acquired compared with those in whom they were not (data not shown).

There was no evidence of laterality in the anterior white matter, posterior white matter, or deep gray matter ADC values (all P > .05). Data were therefore combined to 3 variables: mean anterior white matter, mean posterior white matter, and mean deep gray matter ADC. There was no difference in mean (SD) ADC values for anterior white matter, posterior white matter, and deep gray matter in women with diabetes mellitus compared with controls (1.16 [0.12] versus 1.16 [0.08], P = .96; 1.54 [0.16] versus 1.59 [0.20], P = .56; and 1.49 [0.23], versus 1.52 [0.23], P = .89, respectively) (Fig 3).

There was good interrater agreement between the 2 independent investigators for ADC values. The mean difference and 95% confidence intervals between investigators for anterior white matter, posterior white matter, and deep gray matter measurements are reported in Table 2.

**Brain Volumes**

Tissue segmentation data suitable for analysis were used to assess the macrostructure of the fetal brain in 24/52 (46%) of the study population (9/26 [35%] women with diabetes, 15/26 [58%] healthy controls). Fetal motion or data quality prevented interpretable data from being obtained from 28/52 (54%) women of the study population. There was no difference in mean cerebrum volume/milliliter in women with diabetes compared with controls (243.0 ± 22.7 mL³ versus 253.8 ± 31.6 mL³, P = .39). There was no difference in mean intracranial volume in fetuses of women with diabetes compared with controls (265.0 ± 22.5 mL³ versus 274.5 ± 32.3 mL³, P = .47).

**DISCUSSION**

In this study, we demonstrated that it is feasible to recruit pregnant women with diabetes to undergo MR imaging at 3T during the third trimester of pregnancy for measurements of NAA/Cho, NAA/Cr, and Cho/Cr ratios; regional ADC measurements; and cerebrum and intracranial volumes. We chose to acquire ¹H-MRS, DWI, and sMRI because of their use as markers of tissue injury/altered metabolism in the neonatal period and their relationships with long-term outcome. The values we acquired con-
tribute useful normative data for future fetal brain studies performed with 3T systems.

Although this feasibility study was not powered to detect group differences, we observed a marginal-but-significant reduction in Cho/Cr in the brains of fetuses of diabetic mothers during the third trimester. The MR spectroscopy choline peak included free choline, phosphocholine, and glycerophosphocholine, so these data raise the possibility that brain metabolism and neuronal membrane phospholipid turn-over are altered in pregnancies with women with diabetes. While this finding requires confirmation in a larger study, alterations in the Cho/Cr ratio in brains of adults with type 2 diabetes have been reported.44

A strength of our study is that we recruited a cohort of women with well-characterized diabetes with all participants being scanned within a 4-week time window and gestation matched to our control group. This feature is important because 1H-MRS spectra and ADC values are dynamic during this period of brain development.45-47 We also acquired sMRI suitable for conventional clinical reporting, which was possible for all participants. A limitation of our study is that we were unable to acquire data amenable to quantitative analysis on all fetuses scanned. Despite ensuring the comfort of the women in a large-bore scanner, data could not be processed from 1H-MRS in 21% of cases: DWI in 29% of cases and sMRI in 54% of cases. The low data yield for sMRI was partly because acquisition of 1H-MRS and DWI was prioritized over sMRI. For future study designs that require fetal brain segmentation, yield may be increased by modifications to the acquisition protocol such as the increasing the number of stacks per plane and accepting the idea that time constraints required for safety may curtail other acquisitions (we capped imaging at 45 minutes). sMRI suitable for conventional clinical reporting was available for all participants.

We chose to recruit a heterogeneous population of women with diabetes to assess the feasibility of dissecting the effect of different in utero exposure to T1DM, T2DM, and GDM in a future study. Recruitment of women with T1DM and GDM was relatively easy; thus, recruitment to a future study assessing the effect of in utero exposure to T1DM and GDM on the fetal brain would be feasible. In contrast, we were able to recruit only 1 woman with T2DM, due to the lower prevalence of this condition. Thus, targeting recruitment of women with T2DM to a future study will not be practical unless recruitment occurs across multiple sites.

Our data were acquired by using a 3T system as opposed to 1.5T. For the advanced imaging techniques used in this study, there are advantages of acquiring data with the higher field strength of 3T.48 Compared with lower field strengths, imaging at higher field strengths increases the signal-to-noise ratio. This increase improves the spectral quality obtained in 1H-MRS and the ability to differentiate closely located metabolites, particularly at short TEs. The inability to complete data acquisition within the time available due to fetal movement is a major limitation of MR imaging in pregnancy. Acquiring data more rapidly by using more advanced imaging methodologies, using methods of motion correction to compensate for fetal movement, and using alternative sampling techniques such as compressed sensing is likely to greatly increase data yield in the future. Finally, one advantage of 3T is the ability to acquire images with higher spatial resolution (depending on the imaging coil used), potentially increasing diagnostic accuracy.49

Perinatal image metrics are sensitive to tissue injury and neuroprotective treatment strategies. They are therefore increasingly used to address the “gap in translation” in perinatal neuroscience to assess therapies that show promise in preclinical studies at lower economic and opportunity costs than randomized controlled trials powered on clinical outcomes.50 The normative data provided here may inform the development of fetal brain biomarkers for use in interventional perinatal neuroprotective outcome studies.

**CONCLUSIONS**

The data provide proof-of-concept that comprehensive assessment of the fetal brain using measures derived from images acquired at 3T from women with diabetes and healthy controls is achievable. In addition, they suggest that fetal brain MR spectroscopy may provide a promising image marker of altered brain development in maternal diabetes. Finally, although we studied fetuses of mothers with diabetes, this research pipeline and the normative values obtained could be applied to any paradigm in which fetal origins of brain development are being investigated by using 3T MR imaging.

**ACKNOWLEDGMENTS**

We are grateful to the women who consented to take part in the study, to the research midwives, and to the nursing and radiography staff at the Clinical Research Imaging Centre, University of Edinburgh (http://www.cric.ed.ac.uk) who participated in scanning the women.

Disclosures: Fiona C. Denison—RELATED: Grant: TheirWorld, Comments: block grant, research funded by TheirWorld to cover work undertaken and reported in the article*. UNRELATED: Board Membership: Vice-Chair of the National Institute of Clinical Excellence Committee, Comments: I am Vice-Chair of the National Institute of Clinical Excellence Medical Technologies Advisory Committee. My travel expenses are paid to attend this monthly meeting. I am not paid directly to attend; Grants/Grants Pending: National Institute for Health and Care Excellence, Medical Research Council, Edinburgh and Lothians Health Foundation, and Tommy’s. Comments: I am Co-Investigator or Principal Investigator of several research grants funded by the National Institute for Health and Care Excellence, Medical Research Council, Edinburgh, and Lothians Health Foundation. All of these grants fund separate studies, and there is no overlap with the work undertaken in article. Gillian Macnaught—UNRELATED: Employment: GlaxoSmithKline, Comments: Forty percent of my previous post was paid by GlaxoSmithKline; however, they were not involved in this submitted work.* Scott I.K. Semple—UNRELATED: Other: GlaxoSmithKline, Comments: GlaxoSmithKline sponsored a research fellow in my group, not connected to this work.*
REFERENCES


23. Hanrahan JD, Cox II, Azzopardi D, et al. Relation between proton magnetic resonance spectroscopy within 18 hours of birth asphyxia and neurodevelopment at 1 year of age. Dev Med Child Neurol 1999;41:76–82 CrossRef Medline


ABSTRACT

BACKGROUND AND PURPOSE: Arthrogryposis is among the malformations of congenital Zika syndrome. Similar to the brain, there might exist a spectrum of spinal cord abnormalities. The purpose of this study was to explore and describe in detail the MR imaging features found in the spinal cords, nerve roots, and brains of children with congenital Zika syndrome with and without arthrogryposis.

MATERIALS AND METHODS: Twelve infants with congenital Zika syndrome (4 with arthrogryposis and 8 without) who had undergone brain and spinal cord MR imaging were retrospectively selected. Qualitative and quantitative analyses were performed and compared between groups.

RESULTS: At visual inspection, both groups showed reduced thoracic spinal cord thickness: 75% (6/8) of the group without arthrogryposis and 100% (4/4) of the arthrogryposis group. However, the latter had the entire spinal cord reduced and more severely reduced conus medullaris anterior roots (respectively, \(P = .002\) and \(P = .007\)). Quantitative differences were found for conus medullaris base and cervical and lumbar intumescences diameters (respectively, \(P = .008\), \(P = .048\), \(P = .008\)), with more prominent reduction in arthrogryposis. Periventricular calcifications were more frequent in infants with arthrogryposis (\(P = .018\)).

CONCLUSIONS: Most infants had some degree of spinal cord thickness reduction, predominant in the thoracic segment (without arthrogryposis) or in the entire spinal cord (with arthrogryposis). The conus medullaris anterior roots were reduced in both groups (thinner in arthrogryposis). A prominent anterior median fissure of the spinal cord was absent in infants without arthrogryposis. Brain stem hypoplasia was present in all infants with arthrogryposis, periventricular calcifications, in the majority, and polymicrogyria was absent.

ABBREVIATIONS: AACD = Association for Assistance of Disabled Children; GRE = gradient recalled-echo; IgM = immunoglobulin M
and how the Brazilian Zika virus strain could have developed this neurotropism for the central nervous system is still unknown.

The disease has already spread and, according to the World Health Organization, 76 countries and territories, particularly in Latin America, have reported evidence of transmission of the Zika virus by mosquitoes. Cases of microcephalic infants have been reported in 29 countries.10

In addition to microcephaly, other serious brain abnormalities were observed, especially brain calcifications, predominantly in the cortical and subcortical white matter junction, associated with malformations of cortical development (often polymicrogyria or pachgyria with predominant frontal lobe involvement) and a simplified cortical gyral pattern. Other frequent imaging findings are ventriculomegaly; decrease in brain, brain stem, and cerebellar volumes; enlargement of the cisterna magna and the extraaxial subarachnoid space; corpus callosum abnormalities (hypogenesis and hypoplasia); and delayed myelination.11

The congenital Zika syndrome is an entity without a well-known clinical spectrum, probably with only the most severe cases of the spectrum recognized. Other malformations have been described in some infants, such as ophthalmologic alterations12,13 and arthrogryposis.14,15 Currently, 8% of the children with presumed congenital Zika virus infection followed by the Association for Assistance of Disabled Children (AACD) in Recife, Brazil have arthrogryposis. Among the children with CSF immunoglobulin M (IgM) who tested positive for Zika virus, 6.6% have arthrogryposis.

Arthrogryposis multiplex congenita, often known simply as arthrogryposis, is a syndrome characterized by joint contractures, present since birth, affecting ≥2 areas of the body.15-20 These joint malformations can be attributed to different disorders, such as defects of uterine environment, disorders of connective tissues, muscular dystrophies, and other abnormalities or conditions that affect the central or peripheral nervous systems in at least one of the components of the motor pathways from the spinal cord to muscles.16,18 Regardless of the cause, children affected by arthrogryposis have onset and severe weakness early in intrauterine life, with immobilization of joints at different developmental stages.16

No study has yet analyzed qualitatively and quantitatively MR imaging of the spinal cord of children with congenital Zika syndrome, to our knowledge. Because there is a spectrum of congenital Zika syndrome for brain abnormalities, a similar spectrum might occur in the spinal cord. Therefore, the aim of this study was to explore and describe in detail the MR imaging features found in the spinal cord and nerve roots of infants with congenital Zika syndrome with or without arthrogryposis.

**MATERIALS AND METHODS**

This retrospective series included 12 infants diagnosed with congenital Zika syndrome, based on the Brazilian government protocol,8 who had or did not have arthrogryposis and met the following inclusion criteria: 1) brain imaging suggestive of any congenital infection; 2) a complete investigation with negative findings, including laboratory testing, of the other 6 main infectious causes of primary microcephaly (toxoplasmosis, cytomegalovirus, rubella, syphilis, herpes simplex, and HIV); 3) negative results of an investigation, through clinical examination and family history, for causes of microcephaly and brain calcifications such as prenatal and perinatal complications; exposure to licit and illicit drugs, toxic substances, and ionizing radiation; genetic and metabolic diseases; and congenital infections; and 4) the existence of previously performed brain and spinal cord MR imaging.

Since October 2015, infants with suspected microcephaly born in Pernambuco are reported on a government Web site (cievspe.com/microcefalia) by health professionals. The criterion for microcephaly up to December 2015 was a head circumference of ≤33 cm. Therefore, all the infants born before December 2015, including all the ones included in this study, were referred for investigation of congenital Zika syndrome because they had a head circumference of up to 33 cm (currently, the criterion for microcephaly is based on the International Fetal and Newborn Growth Consortium for the 21st Century [INTERGROWTH-21st]). Once microcephaly was detected, the infants in this study were referred to the AACD, where they were evaluated, as needed, by a multidisciplinary team. The AACD has become one of the major referral centers in Pernambuco for the diagnosis and follow-up of children with microcephaly and congenital Zika syndrome.

Besides a noncontrast brain CT scan, obtained in all suspected cases of microcephaly as part of the government protocol, all infants in this study also underwent MR imaging, based on clinical evaluation. The major indications for brain MR imaging were the presence of refractory seizures or suspicion of hydrocephalus, while the major indications for spinal cord MR imaging were suspicions of arthrogryposis or the presence of early-onset abnormal posture of the limbs. A reduced essential protocol for neuroaxis MR imaging was created in our service with a reduced sedation time.

This retrospective study was approved by the Federal University of Pernambuco Research Ethical Committee, and the children’s mothers or guardians gave their consent for the publication of the results and images. Infants included in this study were chosen by convenience, following the inclusion criteria. Up to the beginning of the study, 12 infants with congenital Zika syndrome (4 with arthrogryposis and 8 without it) had undergone, postnatally, brain and spinal cord MR imaging at our service.

