Medicare for all: considerations for neuroradiologists
MRI discrimination of brain stem nuclei and pathways
Fast stent retrieval improves recanalization rates of thrombectomy
Tumor response assessment in diffuse intrinsic pontine glioma

Official Journal ASNR • ASFNR • ASHNR • ASPNR • ASSR
FRED™
Flow Re-Direction
Endoluminal Device

FLOW DIVERSION.
SIMPLIFIED.
The New Standard of Ease and Simplicity in Flow Diversion Treatment.

The Flow Re-Direction Endoluminal Device (FRED™) System is indicated for use in the internal carotid artery from the petrous segment to the terminus for the endovascular treatment of adult patients (22 years of age or older) with wide-necked (neck width ≥ 4 mm or dome-to-neck ratio < 2) saccular or fusiform intracranial aneurysms arising from a parent vessel with a diameter ≥ 2.0 mm and ≤ 5.0 mm.

Use of the FRED™ System is contraindicated under these circumstances: Patients in whom anticoagulant, anti-platelet therapy, or thrombolytic drugs are contraindicated. Patients with known hypersensitivity to metal such as nickel-titanium and metal jewelry. Patients with anatomy that does not permit passage or deployment of the FRED™ System. Patients with an active bacterial infection. Patients with a pre-existing stent in place at the target aneurysm. Patients in whom the parent vessel size does not fall within the indicated range. Patients who have not received dual anti-platelet agents prior to the procedure. For complete indications, contraindications, potential complications, warnings, precautions, and instructions, see instructions for use (IFU provided in the device).

RX Only: Federal (United States) law restricts this device to sale by or on the order of a physician.

MICROVENTION™ and FRED™ are registered trademarks of MicroVention, Inc. in the United States and other jurisdictions. © 2020 MicroVention, Inc. 04/2020.
737 PERSPECTIVES A. Gordhan

REVIEW ARTICLES

738 Tyrosine Kinase Inhibitor Therapy for Brain Metastases in Non-Small-Cell Lung Cancer: A Primer for Radiologists C. Dodson, et al.

751 Advanced Multicompartiment Diffusion MRI Models and Their Application in Multiple Sclerosis D.A. Lakhani, et al.

LEVEL 1 EBM EXPEDITED PUBLICATION

758 Methylphenidate Effects on Cortical Thickness in Children and Adults with Attention-Deficit/Hyperactivity Disorder: A Randomized Clinical Trial K.B. Walhovd, et al.

RESEARCH PERSPECTIVES


PRACTICE PERSPECTIVES

772 Medicare for All: Considerations for Neuroradiologists T.H. Nguyen, et al.

GENERAL CONTENTS

777 Direct In Vivo MRI Discrimination of Brain Stem Nuclei and Pathways T.M. Shepherd, et al.

785 Acetazolamide-Loaded Dynamic 7T MR Quantitative Susceptibility Mapping in Major Cerebral Artery Steno-Occlusive Disease: Comparison with PET K. Fujimoto, et al.