Inclusion of controls was not possible because the techniques and section thickness of spinal MR imaging of infants up to 1 year found retrospectively at our service (5 cases) were different from the 12 already-included cases, making comparison very difficult. Nevertheless, a brief qualitative comparison, by visual inspection, was still performed (Fig 1).

MR imaging of the whole spinal cord was performed in all 12 children. All the images were T2-weighted. The protocol was optimized to decrease the examination time. Therefore, T1 was not performed in the spinal cord, and the duration of an MR imaging procedure for the brain and spinal cord was 50 minutes, approximately. Of the 12 children, 2 underwent MR imaging in a 1.5T scanner (Intera; Philips Healthcare, Best, the Netherlands); and 10, in a 1.5T scanner (Signa HDxt; GE Healthcare, Milwaukee, Wisconsin). The 2 examinations performed in the Philips MR imaging scanner had technical parameters similar to those of the GE scanner, while measurements were performed on balanced fast-field echo sequences. No contrast was used in any of the infants, while sedation was necessary for all of them. Table 1 shows
the parameters used for brain and spinal cord MR imaging. The images were analyzed and measured by 2 experienced neuroradiologists (M.F.V.V.A. and A.M.B.-L.), with the final interpretation determined by a consensus between them. Discrepancies between observers were not common, and when they occurred, they were resolved jointly after further review of the images.

Measurements of the spinal cord and spinal canal of these infants were taken by using the PACS software. To account for the spinal cord and spine curvatures, we performed all the measurements on sagittal images, as shown in Fig 2 (sagittal T2 volumetric GRE (E–H), we can observe reduction of the conus medullaris anterior roots (long white arrows) compared with posterior roots (short white arrows).

Table 1: MR imaging parameters for brain and spinal cord image acquisition for all of the participants

<table>
<thead>
<tr>
<th>MRI Parameters</th>
<th>Brain</th>
<th>Spinal Cord</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Axial 3D</td>
<td>Axial 3D</td>
</tr>
<tr>
<td>TE (ms)</td>
<td>50</td>
<td>Minimum</td>
</tr>
<tr>
<td>TR (ms)</td>
<td>78.3</td>
<td>–</td>
</tr>
<tr>
<td>Flip angle</td>
<td>15°</td>
<td>12°</td>
</tr>
<tr>
<td>Bandwidth (kHz)</td>
<td>41.67</td>
<td>31.25</td>
</tr>
<tr>
<td>FOV (cm)</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Section thickness (mm)</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Spacing (mm)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Frequency (Hz)</td>
<td>288</td>
<td>256</td>
</tr>
<tr>
<td>Phase</td>
<td>224</td>
<td>256</td>
</tr>
<tr>
<td>NEX</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>Frequency direction</td>
<td>AP</td>
<td>AP</td>
</tr>
</tbody>
</table>

Note: —AP indicates anteroposterior; R-L, right-left; FSPGR, fast-spoiled gradient recalled; SWAN, susceptibility-weighted angiography.

FIG 1. Comparison of spinal cord MR imaging between a control 12-month-old child (A–D) and a 4-month-old infant with presumed congenital Zika syndrome and arthrogryposis (E–H). Sagittal T2 (A) shows a normal-sized spinal cord and conus medullaris and no abnormal signal. Axial reformatted T2 reveals symmetric and normal-sized anterior and posterior nerve roots in the conus medullaris (B and C) and cauda equina (D). Meanwhile, sagittal T2-weighted volumetric GRE image (E) shows reduced spinal cord thickness, especially in the thoracic region (short white arrows). On the axial reconstruction of T2-weighted volumetric GRE (F–H), we can observe reduction of the conus medullaris anterior roots (long arrows) compared with posterior roots (short arrows).

by visual inspection. The anteroposterior diameter of the conus medullaris base was also measured. It is defined as the inferior portion of the spinal cord, at the location where the spinal cord begin to reduce its thickness to form the conus medullaris. The medullaris cone is usually located, in infants, in the level of L2–L3 or above.22 Similarly, the vertebral canal was measured at approximately the same levels as the spinal cord measurements, to obtain a spinal cord/vertebral canal ratio for each measure of the spinal cord. These measurements were compared between children with and without arthrogryposis.

Other features of the spinal cord were qualitatively evaluated and compared between the 2 groups on axial T2 FSE or volumetric GRE and balanced fast-field echo reconstructions. The features evaluated in the axial plane were reduction of conus medullaris roots (graded as mild, moderate, and severe) and prominence of the anterior median fissure of the spinal cord (absent or present). On the sagittal plane, the thickness for each spinal cord segment (normal or decreased) was evaluated. The hip joints were assessed for developmental dysplasia on coronal FSE T2WI.

Additionally, brain MR imaging findings were compared between the 2 groups. We reviewed the MR images for the following: decreased brain volume (graded as mild, moderate, and severe); cerebral ventricular enlargement due to white matter hypoplasia; malformations of cortical development and sulcation and their locations; abnormalities of the corpus callosum (classified as agenesis, hypogenesis, and hypoplasia); myelination (normal or delayed, based on a previous study); the presence and location of brain calcifications; decreased brain stem and cerebellar volume; an enlarged cisterna magna; an enlarged anterior supratentorial subarachnoid space; and the presence of intraparenchymal cysts.

Symmetry of brain damage was also evaluated. It was determined by visual comparison between the cerebral hemispheres. The criteria used to describe pathologic asymmetry were those used in clinical practice, mainly ventricle size, sulci enlargement, cerebral lobe size, and spatial displacement of the left and right hemispheres with respect to each other. When the differences between the hemispheres were according to the normal pattern, even though they were not equally sized, the damage was considered symmetric.

Statistical Analysis

Absolute and percentage values were calculated to describe the qualitative variables and median and interquartile ranges for con-

---

tinuous variables. For association, we used the Fisher exact test. The Mann-Whitney test was applied to compare continuous variables. A \( P \) value \(< .05\) was significant, while a \( P \) value \(< .1\) was considered a statistical trend. The authors opted to include the data with a statistical trend because it points to the possibility of finding statistically significant results if the sample size is increased. The statistical analyses were performed with the Statistical Package for Social Sciences software, Version 21.0 (IBM, Armonk, New York).

**RESULTS**

Of the analyzed sample of 12 children, 7 (58%) were boys, and of the total, 4 (33%) had arthrogryposis. Maternal mean age was 27.4 ± 8.5 years. Gestational age ranged from 35 to 40 weeks, with an average of 38.2 ± 1.6 weeks. Nine of the mothers reported a rash, 6 (66.7%) in the first trimester and 3 (33.3%) in the second trimester. Forty percent of girls in this study had arthrogryposis, while 29% of boys had arthrogryposis, with no statistical difference \( (P = 1.000)\). Table 2 shows the individual clinical data of the 12 children and the IgM CSF status positive for Zika virus for all 8 children tested. The mean age at MR imaging was 135.83 days (131.25 days for the arthrogryposis group and 138.16 days for the group without arthrogryposis). The On-line Table shows the individual radiologic data of the 12 children.

At visual inspection, both groups showed reduction of high thoracic spinal cord thickness: 75% (6/8) without arthrogryposis and 100% (4/4) with arthrogryposis. The thoracic spinal cord

**Table 2: Individual clinical data of the 12 infants, 8 without and 4 with arthrogryposis, included in the study**

<table>
<thead>
<tr>
<th>No.</th>
<th>Sex</th>
<th>Gestational Age (wk)</th>
<th>HC at Birth (cm)</th>
<th>Mother’s Rash during Pregnancy</th>
<th>Joints Affected</th>
<th>IgM ZIKV CSF Statusa</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>F</td>
<td>35</td>
<td>29.5</td>
<td>2 mo</td>
<td>Feet, knees, hips, elbows, wrists, fingers</td>
<td>Positive</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>36</td>
<td>31.5</td>
<td>4 mo</td>
<td>Feet, knees, hips, wrists, fingers</td>
<td>No data</td>
</tr>
<tr>
<td>3</td>
<td>F</td>
<td>40</td>
<td>30</td>
<td>No rash</td>
<td>Feet, knees, hips</td>
<td>No data</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>39</td>
<td>26</td>
<td>3 mo</td>
<td>Feet, knees, hips</td>
<td>Positive</td>
</tr>
<tr>
<td>5</td>
<td>M</td>
<td>39</td>
<td>31</td>
<td>3 mo</td>
<td>Feet, knees, hips</td>
<td>Positive</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>39</td>
<td>28.5</td>
<td>2.5 mo</td>
<td>Feet, knees, hips</td>
<td>Positive</td>
</tr>
<tr>
<td>7</td>
<td>M</td>
<td>39 wk 5 days</td>
<td>32</td>
<td>4 mo</td>
<td>Feet, knees, hips</td>
<td>Positive</td>
</tr>
<tr>
<td>8</td>
<td>M</td>
<td>39 wk 5 days</td>
<td>28</td>
<td>3 mo</td>
<td>Feet, knees, hips</td>
<td>Positive</td>
</tr>
<tr>
<td>9</td>
<td>F</td>
<td>37</td>
<td>29</td>
<td>No rash</td>
<td>Feet, knees, hips, elbows, wrists, fingers</td>
<td>Positive</td>
</tr>
<tr>
<td>10</td>
<td>M</td>
<td>37</td>
<td>26</td>
<td>4 mo</td>
<td>Feet, knees, hips, wrists, fingers</td>
<td>No data</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>40</td>
<td>27</td>
<td>No rash</td>
<td>Feet, knees, hips</td>
<td>No data</td>
</tr>
<tr>
<td>12</td>
<td>F</td>
<td>38</td>
<td>30</td>
<td>2 mo</td>
<td>Feet, hips, wrists, fingers</td>
<td>Positive</td>
</tr>
</tbody>
</table>

Note:—HC indicates head circumference; ZIKV, Zika virus.

a Test performed on the infant.
In addition, half of these infants had congenital hip dysplasia. In addition, half of these infants were noted to have severe reduction of the anterior nerve roots of the conus medullaris compared with the children without arthrogryposis, with statistical significance (respectively, \( P = .002 \) and \( .007 \)). However, 6 of the children without arthrogryposis also had reductions of the anterior conus medullaris roots (4 had mild and 2 had moderate reduction), and 6 had congenital hip dysplasia. The groups were also found to be different, with a statistical trend \( (P = .091) \), regarding the prominence of the anterior median fissure of the spinal cord, with none of the children without arthrogryposis having this feature.

Thus, in summary, the 4 infants with congenital Zika syndrome and arthrogryposis had, at visual inspection, reduced thickness of all segments of the spinal cord, with the thoracic segment its most compromised portion. All of them also had a severe reduction of the anterior nerve roots of the conus medullaris and congenital hip dysplasia. In addition, half of these infants had a prominent anterior median fissure, a feature not identified in any of the children without arthrogryposis.

Overall, the anteroposterior diameters of different levels of the spinal cord were smaller in infants with arthrogryposis compared with those without (Table 4). Statistically significant differences were found between the children with and without arthrogryposis regarding the conus medullaris base and lumbar and cervical intumescence measurements (respectively, \( P = .008, .008, .048 \)). In addition, a statistical trend was observed regarding the smallest cervical spinal cord diameter \( (P = .073) \). No statistical difference was found in the smallest high thoracic spinal cord anteroposterior diameters between the 2 groups.

The anteroposterior diameter of the vertebral canal was statistically similar between the 2 groups. The ratio between the spinal cord diameters and these vertebral canal measurements, at similar levels of the spinal cord, revealed statistically significant differences in the conus medullaris base and the lumbar intumescence between children with and without arthrogryposis \( (P = .016 \) and \( .008) \) (Table 4).

Figures 3 and 4 show children with congenital Zika syndrome with arthrogryposis, while Fig 5 shows a child with congenital Zika syndrome without arthrogryposis. Reduced spinal cord thickness (Figs 3D and 4C), severe reduction of anterior conus medullaris roots (Figs 3E–G and 4G, H), and congenital hip dysplasia (Fig 3I) were found in all children with arthrogryposis.

Statistically significant differences between the groups with and without arthrogryposis were found only in periventricular calcifications \( (P = .018) \) when brain abnormalities were compared, and only a statistical trend was observed for cerebellar or brain stem hypoplasia, cerebellar calcifications, and brain stem calcifications (Table 5); those findings were more frequent in infants with arthrogryposis than without it, being brain stem hypoplasia found in all 4 infants, and cerebellar hypoplasia in 2 of them. The cerebral damage was severe in all the children with arthrogryposis, while only 3 without

---

**Table 3: Comparison between groups with and without arthrogryposis of features identified on MRI by visual inspection**

<table>
<thead>
<tr>
<th>Sites with spinal cord reduction</th>
<th>Ag (n = 4)</th>
<th>No Ag (n = 8)</th>
<th>( P^b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>0 (0.0%)</td>
<td>2 (25.0%)</td>
<td>.002^c</td>
</tr>
<tr>
<td>Thoracic</td>
<td>0 (0.0%)</td>
<td>5 (62.5%)</td>
<td></td>
</tr>
<tr>
<td>Cervical, thoracic</td>
<td>0 (0.0%)</td>
<td>1 (12.5%)</td>
<td></td>
</tr>
<tr>
<td>Cervical, thoracic, conus medullaris</td>
<td>4 (100.0%)</td>
<td>0 (0.0%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Site of most severe spinal cord reduction</th>
<th>Ag (n = 4)</th>
<th>No Ag (n = 8)</th>
<th>( P^b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>0 (0.0%)</td>
<td>2 (25.0%)</td>
<td>.999</td>
</tr>
<tr>
<td>Thoracic</td>
<td>4 (100.0%)</td>
<td>6 (75.0%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prominence of the anterior median fissure of the spinal cord</th>
<th>Ag (n = 4)</th>
<th>No Ag (n = 8)</th>
<th>( P^b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>2 (50.0%)</td>
<td>8 (100.0%)</td>
<td>.091</td>
</tr>
<tr>
<td>Present</td>
<td>2 (50.0%)</td>
<td>0 (0.0%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reduction of the anterior nerve roots of the conus medullaris</th>
<th>Ag (n = 4)</th>
<th>No Ag (n = 8)</th>
<th>( P^b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>No reduction</td>
<td>0 (0.0%)</td>
<td>2 (25.0%)</td>
<td>.007^c</td>
</tr>
<tr>
<td>Mild</td>
<td>0 (0.0%)</td>
<td>4 (50.0%)</td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>0 (0.0%)</td>
<td>2 (25.0%)</td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>4 (100.0%)</td>
<td>0 (0.0%)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Congenital hip dysplasia</th>
<th>Ag (n = 4)</th>
<th>No Ag (n = 8)</th>
<th>( P^b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>0 (0.0%)</td>
<td>2 (25.0%)</td>
<td>.515</td>
</tr>
<tr>
<td>Present</td>
<td>4 (100.0%)</td>
<td>6 (75.0%)</td>
<td></td>
</tr>
</tbody>
</table>

**Note:**—Ag indicates arthrogryposis.

^a Data are number of patients (%).