804 The Impact of Intracortical Lesions on Volumes of Subcortical Structures in Multiple Sclerosis I. Kalinin, et al.

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Journal</th>
</tr>
</thead>
<tbody>
<tr>
<td>815</td>
<td>Prognostic Value of Preoperative MRI Metrics for Diffuse Lower-Grade</td>
<td>ADULT BRAIN</td>
</tr>
<tr>
<td></td>
<td>Glioma Molecular Subtypes</td>
<td>P. Darvishi, et al.</td>
</tr>
<tr>
<td>822</td>
<td>A Critical Assessment of the Golden Hour and the Impact of Procedural</td>
<td>INTERVENTIONAL</td>
</tr>
<tr>
<td>828</td>
<td>Predictors of Cerebral Aneurysm Rupture after Coil Embolization:</td>
<td>INTERVENTIONAL</td>
</tr>
<tr>
<td></td>
<td>Single-Center Experience with Recanalized Aneurysms</td>
<td>Y. Funakoshi, et al.</td>
</tr>
<tr>
<td>836</td>
<td>Signal of Carotid Intraplaque Hemorrhage on MR T1-Weighted Imaging:</td>
<td>EXTRACRANIAL VASCULAR</td>
</tr>
<tr>
<td></td>
<td>Association with Acute Cerebral Infarct</td>
<td>D. Yang, et al.</td>
</tr>
<tr>
<td>844</td>
<td>Distinguishing Recurrent Thyroid Cancer from Residual Nonmalignant</td>
<td>HEAD &amp; NECK</td>
</tr>
<tr>
<td></td>
<td>Thyroid Tissue Using Multiphasic Multidetector CT</td>
<td></td>
</tr>
<tr>
<td>852</td>
<td>MRI Signal Intensity and Electron Ultrastructure Classification</td>
<td>HEAD &amp; NECK</td>
</tr>
<tr>
<td></td>
<td>Predict the Long-Term Outcome of Skull Base Chordomas</td>
<td>J. Bai, et al.</td>
</tr>
<tr>
<td>859</td>
<td>Inter- and Intrareader Agreement of NI-RADS in the Interpretation of</td>
<td>HEAD &amp; NECK</td>
</tr>
<tr>
<td></td>
<td>Surveillance Contrast-Enhanced CT after Treatment of Oral Cavity and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oropharyngeal Squamous Cell Carcinoma</td>
<td></td>
</tr>
<tr>
<td>866</td>
<td>Tumor Response Assessment in Diffuse Intrinsic Pontine Glioma:</td>
<td>PEDIATRICS</td>
</tr>
<tr>
<td></td>
<td>Comparison of Semiautomated Volumetric, Semiautomated Linear, and</td>
<td>FUNCTIONAL</td>
</tr>
<tr>
<td>874</td>
<td>MR Imaging Correlates for Molecular and Mutational Analyses in</td>
<td>PEDIATRICS</td>
</tr>
<tr>
<td>882</td>
<td>Synthetic MRI of Preterm Infants at Term-Equivalent Age: Evaluation</td>
<td>PEDIATRICS</td>
</tr>
<tr>
<td></td>
<td>of Diagnostic Image Quality and Automated Brain Volume Segmentation</td>
<td>FUNCTIONAL</td>
</tr>
<tr>
<td>889</td>
<td>Disrupted Functional and Structural Connectivity in Angelman Syndrome</td>
<td>PEDIATRICS</td>
</tr>
<tr>
<td></td>
<td>H.M. Yoon, et al.</td>
<td>FUNCTIONAL</td>
</tr>
<tr>
<td>898</td>
<td>Balanced Steady-State Free Precession Techniques Improve Detection</td>
<td>PEDIATRICS</td>
</tr>
<tr>
<td>904</td>
<td>Neuroimaging Findings in Children with Constitutional Mismatch Repair</td>
<td>PEDIATRICS</td>
</tr>
<tr>
<td></td>
<td>Deficiency Syndrome</td>
<td>A. Kerpel, et al.</td>
</tr>
<tr>
<td>911</td>
<td>Utility of Pre-Hematopoietic Cell Transplantation Sinus CT Screening</td>
<td>PEDIATRICS</td>
</tr>
<tr>
<td>917</td>
<td>The Perirolandic Sign: A Unique Imaging Finding Observed in</td>
<td>PEDIATRICS</td>
</tr>
<tr>
<td></td>
<td>Association with Polymerase γ-Related Disorders</td>
<td>F.G. Gonçalves, et al.</td>
</tr>
<tr>
<td>923</td>
<td>Cerebellar Watershed Injury in Children</td>
<td>PEDIATRICS</td>
</tr>
<tr>
<td></td>
<td>J.N. Wright, et al.</td>
<td></td>
</tr>
<tr>
<td>929</td>
<td>Sensitivity of the Inhomogeneous Magnetization Transfer</td>
<td>SPINE</td>
</tr>
<tr>
<td></td>
<td>Imaging Technique to Spinal Cord Damage in Multiple Sclerosis</td>
<td>FUNCTIONAL</td>
</tr>
<tr>
<td>938</td>
<td>MR Myelography for the Detection of CSF-Venous Fistulas</td>
<td>SPINE</td>
</tr>
<tr>
<td>941</td>
<td>35 YEARS AGO IN AJNR</td>
<td></td>
</tr>
</tbody>
</table>
ONLINE FEATURES

LETTERS
E26 Identification of Vortex Cores in Cerebral Aneurysms on 4D Flow MRI  S. Meckel, et al.
E27 Reply  K. Futami, et al.
E29 Reply  K. Nael
E30 How Far Can We Take Vessel Wall MRI for Intracranial Atherosclerosis? The Tissue is Still the Issue  M.D. Alexander, et al.
E32 Reply  M.H.T. Zwartbol, et al.
E33 Time to Discontinue Use of the Term Stroke in Neuroradiology  D.A.L. Ortega, et al.
E34 Reply  M.A. Aftab

BOOK REVIEWS  R.M. Quencer, Section Editor
Please visit www.ajnrblog.org to read and comment on Book Reviews.
AJNR urges American Society of Neuroradiology members to reduce their environmental footprint by voluntarily suspending their print subscription.

The savings in paper, printing, transportation, and postage directly fund new electronic enhancements and expanded content.

The digital edition of AJNR presents the print version in its entirety, along with extra features including:

- Publication Preview
- Case Collection
- Podcasts
- The AJNR News Digest
- The AJNR Blog

It also reaches subscribers much faster than print. An electronic table of contents will be sent directly to your mailbox to notify you as soon as it publishes.

Readers can search, reference, and bookmark current and archived content 24 hours a day on www.ajnr.org.

ASNR members who wish to opt out of print can do so by using the AJNR Go Green link on the AJNR Website (http://www.ajnr.org/content/subscriber-help-and-services). Just type your name in the email form to stop print and spare our ecosystem.
and visualize areas with disruption of the blood brain barrier (BBB) and/or MRI in brain (intracranial), spine and associated tissues in adult and pediatric patients (including term neonates) to detect and visualize areas with disruption of the blood brain barrier (BBB) and/or abnormal vascularity.

**CONTRAINDICATIONS**

History of clinically important hypersensitivity reactions to DOTAREM.

**WARNINGS AND PRECAUTIONS**

- **Hypersensitivity Reactions:** Anaphylactic and anaphylactoid reactions have been reported with DOTAREM, involving cardiovascular, respiratory, and/or cutaneous manifestations. Some patients experienced circulatory collapse and died. In most cases, initial symptoms occurred within minutes of DOTAREM administration and resolved with prompt emergency treatment.

- Before DOTAREM administration, assess all patients for any history of a reaction to contrast media, bronchospasm, coma and convulsion. These patients may have an increased risk for a hypersensitivity reaction to DOTAREM.

- Administer DOTAREM only in situations where trained personnel and therapies are promptly available for the treatment of hypersensitivity reactions, including personnel trained in resuscitation.

- **Gadolinium Retention:** Gadolinium is retained for months or years in several organs. The highest concentrations have been identified in the bone, followed by brain, skin, kidney, liver and spleen. The duration of retention also varies by tissue, and is longest in bone. Linear GBCAs cause more retention than macrocyclic GBCAs.

- Consequences of gadolinium retention in the brain have not been established. Adverse events involving multiple organ systems have been reported in patients with normal renal function without an established causal link to gadolinium retention.

- **Acute Kidney Injury:** In patients with chronically reduced renal function, acute kidney injury requiring dialysis has occurred with the use of GBCAs. The risk of acute kidney injury may increase with increasing dose of the contrast agent; administer the lowest dose necessary for adequate imaging.

- Extravasation and Injection Site Reactions: Ensure catheter and venous patency before the injection of DOTAREM. Extravasation into tissues during DOTAREM administration may result in tissue irritation.

**ADVERSE REACTIONS**

- The most common adverse reactions associated with DOTAREM in clinical trials were nausea, headache, injection site pain, injection site coldness and rash.

- Serious adverse reactions in the Postmarketing experience have been reported with DOTAREM. These serious adverse reactions include but are not limited to: arrhythmia, cardiac arrest, respiratory arrest, pharyngeal edema, laryngospasm, bronchospasm, coma and convulsion.