^b P = Fisher Exact test.

^c Statistically significant.

---

**Table 4: Comparison between groups with and without arthrogryposis of features identified on MRI by visual inspection**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AP Diameters (mm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Smallest spinal cord diameters</th>
<th>Cervical</th>
<th>High thoracic</th>
<th>Conus medullaris base</th>
<th>Largest spinal cord diameters</th>
<th>Lumbar intumescence</th>
<th>Cervical intumescence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag (n = 4)</td>
<td>4.5 (4.4–4.9)</td>
<td>3.6 (2.8–4.2)</td>
<td>4.3 (3.5–4.8)</td>
<td>4.7 (4.5–5.0)</td>
<td>5.0 (3.9–5.8)</td>
<td></td>
</tr>
<tr>
<td>No Ag (n = 8)</td>
<td>5.4 (4.7–5.9)</td>
<td>3.9 (3.6–4.5)</td>
<td>6.3 (5.7–6.5)</td>
<td>5.6 (5.4–6.2)</td>
<td>7.3 (7.1–7.5)</td>
<td></td>
</tr>
<tr>
<td>( P^b )</td>
<td>.073</td>
<td>.214</td>
<td>.008^c</td>
<td>.048^c</td>
<td>.008^c</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spinal Cord/Vertebral Canal Ratio</th>
<th>Spinal Cord/Vertebral Canal Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ag (n = 4)</td>
<td>0.44 (0.41–0.47)</td>
</tr>
<tr>
<td>No Ag (n = 8)</td>
<td>0.57 (0.44–0.60)</td>
</tr>
<tr>
<td>( P^b )</td>
<td>.154</td>
</tr>
</tbody>
</table>

**Note:**—AP indicates anteroposterior; Ag, arthrogryposis.

^a Data are median (25th–75th percentile).

^b P = Mann-Whitney test.

^c Statistically significant.
arthrogryposis had severe damage, though there were no statistical differences between the groups.

Regarding malformations of cortical development, no statistical differences were found. None of the infants with arthrogryposis had polymicrogyria, while this malformation was present in 2 children without arthrogryposis (P = .515). Pachygyria and a simplified gyral pattern were more often seen among the infants with arthrogryposis (both P = .576): One had diffuse pachygyria, one had diffuse simplified gyral pattern, and the other 2 had both malformations, with frontal pachygyria and a simplified parieto-occipital gyral pattern.

DISCUSSION

We analyzed the spinal cords of children with microcephaly with congenital Zika syndrome with and without arthrogryposis. One-third of our sample had arthrogryposis, a figure that is not representative of the entire sample of patients with congenital Zika syndrome in the AACD.

By visual inspection, the arthrogryposis group had significant qualitative reduction of the entire spinal cord and severe reduction in the anterior conus medullaris roots. However, most of the children without arthrogryposis also had mild reduction of the anterior conus medullaris roots, and often just the thoracic spinal cord was reduced. This finding could explain the lack of statistical differences in the thoracic spinal cord segment between the groups evaluated in quantitative analysis.

Significant quantitative differences were found between the groups, with arthrogryposis showing a thinner conus medullaris base, lumbar intumescence, and cervical intumescence. The arthrogryposis group was also significantly thinner at the conus medullaris base and at the lumbar intumescence by the evaluation of ratios with the canal diameter (anteroposterior diameters of the spinal cord/vertebral canal).

These findings support the hypothesis that the congenital Zika syndrome has a disease-severity spectrum. The spectrum is not restricted to the brain, but a disease spectrum is also present in the spinal cord and spinal roots. Clinically, this spectrum would range from absent or mild manifestations to arthrogryposis. Thus, arthrogryposis would be the most severe extreme of the spectrum of spinal cord damage, with thinner thickness of the entire spinal cord and severe anterior nerve root reduction. However, the MRIs of infants without arthrogryposis have also demonstrated some grade of damage in the spinal cord, mainly in the thoracic segment, with some mild anterior spinal root reduction.

Our study has no control group for quantitative analysis. In addition, we have found no references in the literature to normal spinal cord measurements in infants. The control group is a group difficult to obtain retrospectively, especially with the same MR imaging technique for accurate comparison.

Arthrogryposis has been associated with microcephaly in the spectrum of congenital Zika virus infection. Schuler-Faccini et al identified arthrogryposis in 4 of 27 children, while Oliveira Melo et al mentioned 1 child with the condition. In addition, Melo et al described 3 neonates who died shortly after birth. The MR imaging findings in arthrogryposis were recently reported, but that series of cases did not evaluate the spinal cord of infants without arthrogryposis. This study identified apparently reduced spinal cord thickness and reduced ventral roots in comparison with the dorsal roots. Regarding neurogenic arthrogryposis in patients without congenital Zika virus infection, Fedrizzi et al described brain and spinal cord MR imaging findings in 10 patients without an identified etiology.

The histopathologic changes in neurogenic arthrogryposis are dysgenesis of the anterior medullary horns and cytoarchitectural disorganization, which are more prominent in cervical and lumbar intumescence.
The above-mentioned features were previously suggested by brain and spinal MR imaging evaluation of infants with arthrogryposis and reported by the authors and are consistent with what was found in the present study, reinforcing MR imaging being able to demonstrate the physiopathology of the congenital Zika syndrome. In this context, Zika virus probably has a tropism for the brain and also for the motor spinal cord neurons, which could occur initially in the thoracic region and achieve cervical and lumbar intumescences and conus medullaris in more severe cases.

Tropism of the Zika virus for neurons, leading eventually to their death, has been shown in the literature. The morphologic brain alterations with multiple calcifications, mainly in the junction between the cortical and subcortical white matter, induced by the Zika virus, suggest serious damage that may cause sudden arrest in the development of the nervous system, resulting in a simplified pattern of cortical circulations, malformations of cortical development (mainly in the frontal lobes), associated with ventriculomegaly, and corpus callosum and brain stem hypoplasia. Long tract and spinal cord neuron alterations may occur in congenital Zika virus infection, even in patients without arthrogryposis. This hypothesis is supported by Mlakar et al., who described infant brain abnormalities and Wallerian degeneration in the descending tracts of the brain stem and spinal cord, while the ascending tracts of the dorsal columns were well-preserved, in a 32-week-old fetus with confirmed real-time polymerase chain reaction for Zika virus infection without arthrogryposis.

In our study, the arthrogryposis group showed significantly more frequent periventricular calcifications and a trend toward more frequent cerebellar and brain stem hypoplasia and calcifications. One hypothesis that could be taken from these data is that the Zika virus damage to the human neural progenitor cells in the periventricular zone will be more pronounced in children with arthrogryposis. Another possible explanation is major damage in the basal gangla and pyramidal tracts, which could lead to or be associated with brain stem and cerebellar hypoplasia and spinal cord abnormalities. We hypothesized that both primary and secondary damage to the spinal cord are possible. We observed that the virus probably has great tropism for motor neurons, both in the brain (frontal lobes) and spinal cord (ventral spinal cord). In this context, MR imaging is important in understanding the physiopathology of congenital Zika syndrome and clarifying that the joint malformations found in these children are due to the virus tropism for specific motor neurons in related areas of the brain stem, cerebellum, and spinal cord and not to direct action of the virus in the osteoarticular system.

Although without significance, the overall brain damage caused by the Zika virus was more pronounced in the arthrogryposis group. Another interesting observation is that polymicrogyria was absent in the 4 children with arthrogryposis, while pachygyria was present in almost all the infants with arthrogryposis. Polymicrogyria is a feature related to an interruption in the late stages of neuronal migration and cortical organization that originates only after the twentieth gestational week, while pachygyria is believed to originate in the early phase of pregnancy, between the twelfth and sixteenth gestational weeks. These data can indicate that congenital Zika syndrome with arthrogryposis is more often associated with earlier fetal infection, further supported by the more severe damage found in these infants. However, we did not find an association of the type of malformation of cortical development with the time of the mother’s rash.

Nevertheless, the literature has shown that there is an interval between the maternal infection and the sonographic evidence of fetal abnormalities from 2 to 27 weeks. We suppose that the month of the mother’s rash does not necessarily indicate exactly...
when the embryo or fetus was infected, whether early or late during infection of the mother in the pregnancy.

On the basis of the findings described in this study, it is important to consider Zika virus infection in the differential diagnosis of congenital spinal cord and anterior nerve root diseases if the infant and mother have a positive epidemiologic context. This is especially important in mild cases in which microcephaly is absent and the only clinical manifestation is, for example, abnormal joints. On the other hand, health professionals should pay close attention during the follow-up of children from an epidemic area with mild or no clinical signs of spinal cord and anterior nerve root damage because they could possibly have future problems in their neuropsychomotor development. This can also be true for children without microcephaly, born in regions with the Zika virus epidemic.

It is difficult to determine the prognosis of the different degrees of the congenital Zika syndrome due to the lack of follow-up studies; however, congenital Zika virus infection with severe brain damage should have a poor prognosis. Knowledge of the spectrum of this syndrome can be helpful in identifying which cases could have higher chances of worse outcomes. It is probable that infants with arthrogryposis will have worse prognoses, especially in motor development, even if they do not have severe brain lesions or microcephaly.

Despite the limitations, especially regarding the small number of patients and lack of a control group for quantitative analysis, this study is the first to analyze spinal cord MR imaging abnormalities in children with congenital Zika virus infection without arthrogryposis. In addition, this study raises the alarming hypothesis that children without clear signs of impairment (e.g., microcephaly and arthrogryposis) can have brain and spinal cord imaging abnormalities probably caused by the Zika virus, a possibility that is starting to be seen in clinical practice.

The 8 children with congenital Zika syndrome without arthrogryposis did not have this major clinical manifestation of spinal cord impairment but had mild radiologic spinal cord abnormalities, such as a qualitative decrease in

---

Table 5: Comparison between children with and without arthrogryposis regarding some of the brain abnormalities found on MRI

<table>
<thead>
<tr>
<th>Variables</th>
<th>No (n = 8)</th>
<th>Yes (n = 4)</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decreased brain volume</td>
<td>6 (75.0%)</td>
<td>4 (100.0%)</td>
<td>.515</td>
</tr>
<tr>
<td>Grade of brain volume decrease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>1 (20.0%)</td>
<td>0 (0.0%)</td>
<td>.876</td>
</tr>
<tr>
<td>Moderate/Severe</td>
<td>4 (80.0%)</td>
<td>3 (100.0%)</td>
<td></td>
</tr>
<tr>
<td>Degree of cerebral damage</td>
<td>1 (12.5%)</td>
<td>0 (0.0%)</td>
<td>.394</td>
</tr>
<tr>
<td>Absent</td>
<td>2 (25.0%)</td>
<td>0 (0.0%)</td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>2 (25.0%)</td>
<td>0 (0.0%)</td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>3 (37.5%)</td>
<td>4 (100.0%)</td>
<td></td>
</tr>
<tr>
<td>Symmetry</td>
<td>6 (75.0%)</td>
<td>3 (75.0%)</td>
<td>&gt; .999</td>
</tr>
<tr>
<td>Cortical development abnormalities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pachygyria</td>
<td>3 (37.5%)</td>
<td>3 (75.0%)</td>
<td>.545</td>
</tr>
<tr>
<td>Polymicrogyria</td>
<td>2 (25.0%)</td>
<td>0 (0.0%)</td>
<td>.515</td>
</tr>
<tr>
<td>Simplified gyral pattern</td>
<td>4 (50.0%)</td>
<td>3 (75.0%)</td>
<td>.576</td>
</tr>
<tr>
<td>Corpus callosum</td>
<td></td>
<td></td>
<td>&gt; .999</td>
</tr>
<tr>
<td>Normal</td>
<td>2 (25.0%)</td>
<td>0 (0.0%)</td>
<td></td>
</tr>
<tr>
<td>Hypoplasia</td>
<td>5 (62.5%)</td>
<td>3 (75.0%)</td>
<td></td>
</tr>
<tr>
<td>Hypogenesis</td>
<td>1 (12.5%)</td>
<td>1 (25.0%)</td>
<td></td>
</tr>
<tr>
<td>Cortical and subcortical junction calcifications</td>
<td>7 (85.5%)</td>
<td>4 (100.0%)</td>
<td>&gt; .999</td>
</tr>
<tr>
<td>Basal ganglia calcifications</td>
<td>3 (37.5%)</td>
<td>3 (75.0%)</td>
<td>.545</td>
</tr>
<tr>
<td>Periventricular calcifications</td>
<td>0 (0.0%)</td>
<td>3 (75.0%)</td>
<td>.018</td>
</tr>
<tr>
<td>Brain stem calcifications</td>
<td>1 (12.5%)</td>
<td>3 (75.0%)</td>
<td>.067</td>
</tr>
<tr>
<td>Cerebellum calcifications</td>
<td>0 (0.0%)</td>
<td>2 (50.0%)</td>
<td>.091</td>
</tr>
<tr>
<td>Cerebellum or brain stem hypoplasia</td>
<td>2 (25.0%)</td>
<td>4 (100.0%)</td>
<td>.061</td>
</tr>
<tr>
<td>Increased cisterna magna</td>
<td>8 (100.0%)</td>
<td>4 (100.0%)</td>
<td></td>
</tr>
<tr>
<td>Delayed myelination</td>
<td>5 (50.0%)</td>
<td>4 (100.0%)</td>
<td>.208</td>
</tr>
</tbody>
</table>

* Data are number of patients (%).
* *P* = Fisher Exact test.