**INDICATIONS AND USAGE**

DOTAREM® (gadoterate meglumine) injection is a prescription gadolinium-based contrast agent indicated for intravenous use with magnetic resonance imaging (MRI) in brain (intracranial), spine and associated tissues in adult and pediatric patients (including term neonates) to detect and visualize areas with disruption of the blood brain barrier (BBB) and/or abnormal vascularity.

**USE IN SPECIFIC POPULATIONS**

- **Pregnancy:** GBCAs cross the human placenta and result in fetal exposure and gadolinium retention. Use only if imaging is essential during pregnancy and cannot be delayed.

- **Lactation:** There are no data on the presence of gadoterate in human milk, the effects on the breastfed infant, or the effects on milk production. However, published lactation data on other GBCAs indicate that 0.01 to 0.04% of the maternal gadolinium dose is present in breast milk.

- **Pediatric Use:** The safety and efficacy of DOTAREM at a single dose of 0.1 mmol/kg has been established in pediatric patients from birth (term neonates ≥ 37 weeks gestational age) to 17 years of age based on clinical data. The safety of DOTAREM has not been established in preterm neonates. No cases of NSF associated with DOTAREM or any other GBCA have been identified in pediatric patients age 6 years and younger.

You are encouraged to report negative side effects of prescription drugs to the FDA. Visit www.fda.gov/medwatch or call 1-800-FDA-1088.

*Please see the full Prescribing Information, including the patient Medication Guide, for additional important safety information.*

**References:**

2. Maravilla K et al. Comparison of Gadotere Meglumine and Gadobutrol in the Diagnosis of Primary Brain Tumors: A Double-Blind Randomized Controlled Intraindividual Crossover Study (the REMIND Study). 2017 June 29. doi: 10.3174/ajnr.A5316. [Epub ahead of print].

DOTAREM® is a registered trademark of Guerbet LLC, and is available by prescription only.

Gadavist® is a registered trademark of the Bayer group of companies, and is available by prescription only.
Despite the difference in relaxivity between these 2 GBCAs

**There Is No Measurable Difference in Clinical Benefit Observed Between Dotarem® and Gadavist®**

as demonstrated by the REMIND Study, a multicenter, double-blind, randomized, controlled intraindividual crossover study.²

This study “demonstrates the non-inferiority of gadoterate meglumine [Dotarem®] versus gadobutrol [Gadavist®] for overall visualization and characterization of primary brain tumors”.² Additionally, there was no preference of the readers for either contrast agent, in most cases, regarding border delineation, internal morphology, and the qualitative degree of contrast enhancement, despite quantitative mean lesion percentage enhancement being higher with gadobutrol.²

- For all readers in the REMIND Study, more than 90% of patients presented with good or excellent overall lesion visualization and characterization with either Dotarem® or Gadavist®.²
- The REMIND Study also demonstrated a low incidence of immediate reported AEs with Dotarem® and with Gadavist®, as shown in multiple previous studies.²³⁴
- Dotarem® is the first imaging contrast with a macrocyclic and ionic structure for high thermodynamic and kinetic stability.⁵⁶
- Dotarem® is not only trusted for high molecular stability⁵⁶; the REMIND Study demonstrates that it is as effective as Gadavist® for MRI diagnosis of primary brain tumors.²

Please see Important Safety information on opposite page. For more information on Dotarem®, please see Full Prescribing Information including Boxed Warning and Medication Guide.
Introducing

Eclipse 2L

DUAL LUMEN BALLOON CATHETER

Balt’s Access Portfolio Delivers Unique Solutions

Visit www.balt-usa.com to learn more
Title: Victory of the Healing Hand in the Fight Against Covid 19, oil painting.
Ajeet Gordhan, MD, Neurointerventional Radiologist, Bloomington, Illinois

To keep fighting against a seemingly invincible enemy at a reminiscent time the world faced a common threat, Winston Churchill said,

“You ask, what is our aim? I can answer in one word. It is victory. Victory at all costs — Victory in spite of all terrors — Victory, however long and hard the road may be, for without victory there is no survival.