---

FIG 5. MR imaging of the brain and the spinal cord of an infant with microcephaly confirmed to be caused by the Zika virus without arthrogryposis. Sagittal T2-weighted images (A) shows hypogenesis of the corpus callosum (white arrow) and an enlarged cisterna magna (black arrow). Coronal T2-weighted image (B) shows left cerebellar hemisphere hypoplasia, with cortical malformation and microcysts (white arrows). Axial SWI (C) shows small dystrophic calcifications in the junction between the cortical and subcortical white matter and in the basal ganglia (black arrows). Axial T2-weighted image (D) shows a simplified gyral pattern, bilateral cortical thickness in the pachygyric frontal lobe (white arrows), and ventriculomegaly (black arrows). The spinal cord and conus medullaris are normal-sized and show no abnormal signal on the sagittal T2-weighted volumetric GRE (E). Axial reformatted T2-weighted volumetric GRE reveals normal-sized anterior and posterior nerve roots in the conus medullaris (F and G) and cauda equina (H).
spinal cord thickness, especially at the thoracic segment, and mildly reduced anterior nerve roots at the conus medullaris. Therefore, we can suppose that there are, currently unidentified in the normal population of the epidemic area, more children with a mild degree of damage not only in the brain but also in the spinal cord. The identified cases could correspond to only the “tip of the iceberg,” represented by microcephaly and arthrogryposis, of the congenital Zika syndrome.

CONCLUSIONS

Most of the infants with congenital Zika syndrome had some degree of spinal cord thickness reduction, which is predominant in the thoracic segment in cases without arthrogryposis and in the entire spinal cord in cases with arthrogryposis. In addition, there is thickness reduction of anterior nerve roots of the conus medullaris in both groups, being more severe in infants with arthrogryposis. With regard to brain lesions, periventricular calcifications were more frequent in infants with arthrogryposis. Although without statistical significance, the prominence of the anterior median fissure of the spinal cord was found only in infants with arthrogryposis; brain stem hypoplasia was present in all infants with arthrogryposis; brain stem and cerebellum calcifications were more frequent; and polymicrogyria was absent in this group.

ACKNOWLEDGMENTS

We are especially grateful to the anesthesiology staff, Drs Antonio Monteiro and Cristovam A. de Lira Terceiro; the neuropediatric group. We are especially grateful to the anesthesiology staff, Drs Antonio Monteiro and Cristovam A. de Lira Terceiro; the neuropediatrician, Dr Ana van der Linden; the medical student, Julia Sales Machado; the MR imaging technician, Edineide Cristina Leite Lopes; and the Centro Diagnostico Multimagem for their help.

REFERENCES


Intraforaminal Location of Thoracolumbar Radicular Arteries Providing an Anterior Radiculomedullary Artery Using Flat Panel Catheter Angiotomography

L. Gregg, D.E. Sorte, and P. Gailloud

ABSTRACT

BACKGROUND AND PURPOSE: Flat panel catheter angiotomography performed during the selective injection of intersegmental arteries offers a multiplanar assessment of the intraforaminal course of the radicular arteries providing an anterior radiculomedullary artery. Injury of anterior radiculomedullary arteries during transforaminal epidural steroid injections can result in spinal cord damage. Evaluations of the intraforaminal location of these arteries have so far been limited to anteroposterior views or the examination of cadaveric material. This study documents the in vivo intraforaminal location of thoracolumbar arteries providing an anterior radiculomedullary artery with flat panel catheter angiotomography.

MATERIALS AND METHODS: Ninety-four flat panel catheter angiotomography acquisitions obtained during the selective injection of intersegmental arteries providing an anterior radiculomedullary artery were reviewed. Measurements obtained from sagittal reconstructions were converted into a scatterplot visualization. Patients’ age, sex, and side and level of the injection were recorded.

RESULTS: The location of radicular arteries could be ascertained in 78 of 94 flat panel catheter angiotomography acquisitions (33 women and 45 men, 22–82 years of age). Fifty-three acquisitions (67.9%) were on the left side, and 25 (32.1%), on the right, between T2 and L3. The arteries were found in the anterosuperior quadrant of the neural foramen in 75 cases (96.2%), in the posterosuperior quadrant in 2 cases (2.6%), and in the anteroinferior quadrant in 1 case (1.3%). None were located in the posteroinferior quadrant. No differences in location were observed with age, sex, side of injection, or vertebral level.

CONCLUSIONS: Avoiding needle placement in the superior half of the neural foramen, specifically the anterosuperior quadrant, can reduce the risk of spinal cord injury during transforaminal epidural steroid injection.

ABBREVIATIONS: ARMA = anterior radiculomedullary artery; FPCA = flat panel catheter angiotomography; ISA = intersegmental arteries; NF = neural foramen; TFESI = transforaminal epidural steroid injection

Flat panel catheter angiotomography (FPCA) is a recently developed angiographic technique in which rotational datasets are used to generate high-resolution multiplanar reconstructions. FPCA performed during the injection of the thoracolumbar intersegmental arteries (ISA) allows documenting with precision the intraforaminal course of radicular branches providing an anterior radiculomedullary artery (ARMA). The location of these vessels within the neural foramen (NF) is relevant to needle placement during transforaminal epidural steroid injection (TFESI), a commonly performed image-guided procedure consisting of the injection of corticosteroids at the NF. Inadvertent injection or injury of branches contributing to the supply of the anterior spinal artery during TFESI can result in spinal cord damage and can lead to paralysis and death. Although several cases of paralysis after TFESI have now been published, such complications are likely under-reported.

Targeting the “safe triangle,” an area located between the pedicle and the nerve root, has been recommended to avoid nerve injury during needle placement (Fig 1). However, this approach does not take into account the position of branches providing an ARMA. Evaluations of the intraforaminal location of these arteries critical to the spinal cord vascularization have so far been limited to anteroposterior views only or to the examination of ca-
MATERIALS AND METHODS

FPCA studies performed during routine spinal angiography between November 2009 and September 2015 were collected retrospectively from an institutional review board–approved clinical data base. Additional approval for publication was granted. Ninety-four datasets obtained during the selective injection of the thoracolumbar ISAs providing an ARMA were reviewed. Acquisitions documenting a radicular branch and the ARMA course within the NF were included, while those with pathologies that obscured or distorted the vessel of interest were excluded. The patient’s age, sex, and side and level of injection were recorded.

Datasets generated by a 20-second nonsubtracted rotational acquisition (DynaCT, Artis Zee; Siemens, Erlangen, Germany) were reconstructed on a dedicated workstation (0.13- to 0.43-mm voxel size, depending on the size of the FOV selected for reconstruction) and were studied with MIP rendering. Measurements were obtained by using commercially available software (OsiriX 64-bit Imaging Software; http://www.osiriX-viewer.com). In a sagittal view of the injected ISA, the working projection was adjusted medially to follow the radicular artery and/or ARMA until the osseous borders of the NF were in view (On-line Video) (Fig 2). Measurements included the NF height, defined as the longest distance between the superior and inferior osseous NF borders (Fig 2C, B to B’), and the NF width, measured perpendicular to and at the midpoint of the height line (Fig 2C, C to C’). The NF was then divided into quadrants based on the divisions created by the height and width measurement lines (Fig 2C). When osteophytes were present, the NF width was measured from the estimated original NF boundaries.

The intraforaminal location of branches providing an ARMA was evaluated by measuring the distance from the center of the artery to the inferior edge of the corresponding pedicle (Fig 2C, D), the posterior edge of the vertebral body (Fig 2C, E), the horizontal NF division line (Fig 2C, F), and the vertical NF division line (Fig 2C, G). These data were used to generate a scatterplot representation. The NF height measurement was used to normalize the measurement data, thus providing a comparable point of reference for ARMAs located in NFs of various sizes.

RESULTS

The intraforaminal location of branches providing an ARMA could be ascertained in 78 of the 94 reviewed FPCA acquisitions (83.0%), including 33 women (42.3%) and 45 men (57.7%) with an average age of 53 years (range, 22–82 years). Precise analysis of the ARMA course was prevented by the presence of a vascular malformation (2 cases), technical artifacts (4 cases), anatomic distortion by a tumor or prior surgery (4 cases), or incomplete documentation of the NF (6 cases).

Of the 78 FPCA acquisitions included in the study, 53 (67.9%) were on the left side and 25 (32.1%) were on the right. Acquisitions included vertebral levels extending from T2 to L3 (Table 1). The branches providing an ARMA were found in the anterosuperior quadrant of the NF in 75 cases (96.2%), in the posterosuperior quadrant in 2 cases (2.6%), and in the anteroinferior quadrant in 1 case (1.3%) (Fig 3). Branches providing an ARMA were not found in the posteroinferior quadrant.

Most arteries providing an ARMA were close to the bone near the subpedicular notch, at the inferior border of the junction of the pedicle and the vertebral body. The center point of the arteries in 51 cases (65.4%) was located ≤2 mm from the osseous borders of the anterosuperior quadrant of the NF.

The scatterplot of the raw data (Fig 3A) was comparable with that of the normalized data (Fig 3B), both of which documented a close association of branches providing an ARMA with the anterosuperior quadrant of the NF near the junction of the pedicle and the vertebral body.

The mean caliber of arteries providing an ARMA was 1 mm (range, 0.6–1.7 mm). A summary of the NF and artery caliber measurements is shown in Table 2. No differences in artery location were observed with age, sex, side of injection, or vertebral level. Illustrative cases are shown in Figs 4 and 5.

DISCUSSION

Methodologic Considerations

To assess the location of the branches providing an ARMA in a clinically relevant manner, our measurements were performed in a sagittal reconstruction of the FPCA datasets matching the fluoroscopic landmarks used during TFESI (ie, a lateral projection perpendicular to the entrance of the NF).18 The vessel measured can be considered as either a radicular artery or an ARMA because the exact point of transition from one to the other remains difficult to assess by imaging and nomenclatures show slight variations.11,12 All the branches investigated for this study continued intradurally as contributors to the anterior spinal arterial axis.

The NF variability in size between patients and vertebral levels was accounted for by adopting the NF height as the normalization...
FIG 2. Depiction of the arterial anatomy of the NF showing the measurements used in this study. A, Illustration of ISA anatomy including the principal supply to the thoracolumbar spinal cord from an anterior oblique view. The ISA originates from the aorta and divides into medial, dorsal, and lateral branches. A complete spinal branch is shown entering the NF at the left L1 vertebral level and providing a retrocorpororeal artery, a prelaminal artery, and a radicular artery; the latter crosses the dura to continue as an ARMA and anastomoses with the anterior spinal artery. The shaded regions and quadrant grids depicted at the L1 NF clarify the working projection and measurement definitions depicted in B and C. B, FPCA, left L1 ISA injection, sagittal reconstruction (thickness = 0.2 mm) demonstrates the working projection used to take measurements. The image documents the location of an ARMA within the NF (red arrow) in a 59-year-old women investigated for acute onset paraplegia. Spinal angiography and FPCA revealed severe stenosis of the left proximal L1 ISA, resulting in an episode of spinal cord ischemia. Measurements were performed with the anterior surface to the left side. C, Measurements included the NF height (B to B line), which was defined as the long axis from the posterior edge of the vertebral pedicle to the superior edge of the pedicle below. The NF was divided into quadrants by measuring the width (C to C line) at the midpoint of the height. The ARMA location was evaluated by measuring the distance from the center of the artery to the superior (D) and anterior (E) walls of the NF and the vertical (F) and horizontal (G) distances to the quadrant divisions. ASQ indicates anterosuperior quadrant; PSQ, posterosuperior quadrant; PIQ, posteroinferior quadrant; AIQ, anteroinferior quadrant. Copyright 2016 Lydia Gregg.

Table 1: Vertebral level and side of included FPCA acquisitions documenting a radicular branch and the ARMA course within the NF

<table>
<thead>
<tr>
<th>Vertebral Level</th>
<th>L Side (No.)</th>
<th>L Side</th>
<th>R Side (No.)</th>
<th>R Side</th>
<th>L and R Sides, Total (No.)</th>
<th>L and R Sides, Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2</td>
<td>1</td>
<td>1.3%</td>
<td>–</td>
<td>–</td>
<td>1</td>
<td>1.3%</td>
</tr>
<tr>
<td>T3</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>T4</td>
<td>–</td>
<td>–</td>
<td>1</td>
<td>1.3%</td>
<td>1</td>
<td>1.3%</td>
</tr>
<tr>
<td>T5</td>
<td>1</td>
<td>1.3%</td>
<td>–</td>
<td>–</td>
<td>2</td>
<td>2.6%</td>
</tr>
<tr>
<td>T6</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>T7</td>
<td>4</td>
<td>5.1%</td>
<td>–</td>
<td>–</td>
<td>4</td>
<td>5.1%</td>
</tr>
<tr>
<td>T8</td>
<td>5</td>
<td>6.4%</td>
<td>3</td>
<td>3.8%</td>
<td>8</td>
<td>10.3%</td>
</tr>
<tr>
<td>T9</td>
<td>10</td>
<td>12.8%</td>
<td>5</td>
<td>6.4%</td>
<td>15</td>
<td>19.2%</td>
</tr>
<tr>
<td>T10</td>
<td>2</td>
<td>2.6%</td>
<td>5</td>
<td>6.4%</td>
<td>7</td>
<td>9.0%</td>
</tr>
<tr>
<td>T11</td>
<td>6</td>
<td>7.7%</td>
<td>1</td>
<td>1.3%</td>
<td>7</td>
<td>9.0%</td>
</tr>
<tr>
<td>T12</td>
<td>7</td>
<td>9.0%</td>
<td>3</td>
<td>3.8%</td>
<td>10</td>
<td>12.8%</td>
</tr>
<tr>
<td>L1</td>
<td>14</td>
<td>17.9%</td>
<td>5</td>
<td>6.4%</td>
<td>19</td>
<td>24.4%</td>
</tr>
<tr>
<td>L2</td>
<td>2</td>
<td>2.6%</td>
<td>1</td>
<td>1.3%</td>
<td>3</td>
<td>3.8%</td>
</tr>
<tr>
<td>L3</td>
<td>1</td>
<td>1.3%</td>
<td>–</td>
<td>–</td>
<td>1</td>
<td>1.3%</td>
</tr>
<tr>
<td>Total</td>
<td>53</td>
<td>67.9%</td>
<td>25</td>
<td>32.1%</td>
<td>78</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Note: — L indicates left; R, right.
standard, considering, in particular, the relative paucity of osteophytes that form on the superior and inferior borders of the vertebral pedicles.\textsuperscript{13} No major differences in artery locations were noted between the raw measurement data and the normalized data (Fig 3).

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
Vertebral Level & No. & Measurement & Mean & Range & SD  \\
\hline
\textbf{Thoracic ARMAs} & 55 & NF height & 16.2 & 9.7–21.2 & 2.4  \\
& & NF width & 8.0 & 5.4–12.7 & 1.8  \\
& & ARMA diameter & 1.0 & 0.6–1.4 & 0.2  \\
\textbf{Lumbar ARMAs} & 23 & NF height & 20.6 & 14.8–25 & 2.3  \\
& & NF width & 10.3 & 8.1–13.6 & 1.4  \\
& & ARMA diameter & 1.1 & 0.7–1.7 & 0.3  \\
\textbf{All ARMAs} & 78 & NF height & 17.5 & 9.7–25 & 3.1  \\
& & NF width & 8.7 & 5.4–13.6 & 2.0  \\
& & ARMA diameter & 1.0 & 0.6–1.7 & 0.2  \\
\hline
\end{tabular}
\caption{NF and radicular branches providing ARMA lumen diameter measurements\textsuperscript{a}}
\end{table}

\textsuperscript{a} Data are in millimeters.

\textit{Intraforaminal Location of Branches Providing an ARMA}
The ISA consists of an aortic stem that divides into medial (or spinal), dorsal, and lateral branches; depending on the level considered, the lateral branch takes the name of posterior intercostal, subcostal, or lumbar artery (Fig 2A). The medial branch, when complete, enters the neural foramen and provides the retrocorporal and prelaminar arteries, which contribute to the vertebral vascularization and the radicular artery.\textsuperscript{14} A radicular artery that supplies the spinal cord crosses the dura to continue as an anterior or posterior radiculomedullary artery or both, which, respectively, contribute to the anterior or posterior spinal arterial chains. Supply to the anterior thoracolumbar spinal cord is typi-
cally limited to a dominant thoracolumbar ARMA, the artery of Adamkiewicz, and a smaller upper thoracic ARMA, the artery of von Haller.15 Reports of paralysis resulting from vascular injuries during TFESI spurred previous studies on the location of these vessels within the NF. The investigation of Murthy et al,7 based on anteroposterior views only, concluded that the radicular segment most often crosses the superior aspect of the NF. Two cadaveric studies reported that branches providing an ARMA are generally located anterosuperior to the dorsal nerve root within the NF.8,9 Our multiplanar analysis of high-resolution angiographic data-sets acquired during routine clinical practice confirms and extends these observations by plotting the precise distribution of branches providing an ARMA within the NF (Fig 3).

The caliber of radicular branches that provide an ARMA measured in our study (average, 1.0 mm; range, 0.6–1.7 mm) was comparable with that in prior investigations, which have reported diameters ranging from 1.2 to 2.5 mm (average, 1.9 mm)8 and from 0.8 to 1.9 mm (average, 1.2 mm)9 (measurements obtained in anatomic specimens, including the vessel wall thickness).

**Clinical Implications for Arterial Injury**

Despite being associated with severe complications related to the intraforaminal location of radicular branches that provide an ARMA,3 TFESI has seen its rate of use steadily increase due to its purported superiority over alternative techniques, such as the caudal or interlaminar approaches.1 Spinal cord ischemia and stroke occurring during TFESI may result from arterial embolization with air or particulate steroids1 or from direct vessel

FIG 4. FPCA, left L2 ISA injection. This case illustrates the intraforaminal course of an ARMA in a 42-year-old woman investigated for progressive myelopathy. Spinal angiography and FPCA findings were unremarkable. A, Sagittal MIP reconstruction (thickness = 0.2 mm) documents the location of an ARMA (white arrow) within the NF. The retrocorporal artery (gray arrow), L2 (white asterisk) and L3 pedicles (black asterisk), and portions of the internal vertebral venous plexus (black arrowheads) are also visible. B, The same sagittal MIP reconstruction with graphics indicates recorded measurements, including the NF height (B to B'), NF width (C to C'), distance of the ARMA to the inferior border of the L2 pedicle (D), the posterior wall of the L2 vertebral body (E), and the vertical (F) and horizontal (G) distances to the quadrant divisions. C, Coronal-oblique MIP reconstruction (thickness = 1.7 mm) documents the same ARMA (white arrows) within the NF and its anastomosis with the anterior spinal artery (black arrows). The L2 (white asterisk) and L3 pedicles (black asterisk) and a portion of the internal vertebral venous plexus (black arrowhead) are also visible. D, Axial MIP reconstruction (thickness = 9.3 mm) documents the same L2 ARMA (white arrows), the anterior spinal artery (black arrow), portions of the internal vertebral venous plexus (black arrowheads), and the dorsal branch of the L2 ISA (white arrowheads).

FIG 5. FPCA, left L1 ISA injection, in a 77-year-old man with severe spinal canal stenosis investigated for myelopathy. Spinal angiography and FPCA findings were normal. This case illustrates how NF width estimates were derived from abnormal foramina. A, Sagittal MIP reconstruction (thickness = 0.2 mm) documents the location of a branch providing an ARMA (white arrow) within the NF. An osteophyte (gray asterisk) on the L1 vertebral body distorts the width of the NF between the L1 (white asterisk) and L2 pedicles (black asterisk), while the NF height appears unaffected. B, The same sagittal MIP reconstruction with graphics indicating recorded measurements includes the following: the NF height (B to B'), estimated NF width (C to C'), distance of the artery to the inferior border of the L1 pedicle (D), the estimated posterior wall of the L1 vertebral body (E), and the vertical (F) and horizontal (G) distances to the quadrant divisions. C, Coronal-oblique MIP reconstruction (thickness = 1.8 mm) documents the same branch providing an ARMA (white arrows) within the NF and its anastomoses with the anterior spinal artery (black arrow). The L1 (white asterisk) pedicle and a dilated anterior median spinal vein (black arrowheads) are also visible. D, Axial MIP reconstruction (thickness = 9.0 mm) documents the same branch providing an ARMA (white arrows) and dorsal branches of the L1 ISA (white arrowheads).
injuries such as spasm, laceration, and dissection. The close association between branches providing an ARMA and the subpedicular notch documented in our study emphasizes the danger in making contact between the needle and the pedicle, as is typical with the subpedicular approach. Multiple attempts at repositioning the needle tip, as reported in several cases of paralysis, likely increase the risk of vessel damage.

The Role of Particulate Steroids
All 19 observations of paralysis reported so far have involved the injection of particulate steroids during TFESI. With particle sizes ranging from 1 to 100 μm, an intra-arterial injection could result in the embolization of spinal arterioles. Direct arterial injection of particulate steroids has caused permanent neurologic damage in animal models, while nonparticulate steroids such as dexamethasone and prednisolone produced no injuries. Although it has been suggested that particulate steroids are more effective than nonparticulate steroids, no significant difference in efficacy after lumbar TFESI has been reported. It is possible that the small number of reported complications and the widespread use of particulate steroids falsely suggest a causative relationship between particulate injectate and the risk of paralysis. Nonetheless, dexamethasone became the recommended injectate for TFESI in 2010.

Posterosilateral Approach
Positioning the needle in the region of the NF known as the Kambin triangle has been recommended as a safer alternative to the subpedicular approach. In this posterior triangle or posterolateral approach, the targeted area is defined by the posterior margin of the exiting nerve root, the endplate of the caudal vertebral level, and the articular facet of the cranial vertebral level. This approach was proposed for arthroscopic discectomy, which requires the transforaminal placement of large-caliber instrumentation. TFESI via the posterolateral approach has shown pain reduction scores similar to the those with the subpedicular approach, while decreasing both the amount of periprocedural pain and the risk of nerve damage. Considering that the posterolateral approach targets the posteroinferior quadrant of the NF, where branches providing an ARMA are least likely encountered, and that it has not yet been associated with reports of paralysis suggests that it should be the preferred technique for TFESI.

Misconceptions Regarding Spinal Vascular Anatomy and the “Safe Triangle”
Several misconceptions have led to a slow rejection of the subpedicular approach, notably the incorrect notions that the “safe triangle” helps prevent vascular injuries in addition to nerve damage and that arterial opacification is reliably detected during contrast injections performed before steroid instillation. Intra-arterial needle tip placement during TFESI can, in fact, not be excluded with certainty by using intermittent fluoroscopy, aspiration, contrast injection, or even digital subtraction angiography. In addition, none of these imaging methods would help avoid other injury mechanisms such as arterial transection or dissection. A limited understanding of spinal vascular anatomy, which is often overlooked or misrepresented, also plays a role.

Simon et al reported the presence of arteries both in the subpedicular region and the posterolateral triangle on contrast-enhanced CT scans, but the vessels observed with that technique cannot be clearly identified as branches providing an ARMA or osteonegeal branches or even venous structures such as radiculomedullary veins or the emissary veins linking the internal and external venous plexuses at each vertebral level. Figures 4 and 5 illustrate the advantage offered by a high-resolution imaging technique to reliably distinguish such minute vascular elements on the basis of their morphology (ie, termination of the radiculomedullary veins into the epidural plexus) or opacification pattern. A similar lack of specificity is found in a report of apparent retrograde flow into a T6 spinal artery noted during a postero lateral approach, which prompted Simon et al to suggest that the injection of any artery near the NF carries potential clinical consequences. However, the images documenting that observation only show the opacification of the external epidural venous plexus with retrograde filling of a basivertebral and intrasosseous venous system. An analysis of reports of paralysis following TFESI in which the needle position was mentioned shows that spinal cord damage only occurred in association with the subpedicular approach, a fact weakening the role of retrograde injections and the notion that any artery in the vicinity of the NF has clinical consequences comparable with those supplying ARMAs when performing TFESI.

Study Limitations
The goal of our study was to describe the typical location of radicular branches providing an ARMA in the NF based on FPCA datasets acquired in patients principally investigated for myelopathy. This patient group may have a lower incidence of degenerative osteodiscal pathology than the population typically treated with TFESI, leading to a more stable distribution of contributors of ARMAs within the NF.

CONCLUSIONS
Thoracolumbar radicular branches providing an ARMA were located in the anterosuperior quadrant of the NF in 75 of 78 patients (96.2%), suggesting that needle placement in that area should be avoided during TFESI. Unlike the subpedicular approach, the posterolateral approach allows placing the needle tip away from the documented position of ARMA contributors within the NF, reducing the risk of intra-arterial injection or injury to the spinal vascularization.

Quantitative Measurement of CSF in Patients with Spontaneous Intracranial Hypotension


ABSTRACT

BACKGROUND AND PURPOSE: CSF hypovolemia is a core feature of spontaneous intracranial hypotension. Spontaneous intracranial hypotension is characterized by orthostatic headache and radiologic manifestations, including CSF along the neural sleeves, diffuse pachymeningeal enhancement, and/or venous engorgement. However, these characteristics are only qualitative. Quantifying intraspinal CSF volumes could improve spontaneous intracranial hypotension diagnosis and evaluation of hypovolemic statuses in patients with spontaneous intracranial hypotension. The purpose of this study was to compare intraspinal CSF volumes across spontaneous intracranial hypotension stages and to test the clinical applicability of these measures.

MATERIALS AND METHODS: A cohort of 23 patients with spontaneous intracranial hypotension and 32 healthy controls was subjected to brain MR imaging and MR myelography with 1.5T imaging. An automatic threshold-based segmentation method was used to calculate intraspinal CSF volumes at initial hospitalization (spontaneous intracranial hypotension-initial), partial improvement (spontaneous intracranial hypotension-intermediate), and complete recovery (spontaneous intracranial hypotension-recovery) stages.

RESULTS: The mean intraspinal CSF volumes observed were the following: 95.31 mL for healthy controls, 72.31 mL for spontaneous intracranial hypotension-initial, 81.15 mL for spontaneous intracranial hypotension-intermediate, and 93.74 mL for spontaneous intracranial hypotension-recovery. Increased intraspinal CSF volumes were related to disease recovery \((P < .001)\). The intraspinal CSF volumes of patients before complete recovery were significantly lower than those of healthy controls. With the estimated intradural CSF volumes as a reference, the intraspinal CSF volume percentage was lower in patients with spontaneous intracranial hypotension with venous engorgement than in those without it \((P = .058)\).

CONCLUSIONS: With a threshold-based segmentation method, we found that spinal CSF hypovolemia is fundamentally related to spontaneous intracranial hypotension. Intraspinal CSF volumes could be a sensitive parameter for the evaluation of treatment response and follow-up monitoring in patients with spontaneous intracranial hypotension.

ABBREVIATIONS: BH = body height; BMI = body mass index; BW = body weight; 3D-SPACE = 3D sampling perfection with application-optimized contrasts by using different flip angle evolution; ECSF = estimated intradural CSF volume; DPE = diffuse pachymeningeal enhancement; HC = healthy control; MRM = MR myelography; SIH = spontaneous intracranial hypotension

Spontaneous intracranial hypotension (SIH) has been diagnosed increasingly as an important cause of headaches. Orthostatic headache is the characteristic symptom of SIH caused by CSF leakage in the spine. Despite the term “SIH,” CSF hypovolemia may be the core of the condition. About 18% of patients with symptoms typical of SIH have normal CSF opening pressure. The diagnosis of SIH depends on various imaging characteristics and clinical symptoms. The commonly reported findings indicating CSF leakage are epidural fluid accumulation with a dilated epidural venous plexus, a triangular-shaped expansion of the neural sleeve, and an irregular linear signal lateral to the neural sleeve on spinal MR imaging or MR myelography (MRM). Conventional brain MR imaging has demonstrated diffuse pachymeningeal enhancement (DPE), venous engorgement, and pituitary hyperemia in patients with SIH. However, DPE and other brain MR imaging abnormalities are not observed in all patients with SIH, especially in early-stage SIH.

As the amount of CSF decreases, the buoyant force from in-
tracranial CSF is reduced and the brain descends while patients are standing. This results in stretching of pain-sensitive structures on the brain surface and causes orthostatic headaches. However, CSF hypovolemia has only been linked indirectly as a causative factor of SIH in the literature.1 Researchers have used imaging signs to determine the severity of CSF hypovolemia, such as the number of vertebral bodies over which leaked CSF had spread in spine MR imaging and DPE in brain MR imaging.1 Therefore, in addition to morphologic alterations in the brain and spine, we hypothesized that quantifying intraspinal CSF volumes would be crucial to evaluating the hypovolemic status of patients with SIH. This important issue of direct monitoring of CSF volume depletion has not been discussed in the SIH literature, to our knowledge.

The aim of this study was to apply an automatic segmentation method to measure the entire CSF volume in the spine and detect CSF hypovolemia in patients with SIH. Furthermore, we compared intraspinal CSF volumes across SIH stages and tested the clinical applicability of these measures.

**MATERIALS AND METHODS**

**Subjects**

Between January 2012 and February 2016, 23 patients diagnosed with SIH (according to the criteria of Schievink et al7 or the “International Classification of Headache Disorders,” 3rd ed8) and treated in the neurology department at our hospital (Taichung Veterans General Hospital, Taichung, Taiwan) were enrolled in our study. During their hospital stays, patients received conservative treatment and a targeted epidural blood patch. Whole-spine MRM and conventional brain MR imaging were performed on patients simultaneously at 3 different time points. All patients underwent MR imaging on presentation/enrollment that was defined as SIH-initial. Some patients received a second MR imaging for an additional epidural blood patch when their clinical symptoms improved but did not resolve completely. These measurements are the SIH-intermediate. The mean time of MR imaging for SIH-intermediate measurement was 8.41 ± 4.34 days. If complete recovery was achieved following treatment, patients received another MR imaging, the SIH-recovery measurement. Complete recovery was defined as patients having no more headaches or other associated clinical orthostatic symptoms and no CSF leakage or other abnormalities detected in MRM or MR imaging. The mean time of MR imaging for SIH-recovery was 104.67 ± 94.00 days.

Thirty-two healthy volunteers with no known neurologic disorders, spine malformations, or previous spine operations were recruited as healthy controls (HCs). Whole-spine MRM was performed for each subject, and the only pathologic finding acceptable for inclusion was minor degenerative changes of the spine. Before enrollment, written informed consent was obtained from all subjects. This study was approved by the institutional review board of Taichung Veterans General Hospital.

**MR Imaging Acquisition**

Subjects underwent MRM and brain MR imaging on a 1.5T MR imaging scanner (Magnetom Aera; Siemens, Erlangen, Germany). Whole-spine MRM was performed with a 3D sampling perfection with application-optimized contrasts by using different flip angle evolution (3D-SPACE) sequence (Siemens). The MR imaging parameters were as follows: TR = 3000 ms, TE = 560 ms, isotropic voxel size = 0.9 mm³, matrix size = 320 × 320 pixels, and FOV = 200 mm. Fat suppression and a generalized autocalibrating partially parallel acquisition imaging reconstruction with an acceleration factor of 2 were used. Images were acquired volumetrically in the coronal plane of the cervical-to-thoracic and thoracic-to-lumbar regions of the spine, parallel to the spinal curve with some overlapping. The acquisition time was 7 minutes 38 seconds for each segment. Conventional brain MR imaging included axial spin-echo T1-weighted images (TR/TE, 500/10 ms) (2 minutes 12 seconds); axial fast-spin-echo T2-weighted images (TR/TE, 3200/115 ms) (2 minutes 39 seconds); and gadolinium-enhanced spine echo T1WI in the axial (2 minutes 25 seconds), sagittal (2 minutes 10 seconds), and coronal planes (1 minute 56 seconds). The contrast-enhanced study was performed by using 1.0 mol/L of gadolinium chelate (gadobutrol, Gadovist 1.0; Bayer Schering Pharma, Berlin, Germany) with 0.1 mmol per kilogram of body weight.

**Intraspinal CSF Volume Measurement and Data Segmentation**

Before segmentation, a preprocessing of bias field correction was performed to reduce intensity inhomogeneity in 2 sets of 3D-SPACE whole-spine MRM studies. The threshold-based algorithm of relative entropy has been proposed for use in clinical applications to enable reproducible and facile image segmentation.9-12 With the advantages of selecting an optimal threshold value with local neighborhood information, the method could extract hyperintense CSF voxels from other hypointense background tissue automatically and robustly in 3D-SPACE whole-spine MRM. After processing, we calculated CSF volumes by identifying all CSF-specific voxels from the tip of the odontoid process to the end of the dural sac. The volumes obtained were defined as intraspinal CSF volumes, including intra- and extradural CSF in the spinal canal at SIH-initial and SIH-intermediate measurements. For SIH-recovery measurements and HCs, intraspinal CSF volumes would be equal to intradural CSF because no CSF leakage occurred. Intraspinal CSF volumes were evaluated by using absolute data. We also estimated each subject’s intradural CSF volumes by correlated parameters. The intraspinal CSF volume percentage was defined as the following: Intraspinal CSF Volumes of SIH–Initial / Estimated Intradural CSF Volumes.

**Brain Imaging Measurements**

Some qualitative indications in brain images were evaluated, including the following: pituitary volumes (in cubic centimeters) defined as the half-value of the product of the gland height, length, and width; venous engorgement, which appears as a convex upper surface of the sinus confluence on a sagittal view of postcontrast T1WI; and any DPE, defined as continuously enhanced dural matter.

**Statistical Analyses**

All data were analyzed in SPSS software (Version 18; IBM, Armonk, New York). Demographic data, intraspinal CSF volumes,
and brain MR imaging findings were compared between HCs and the SIH-initial, SIH-intermediate, and SIH-recovery measurements of patients. We used Kolmogorov-Smirnov tests to verify normal distribution of continuous variables. The age, body mass index (BMI), body height (BH), body weight (BW), and intraspinal CSF volume data for subjects with SIH and HCs were not normally distributed and were therefore analyzed with nonparametric Mann-Whitney tests. The sexes between patients with SIH and HCs were analyzed by using a 2-sample t test. The difference among SIH-initial, SIH-intermediate, and SIH-recovery measurements was analyzed with the paired t test. The relationships between intraspinal CSF volumes of the HCs and SIH-recovery measurements with age, BMI, BH, BW, and sex were subjected to the Spearman rank-based correlation coefficient. For estimated intradural CSF volumes (ECSF), we used linear regression to fit parameters that correlated with the intraspinal CSF volumes in SIH-recovery measurements of patients and HCs. All tests were 2-tailed with statistical significance at $P < .05$.

**RESULTS**

Patient demographics and baseline characteristics are summarized in Table 1. There were no statistical differences in age, sex, BMI, BH, or BW between the patients with SIH and HCs. Most of our patients had a significant amount CSF leakage at the spinal nerve roots with epidural fluid accumulation. The extent of CSF leakage for each patient is summarized in the On-line Table. All our patients had a significant amount CSF leakage at the spinal CSF Volumes

The mean intraspinal CSF volumes obtained for the SIH-initial (72.31 mL) and SIH-intermediate (81.15 mL) stages differed significantly from those of the HCs (95.31 mL; $P < .001$ and $P < .05$, respectively). However, the mean intraspinal CSF volumes obtained for the SIH-recovery stage (93.74 mL) were similar to those of the HCs ($P = .731$). Intraspinal CSF volumes increased stepwise ($P < .001$) in patients with SIH from SIH-initial to SIH-intermediate to SIH-recovery.

**Estimated Intradural CSF Volumes**

Intraspinal CSF volumes of HCs and SIH-recovery measurements correlated moderately well and significantly with BH (correlation coefficient $= 0.444$, $P = .002$), but not with BMI (correlation coefficient $= -0.033$, $P = .822$), BW (correlation coefficient $= 0.073$, $P = .622$), age (correlation coefficient $= -0.215$, $P = .134$), or sex (correlation coefficient $= -0.233$, $P = .164$). Because intradural CSF volumes varied across individuals, we estimated each subject’s intradural CSF volumes by using a linear regression equation according to BH to see whether it could provide useful clinical information for diagnosis and treatment (Fig 1).

Relative to ECSF values, intraspinal CSF volumes in the SIH-initial ($P < .001$) and SIH-intermediate ($P < .05$) groups, but not the SIH-recovery group ($P = .772$), were significantly reduced (Table 3).

**Diffuse Pachymeningeal Enhancement**

Among the 23 patients with SIH, 15 had DPE (Table 4). The mean intraspinal CSF volumes of patients with SIH and DPE (70.14 ± 18.74 mL) did not differ significantly from those of patients with SIH without DPE (76.36 ± 21.10 mL, $P = .388$). The intraspinal CSF volume percentage (ie, quotient of the initial intraspinal CSF volumes and ECSF volumes) also did not differ significantly between patients with SIH with and without DPE ($P = .274$).

**Venous Engorgement**

Among the 23 patients with SIH, 13 showed venous engorgement and the other 10 did not (Table 4). The intraspinal CSF volumes of

---

**Table 1: Demographic data**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Patients with SIH</th>
<th>HCs</th>
<th>$P$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y (mean)$^a$</td>
<td>38.83 ± 11.83</td>
<td>35.56 ± 10.02</td>
<td>.393</td>
</tr>
<tr>
<td>BMI, kg/m$^2$ (mean)$^a$</td>
<td>24.97 ± 5.80</td>
<td>22.07 ± 4.24</td>
<td>.078</td>
</tr>
<tr>
<td>BH, m (mean)$^a$</td>
<td>1.67 ± 0.098</td>
<td>1.67 ± 0.078</td>
<td>.942</td>
</tr>
<tr>
<td>BW, kg (mean)$^a$</td>
<td>72.22 ± 20.72</td>
<td>61.48 ± 12.79</td>
<td>.077</td>
</tr>
<tr>
<td>Sex$^b$</td>
<td></td>
<td></td>
<td>.598</td>
</tr>
<tr>
<td>Male (No.)</td>
<td>8</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Female (No.)</td>
<td>15</td>
<td>22</td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Mann-Whitney U test.

$^b$ Two-sample t test.

---

**Table 2: Spinal MR imaging parameters and brain MRI findings from HCs and patients with SIH at initial, intermediate, and recovery stages**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Intraspinal CSF Volumes (ml)</th>
<th>Pituitary Volumes (cm$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCs (n = 32)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIH Ini ($n = 23$)</td>
<td>72.31 ± 19.35</td>
<td>0.60 ± 0.20</td>
</tr>
<tr>
<td>SIH Int ($n = 18$)</td>
<td>81.15 ± 25.41</td>
<td>0.47 ± 0.15</td>
</tr>
<tr>
<td>SIH Rec ($n = 18$)</td>
<td>93.74 ± 20.33</td>
<td>0.36 ± 0.14</td>
</tr>
<tr>
<td>$P$ value$^a$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCs vs SIH Ini</td>
<td>&lt;.001$^b$</td>
<td></td>
</tr>
<tr>
<td>HCs vs SIH Int</td>
<td>&lt;.05$^b$</td>
<td></td>
</tr>
<tr>
<td>HCs vs SIH Rec</td>
<td>.731</td>
<td></td>
</tr>
<tr>
<td>SIH Ini vs SIH Int</td>
<td>&lt;.001$^b$</td>
<td>&lt;.001$^b$</td>
</tr>
<tr>
<td>SIH Ini vs SIH Rec</td>
<td>&lt;.001$^b$</td>
<td>&lt;.001$^b$</td>
</tr>
<tr>
<td>SIH Int vs SIH Rec</td>
<td>&lt;.001$^b$</td>
<td>.355</td>
</tr>
</tbody>
</table>

$^a$ Mann-Whitney U test.

$^b$ Significant.

---

**FIG 1.** Linear regression equation and figure of intraspinal CSF volumes with BH.
patients with SIH with venous engorgement (66.24 ± 16.40 mL) did not differ from those of patients without venous engorgement (80.19 ± 20.84 mL, \( P = .128 \)). Although the intraspinal CSF volume percentage of patients with SIH with venous engorgement did not differ from that of those without venous engorgement \( (P = .058) \), there seemed to be a trend toward patients with SIH with venous engorgement having a lower intraspinal CSF volume percentage.

**Pituitary Volumes**

As shown in Table 2, pituitary volumes were significantly larger in the SIH-initial than in the SIH-intermediate and SIH-recovery measurements (both \( P < .001 \)). However, there was no significant difference in the SIH-intermediate and SIH-recovery pituitary volumes \( (P = .155) \). Pituitary volumes did not differ significantly between patients with SIH with and without DPE \( (P = .747, \) Table 4). However, the pituitary volumes of patients with SIH with venous engorgement were significantly greater than those of patients with SIH without venous engorgement \( (P < .05, \) Table 4).

**Case Illustration**

MR imaging and MRM images obtained for a 57-year-old woman who was diagnosed with CSF leakage into the spine are shown in Fig 2. The initial brain MR imaging showed no evidence of DPE, venous engorgement, or pituitary hyperemia. The calculated intraspinal CSF volume for this patient was 68.36 mL in the initial stage. The estimated intradural CSF for this patient was 79.11 mL, with an intraspinal CSF volume percentage of 86.5%. The intradural CSF volume at complete recovery was 80.04 mL.

MR imaging and MRM images obtained for a 34-year-old man who had orthostatic headache and was shown to have CSF leakage into the spine are shown in Fig 3. The calculated intraspinal CSF volume for this patient was 88.11 mL in the initial stage. DPE, venous engorgement, and pituitary hyperemia were noted in the patient’s initial brain MR imaging. The estimated intradural CSF for this patient was 103.32 mL, with an intraspinal CSF volume percentage of 85.3%. The intradural CSF volume at complete recovery was 109.15 mL.

**DISCUSSION**

In this study, we obtained direct measurements of intraspinal CSF volumes in patients with SIH at 3 clinical stages and compared them with those of HCs. Although the intraspinal CSF volumes in the SIH-initial and SIH-intermediate measurements included leaked CSF in the extradural space and intradural CSF, stepwise increases in intraspinal CSF volumes were observed during the disease recovery. Furthermore, the mean intraspinal CSF volume was significantly lower in the SIH-initial and -intermediate disease stages than in the SIH-recovery stage or in HCs. These findings support the hypothesis that CSF hypovolemia is a core feature of SIH syndrome.

The normal intraspinal CSF volume range for healthy subjects has not been well-documented. In subjects without CSF leakage, intraspinal CSF volumes should be equal to intradural CSF volumes. Hogan et al\(^\text{11}\) reported 2 healthy volunteers with intradural CSF volumes of the spine plus nerve roots of 95 and 120 mL. The only large-scale report was a study of 22 healthy volunteers having a mean intradural CSF volume of 81 ± 13 mL (range, 52–103 mL).\(^\text{14}\) In that study, the original images were obtained with 3D balanced turbo-field echo pulse sequences and the regions for measurement were drawn on the images manually. For CSF volume calculation, the original images were transformed to an axial image of 1.9-mm thickness with a 0.7-mm spacing covering the whole spine. Manually drawing the outline of the spinal cord, spinal canal, and nerve root on every 10th axial image was performed. The intradural CSF volume of the spine was calculated as the following: Area of Spinal Canal − Nerve Roots and Spinal Cord × Interslice Distance.

Here, we used automatic threshold-based segmentation methods for CSF analysis to estimate the mean intradural CSF volumes of 32 HCs. Slightly larger intradural CSF volumes were observed in our study. The reason for the volume differences might be from the imaging sequences and analytic methods. In our experiment, a heavily T2-weighted 3D-SPACE sequence was used to acquire whole spinal MRM, which provided higher isotropic spatial resolution and tissue contrast of CSF in whole spinal

Table 3: Comparison of intraspinal CSF volumes at each SIH stage relative to the ECSF value

<table>
<thead>
<tr>
<th>Parameter</th>
<th>ECSF (n = 23)</th>
<th>SIH</th>
<th>SIH</th>
<th>SIH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ini (n = 23)</td>
<td>Int (n = 18)</td>
<td>Rec (n = 18)</td>
<td></td>
</tr>
<tr>
<td>Intradural CSF volumes (mL)</td>
<td>93.03 ± 13.23</td>
<td>72.31 ± 19.35</td>
<td>81.15 ± 25.41</td>
<td>93.74 ± 20.33</td>
</tr>
<tr>
<td>( P ) values, ECSF vs SIH (^\text{b})</td>
<td>–</td>
<td>(&lt; .001)</td>
<td>(&lt; .05)</td>
<td>.772</td>
</tr>
</tbody>
</table>

Note: Ini indicates initial; Int, intermediate; Rec, recovery.

\(^{b}\) Paired \( t \) test.

\(^{b}\) Significant.

Table 4: Comparison of MR imaging parameters between patients with SIH versus those without DPE and venous engorgement

<table>
<thead>
<tr>
<th>Volume (Mean)</th>
<th>SIH with Ini DPE (n = 15)</th>
<th>SIH without Ini DPE (n = 8)</th>
<th>( P ) Value(^\text{a})</th>
<th>SIH with Ini Venous Engorgement (n = 13)</th>
<th>SIH without Ini Venous Engorgement (n = 10)</th>
<th>( P ) Value(^\text{a})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intradural CSF volumes (mL)</td>
<td>70.14 ± 18.74</td>
<td>76.36 ± 21.10</td>
<td>.388</td>
<td>66.24 ± 16.40</td>
<td>80.19 ± 20.84</td>
<td>.128</td>
</tr>
<tr>
<td>Intradural CSF volume percentage</td>
<td>78.59 ± 29.27</td>
<td>81.75 ± 19.64</td>
<td>.274</td>
<td>70.87 ± 17.74</td>
<td>91.16 ± 30.99</td>
<td>.058</td>
</tr>
<tr>
<td>Brain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pituitary volumes (cm(^3))</td>
<td>0.61 ± 0.20</td>
<td>0.58 ± 0.19</td>
<td>.747</td>
<td>0.67 ± 0.15</td>
<td>0.51 ± 0.22</td>
<td>(&lt;.05) (^\text{b})</td>
</tr>
</tbody>
</table>

Note: Ini indicates initial; Int, intermediate; Rec, recovery.

\(^{a}\) Mann-Whitney \( U \) test.

\(^{b}\) Significant.
We believed that sequence would be more sensitive for detecting the CSF voxels than using a balanced turbo-field echo sequence because the SPACE sequence has less susceptibility artifact–related image blurring and better contrast-to-noise ratio. Additionally, the automatic threshold-based segmentation method could effectively eliminate the discrepancy from intra- and interoperator variability and increase the reliability in clinical applicability. Our cohort was younger (range, 39–59 years) than those in prior studies. For example, in Edsbagge et al, all participants were older than 64 years of age. Degenerative spinal changes, such as spurs or stenosis, might reduce intradural CSF volumes. However, we did observe a wide variation of intradural CSF volumes across individuals.

Using ECSF volumes as a reference to predict individual intraspinal CSF volumes, we discovered that patients with SIH had significantly lower intraspinal CSF volumes at their initial and intermediate measurements. No significant differences in intraspinal CSF volumes were found between the recovery measurements and the ECSF values. We presented 2 cases of SIH with about 20-mL differences in their intraspinal CSF volumes (68.36 versus 88.11 mL) at the initial stage (Figs 2 and 3). It was difficult to determine whether their intraspinal CSF volumes were reduced by only using the absolute volumes. The male patient had higher ECSF than the female patient (103.32 versus 79.11 mL). With ECSF as a reference, both patients had significantly reduced CSF volumes with the intraspinal CSF percentage of about 85% causing orthostatic headache and other image abnormalities. Their intradural CSF volumes during recovery were 80.04 and 109.15 mL, respectively. Therefore, ECSF volumes might serve as an important reference for the diagnosis and evaluation of disease severity.

SIH-associated brain changes can be explained by the Monro-Kellie doctrine, wherein decreases in intracranial CSF volumes are accompanied by increased volumes of brain tissue (eg, pituitary hyperemia) or vascular structures (DPE or venous engorgement). In our study, intraspinal CSF volumes and volume percentage in patients with SIH with DPE (n = 15) were similar to the values obtained for patients without DPE (n = 8). However, the potential effect of venous engorgement on these parameters was less clear. Although there was no significant difference of intraspinal CSF volumes.
nal CSF volumes between patients with SIH with \( (n = 13) \) and without \( (n = 10) \) venous engorgement, there was a strong trend toward a difference in intraspinal CSF volume percentage between these 2 groups (borderline \( P \) value = .058). This trend may become a robustly significant difference with a larger sample size. Venous engorgement was present in patients with more severe CSF depletion and might not be present in those with mild CSF volume loss. In a previous study, we also demonstrated that DPE or venous engorgement reflected a more severe disease status and may not be seen in patients with mild disease.\(^6\) Therefore, we believe that intraspinal CSF volumes and, especially, intraspinal CSF volume percentage may be sensitive markers of SIH progression. Spine MR imaging has also been reported useful for diagnosing SIH, especially at early stages.\(^1\)

Pituitary hyperemia has also been thought to be a sensitive imaging marker of SIH. All the patients in this study had an enlarged pituitary gland, and the pituitary volumes decreased in relation to symptomatic improvement. Pituitary volumes were significantly greater in patients with venous engorgement than in those without it. Although pituitary volumes may reflect disease severity, pituitary volumes vary greatly across individuals, with larger volumes often being seen in young women.\(^2\) The maximum height was seen in the women between 20 and 40 years of age and declined with age.\(^3\) Twelve female patients achieved complete recovery in our study. Their pituitary volumes did not differ from those of male patients. Six of them are between 20 and 40 years of age, and the other 6 are older than 40 years of age. The height and pituitary volume were not significantly different between these 2 groups. Two patients younger than 40 years of age and 1 patient older than 40 years of age had convex-shape pituitary glands. Even though SIH usually occurred in young women, the pituitary size was still variable among subjects. Therefore, using absolute pituitary volumes as a diagnostic criterion of SIH is inappropriate. However, relative change in pituitary volumes is a good parameter for evaluation of treatment effects.

In this study, we found that intradural spinal CSF volumes correlated positively with BH. No correlations were observed with BMI, BW, age, and sex. In previous reports, no correlation was found between intradural CSF and BH,\(^4\) and inverse correlations of lumbosacral intradural CSF with BMI have been reported.\(^13,22\) The number of vertebrae has been used as an indicator of intradural CSF volume.\(^1\) Therefore, we thought that lower limb length might explain the variance in the relationship between BH and intradural CSF volumes across individuals. Additional large-scale studies may better uncover the relationships among clinical findings, brain imaging findings, and intracranial and intraspinal CSF volumes.

Our study has some limitations. First, because of the difficulty of the autosegmentation method in separating extradural CSF from intradural CSF in current status, the intraspinal CSF measurements included both extra- and intradural spaces in SIH-initial and SIH-intermediate measurements. Hence, the CSF that leaked through dural defects was also included in our calculations using our segmentation method. Even though the intraspinal CSF volumes in SIH-initial and SIH-intermediate measurements were still significantly lower than those of healthy controls and SIH-recovery measurements, most of our patients had a significant amount CSF leakage at the spinal nerve roots with epidural fluid accumulation. Therefore, if we took account of only the intradural CSF, we believe that the CSF volumes would be much lower than those in healthy controls in the patients with SIH. Further study with methods that can differentiate intradural CSF from extradural CSF might elucidate the dynamic changes in patients with SIH.

Second, our sample size was small, and intracraniad CSF volumes were not measured. Studies with larger sample sizes with measurements of both intracraniad and intraspinal CSF volumes may provide more information. Third, although hypovolemia is fundamentally related to SIH, individual variations of intradural CSF volumes do exist. The intraspinal CSF volumes could help monitor treatment effects as well. On the other hand, the method by which intradural CSF volumes are estimated for subjects will be important in making clinical application and accurate diagnoses feasible. ECSF volumes determined on the basis of BH as in our study appear to reflect clinical status well and could be used to help determine whether further treatment is needed. However, the correlation between intradural CSF volumes and BH was not high enough. It still needs to be integrated with clinical symptoms and other imaging findings. How to estimate intradural CSF volumes better still needs to be investigated by using other parameters and larger sample sizes.

**CONCLUSIONS**

With a threshold-based segmentation method, we demonstrated that spinal CSF hypovolemia was fundamentally related to SIH and may serve as an early detection method before abnormalities are detectable by conventional brain MR imaging. Intraspinad CSF volumes can be used as a sensitive parameter in treatment-response monitoring and in SIH follow-up evaluations.

**REFERENCES**

CT Recognition of Lateral Lumbar Disk Herniation

Although computed tomographic (CT) scanning has been useful in diagnosing spondylolisthesis and central lumbar disk herniation, its effectiveness in demonstrating lateral lumbar disk herniation has not been evaluated. The myelographic recognition of lateral lumbar disk herniation is limited, since the lateral aspect of the herniated disk is difficult to demonstrate. A study of 142 CT scans interpreted by a neuroradiologist who was not aware that a spondylolisthesis was present showed that CT visualized lateral disk herniations in 10 of 11 cases (91%). Four of these patients had a CT scan of the lumbar spine, but none of these patients had a spondylolisthesis. Because it can image the disk margins and the nerve root entry zone, the CT may be more effective than myelography for demonstrating the presence and extent of lateral disk herniation.

Materials and Methods

During a 2-year period, 1,350 patients with low back pain and/or sciatica pain were evaluated with a GE scanner (General Electric, Milwaukee, Wisconsin). CT scans were obtained from lumbosacral disks in 11 patients who were known to have a spondylolisthesis.

Results

In the 11 cases, displacement of fat within the intervertebral foramen, identified in each case, although a disc herniation was not demonstrated. The location of the disc was not always the same, and the extent of the herniation varied from one case to another. In some cases, the disk was located posteriorly and laterally, while in other cases, it was located anteriorly and medially. The relationship of the disk to the nerve root was also variable, with some cases showing direct compression and others showing indirect pressure on the nerve root.

Representative Case Report

Case 1: A 45-year-old man had a 4-week history of low back pain and radiculopathy. CT revealed a large, left lateral disk herniation that was compressing the nerve root. The patient was treated with a percutaneous diskectomy, and the symptoms resolved.

Atrial Diverticula in Severe Hydrocephalus

Mild ventriculodilatation causes stretching and deformation of the brain's white matter in a normally functioning brain. This stretching is often seen in patients with severe hydrocephalus. Ventriculocisternography is a sedimentation of the subarachnoid space, which can be used to identify and confirm the location of the diverticula. These diverticula are often associated with increased intracranial pressure and can cause symptoms such as headache, nausea, and vomiting.

Molecular and Methods

Several CT scans of 10 patients with ventriculodilatation were reviewed in a series of 50 patients with severe hydrocephalus. Patients were divided into two groups: those with normal intracranial pressure and those with elevated intracranial pressure. The patients were then evaluated with a combination of CT and MRI scans to identify the location of the diverticula. The results showed that the diverticula were more common in patients with severe hydrocephalus and were associated with increased intracranial pressure. CT and MRI were able to accurately identify the location of the diverticula and provided valuable information for planning the appropriate surgical approach.
We would like to thank Jacob et al for raising an important, relevant issue in their article entitled “MR Imaging of the Cervical Spine in Nonaccidental Trauma: A Tertiary Institution Experience.” They reported a relatively high incidence (69%) of cervical spine injury (CSI) in nonaccidental trauma compared with the literature and concluded that positive findings on MR imaging may affect management and, therefore, recommending routine use of MRI in suspected nonaccidental trauma.

However, it is not clear how the authors reached that conclusion. They have not specified how the MR imaging findings impacted management in their study population. Are the authors advocating the use of MR imaging for the diagnosis of nonaccidental trauma, distinguishing accidental from nonaccidental trauma, or management of patients with nonaccidental trauma? Although ligamentous signal on MR imaging was seen in 67% of patients, this was mostly confined to the interspinous and nuchal ligaments. From the description of findings in the study population, would the authors agree that none of the patients had unstable injury? Although this was a retrospective study, would prospective knowledge of these findings change management and in what circumstances?

Previous studies have shown that MR imaging signs of cervical spine injury did not show a statistical relationship with outcome or help discriminate accidental and abusive head trauma. Although cervical spine injury was seen on MR imaging in 36% of patients, none required surgical intervention in the study by Kadom et al.

Despite the higher soft-tissue contrast resolution, MR imaging has not been shown to detect unstable CSIs in patients with CT with negative findings, either in the pediatric or adult population. Choudhary et al found a higher incidence of ligamentous injury in nonaccidental trauma compared with accidental injuries, but the pattern of injuries was similar. Jacob et al used that study as evidence that MR imaging findings actually reflect pathology because the cohort of patients without trauma did not show similar MR imaging abnormalities. However, there was no blinding for the healthy cohort in the study by Choudhary et al. The few studies that have correlated MR imaging findings of ligamentous signal abnormalities with intraoperative or postmortem findings have found a poor correlation. In a comparison of 4 modalities to assess cervical spine instability in pediatric trauma, Brockmeyer et al found MR imaging to be sensitive but not specific (74%). Because of the low specificity, the authors recommended that MR imaging be performed only in patients with neurologic deficits or to follow up a screening study with positive findings.

Performance of MR imaging is not without challenges, risks, and costs in this study population. As per the authors, only 43% of patients in this study had superior quality imaging. The recommendation by the authors of routinely including cervical spine MR imaging as part of the armamentarium of tests while working up a child with nonaccidental trauma probably needs answers to our questions and more evidence.

REFERENCES

X. Wu
D. Durand
B. Rao
A. Malhotra
Department of Radiology and Biomedical Imaging
Yale School of Medicine
New Haven, Connecticut
We would like to thank the authors of a letter in reference to our recently published study, “MR Imaging of the Cervical Spine in Nonaccidental Trauma: A Tertiary Institution Experience.” The word “clearance” was not in the title of the study as suggested by the preceding letter, and our article did not address whether the MR imaging findings could determine cervical spine stability or instability. However, MR imaging with negative findings has been shown to be useful in clearing the cervical spine, which would have been the case in 31% of our patients. We agree that positive findings relating to instability are less well-defined and posterior ligamentous injury is likely a stable injury. The additional findings of capsular injury with distension (13%) and tectorial ligament injury (4%) are more worrisome, and we believe these would alter management, but again this was not the purpose of our study.

These children do not necessarily have a history of a traumatic incident, and positive findings on MR imaging may be helpful to define the injury as traumatic. We do not advocate the use of MR imaging for defining a ligamentous injury as accidental or nonaccidental because these injuries have similar findings, but the positive findings are significantly more common in the nonaccidental cohort as reported by Choudhary et al. Most interesting, the presence of spinal subdural hemorrhage, likely from redistribution, may be indicative of a nonaccidental mechanism because it was rarely seen in the accidental cohort. Spinal subdural hemorrhage was seen in 18% of our patients.

In our study, the cervical spine injuries were statistically associated with parenchymal brain injury; therefore, poor outcome may not be due to the cervical spine injury but related to the associated brain injury. We realize that the prior study by Kadom et al did not find an association between outcome and cervical spine injury as diagnosed on MR imaging and further study is needed. Additionally, a positive finding on cervical MR imaging would strengthen the case for trauma as an etiology for the brain injury.

The letter indicates that only 43% of the MR imaging studies in our study were of superior quality, but studies of all patients were categorized as superior or diagnostic. It is more important to realize the inherent insensitivity of cervical radiographs in the evaluation of cervical spine injury, especially in young children. Also, most of our injuries were at the cervical-occipital junction, and CT criteria for evaluating this injury have recently been re-evaluated and found to be suboptimal.

In this vulnerable population, MR imaging of the cervical spine may provide additional answers. Many of these examinations can be performed in conjunction with brain MR imaging without additional sedation/anesthesia. A small number of children in our study had normal brain imaging findings and evidence of cervical spine injury. This information would be extremely helpful in the evaluation of a child with suspected nonaccidental trauma. We believe our study supports the use of cervical MR imaging in this cohort of patients.

REFERENCES


http://dx.doi.org/10.3174/ajnr.A5111
I read with great interest the study of Hodel et al,¹ which demonstrated that the combined use of arterial spin-labeling (ASL) and susceptibility-weighted imaging was significantly more sensitive and equally specific compared with conventional angiographic MR imaging for the detection of arteriovenous shunting. In Fig 2 of this article, a lesion with increased signal intensity on ASL images was subsequently confirmed as a developmental venous anomaly (DVA) based on the presence of a classic umbrella-shaped appearance on the venous phase cerebral angiogram.¹

Most DVAs are not associated with perfusion changes on ASL imaging. In a large study of 652 DVAs, only a minority of DVAs (8%) demonstrated signal abnormalities on ASL maps,² and intrinsically increased ASL signal or increased signal in a draining vein associated with a DVA are potentially imaging biomarkers of arteriovenous shunting.²,³ It may be argued that the presence of a typical venous phase cerebral angiogram with a caput medusa appearance will confirm the DVA. However, direct arteriovenous shunting into dilated medullary veins characteristic of a DVA without a typical nidus has been repeatedly reported in the literature. These arterialized DVAs have been described with various terminology, but Im et al⁴ proposed the name of “venous-predominant parenchymal AVMs” on the basis of clinical, angiographic, surgical, and histologic findings in a series of 15 cases. Of note, the cerebral angiograms of these lesions demonstrated an abnormal blush in the early arterial phase, followed by diffuse vascular and capillary parenchymal staining in the mid-arterial phase and immediate drainage into a network of radially arranged dilated medullary veins.⁴ These lesions consistently showed a caput medusa appearance very similar to the appearance of DVAs, with the absence of enlarged arterial feeders and lack of a typical AVM nidus.

In summary, a lesion with increased signal intensity on an ASL map is suspicious for arteriovenous shunting, and a caput medusa appearance on the venous phase of a cerebral angiogram is not sufficient to make the diagnosis of DVA. Careful attention to the arterial phase of the angiogram would be prudent to exclude a venous-predominant parenchymal arteriovenous malformation.

REFERENCES

S.A. Nabavizadeh
Department of Radiology
Hospital of University of Pennsylvania
Perelman School of Medicine at the University of Pennsylvania
Philadelphia, Pennsylvania

http://dx.doi.org/10.3174/ajnr.A5008
We thank Dr Nabavizadeh for his interest in our article, "Intracranial Arteriovenous Shunting: Detection with Arterial Spin-Labeling and Susceptibility-Weighted Imaging Combined." 1

We agree that arterial spin-labeling (ASL) hypersignal in developmental venous anomalies (DVAs) may suggest transitional or mixed malformations with arteriovenous shunting. DVA may also coexist with a true arteriovenous malformation, as reported by Erdem et al.,2 with selective arterial embolization of the AVM while preserving the DVA. As highlighted by Dr Nabavizadeh, the presence of dilated deep medullary veins in a “spoke wheel” or caput medusae at the venous phase of angiograms is not sufficient for the diagnosis of DVA. Careful attention must be paid to the arterial phase to rule out arteriovenous shunting resulting in early opacification of the draining vein at the arterial phase of the angiogram. In all 8 patients with DVAs included in our study, the main draining vein was never opacified in the early and mid-arterial phases of the angiograms. Only a capillary stain was observed in 2 patients, forming a blush at the arterial phase, without any arteriovenous communication. These findings are in agreement with those reported by Roccatagliata et al.,3 suggesting that a large spectrum of lesions exists from typical DVAs to clearly distinct AVMs. These cases probably correspond to DVAs surrounded by dilated arteriolar-capillary channels that may exceptionally lead to ischemic or hemorrhagic complications. This pattern may be due to the lack of adaptability to variations in intracranial venous equilibrium, as suggested by the authors.3

However, the natural history, prognosis, and therapeutic management of such lesions remain unclear and are widely debated in the literature. For a systematic review of digital subtraction angiography images in our study, we chose to classify the vascular malformations as DVAs when angiograms demonstrated a caput medusae appearance without opacification of the main draining vein at the arterial phase of the angiograms. If we had decided to classify these atypical DVAs on DSA as arteriovenous shunting, the specificity of ASL would have been even higher than that reported.

Signal abnormalities on ASL images may be located either in the brain parenchyma surrounding the DVA or within the draining vein of the DVA. We agree that it may be speculated that a hypersignal of a part of the DVA on ASL images may potentially reflect the presence of an arteriovenous shunting, but little data exist in the literature comparing ASL signal with DSA in patients with DVA. In addition, erroneous interpretation of CBF may be made in patients with prolonged arterial arrival times. Thus, further studies are required to understand the real significance of these findings.

REFERENCES


http://dx.doi.org/10.3174/ajnr.A5331
In their recent article, Yuan et al. stated that they “highlight the merits and limitations of various imaging techniques for identifying carotid plaque ulceration.” It is curious, therefore, that they made no mention of 3D sonography (3DUS). Although ulceration seen on angiography predicted risk in the North American Symptomatic Carotid Endarterectomy Trial, it was clear that imaging of the lumen by angiography was not a reliable way to assess ulceration.

Schminke et al. reported that 3DUS was a better way to image ulceration than 2D sonography. In 2011, we reported that the number of carotid ulcers (Fig 1) was as strong a predictor of risk among patients with asymptomatic carotid stenosis; patients with ≥3 ulcers in either carotid artery had a risk equivalent to that of patients with microemboli on transcranial Doppler sonography.

In 2104, we reported that ulcer volume predicted cardiovascular risk among patients attending cardiovascular prevention clinics (Fig 2). It is likely that the best way to image and quantify carotid ulceration is with 3DUS.

Disclosures: J. David Spence—UNRELATED: Other: Dr Spence is an officer of Vascularis Inc.

REFERENCES

http://dx.doi.org/10.3174/ajnr.A5130
FIG 1. Number of carotid ulcers. The carotid artery can be examined in any axis for detection of ulcers in atherosclerotic plaques. White arrows show ulcers. Reproduced with permission of Wolters Kluwer from Madani et al.5
FIG 2. Measurement of ulcer volume and ulcer depth. Contours of ulcers were traced, and the depth of ulcers was measured in cross-sectional views. Each section had a thickness of 1 mm; ulcer volume was computed from the sum of the volumes of all sections in which ulceration was traced. Reproduced with permission of Wolters Kluwer from Kuk et al.5
We thank Dr J. David Spence for his interest in our Review Article, “Imaging Carotid Atherosclerosis Plaque Ulceration: Comparison of Advanced Imaging Modalities and Recent Developments.”1 We also thank the Editor of the *American Journal of Neuroradiology* for giving us the opportunity to respond to his comments.

The principles, usage, advantage, and related studies of 3D sonography (US) are discussed in the last paragraph of the “Sonography” section.1 The application of 3D US in measuring plaque stenosis and volume is mentioned in the same paragraph. An example of comparing 3D and 2D US in depicting carotid ulceration is shown in Fig 4. In the On-line Table, we cite the study by Heliopoulos et al2 (reference 36 in the article), which demonstrates that the 3D US shows the ulceration more frequently than 2D US and has slightly superior interobserver reproducibility.2

The purpose of this Review Article was to compare imaging modalities in the evaluation of carotid ulceration. We primarily reviewed studies that compared multiple methods, as summarized in the On-line Table. We thank Dr Spence for bringing to our attention his valuable work with 3D US in imaging ulcerations.3,4

**REFERENCES**


http://dx.doi.org/10.3174/ajnr.A5146