Against this pandemic, Covid 19. Victory is our aim.
Tyrosine Kinase Inhibitor Therapy for Brain Metastases in Non-Small-Cell Lung Cancer: A Primer for Radiologists

C. Dodson, T.J. Richards, D.A. Smith, and N.H. Ramaiya

ABSTRACT

SUMMARY: Treatment options for patients who develop brain metastases secondary to non-small-cell lung cancer have rapidly expanded in recent years. As a key adjunct to surgical and radiation therapy options, systemic therapies are now a critical component of the oncologic management of metastatic CNS disease in many patients with non-small-cell lung cancer. The aim of this review article was to provide a guide for radiologists, outlining the role of systemic therapies in metastatic non-small-cell lung cancer, with a focus on tyrosine kinase inhibitors. The critical role of the blood-brain barrier in the development of systemic therapies will be described. The final sections of this review will provide an overview of current imaging-based guidelines for therapy response. The utility of the Response Assessment in Neuro-Oncology criteria will be discussed, with a focus on how to use the response criteria in the assessment of patients treated with systemic and traditional therapies.

ABBRévIATiONS: ALK = anaplastic lymphoma kinase; CTLa-4 = cytotoxic T-lymphocyte-associated protein 4; EGFR = epidermal growth factor receptor; EML4 = echinoderm microtubule-associated protein-like 4; ICI = immune checkpoint inhibitor; NSCLC = non-small-cell lung cancer; PD-1 = programmed cell death protein 1; PFS = progression-free survival; PRES = posterior reversible encephalopathy syndrome; QALY = quality-adjusted life years; RANO = Response Assessment in Neuro-Oncology; RECIST = Response Evaluation Criteria in Solid Tumors; ROSi = C-ras oncogene i; SRS = stereotactic radiosurgery; TKi = tyrosine kinase inhibitor; WBRT = whole-brain radiation therapy

In recent years, several systemic therapies have been developed that demonstrate efficacy in non-small-cell lung cancer (NSCLC) metastatic to the brain. It is essential for radiologists to be informed about these novel therapeutic agents to evaluate the intracranial response to therapy on imaging in patients undergoing treatment with these drugs. As the systemic chemotherapeutic options for metastatic NSCLC evolve, the assessment of the response of brain metastases on imaging is becoming ever more critical for evaluating and managing patients with metastatic NSCLC.

Brain metastasis is a serious occurrence in patients with cancer and is observed at the highest frequency in NSCLC among all cancer types, with approximately 50% of patients with NSCLC being affected by brain metastases.1 Historically, the prognosis for patients with NSCLC diagnosed with brain metastases has been dismal. Before the advent of FDA-approved targeted therapies, the 5-year survival rate was 2.3% for patients with NSCLC following a diagnosis of brain metastases.2 Cytotoxic chemotherapeutics are largely ineffective intracranially due to their inability to cross the BBB. Therefore, the standard treatment for CNS metastasis has traditionally been local interventions such as an operation, stereotactic radiosurgery (SRS), and whole-brain radiation therapy (WBRT).3

However, the recent discovery of targetable driver mutations such as those involving the epidermal growth factor receptor (EGFR), anaplastic lymphoma kinase (ALK), and C-ras oncogene 1 (ROS1) genes in NSCLC has led to the creation of small-molecule tyrosine kinase inhibitors (TKIs) that target these driver mutations. These TKIs can also traverse the blood-brain barrier to access the brain parenchyma, which has led to increased survival rates for patients with NSCLC with brain metastases. Before the development of TKIs, studies reported the mean overall survival of patients with NSCLC brain metastases as 5 months with motexafin gadolinium chemotherapy plus WBRT and 8 months with cisplatin and etoposide chemotherapy.4,5 The concept of using tyrosine kinase inhibitors to target driver mutations has also been investigated in brain metastases from cancer types other than NSCLC. For instance, the EGFR inhibitor lapatinib has shown efficacy in combination with capecitabine in patients with brain metastases originating from HER-2-positive breast cancer.6 TKI therapy has also been shown to reduce the incidence of brain metastases in patients with metastatic renal cell cancer.7
In the era of targeted therapy, selection of the most effective course of therapy relies on molecular genetic testing of a biopsy or cytology sample of malignant tissue. For preservation, tissue samples are typically processed in a formalin-fixed, paraffin-embedded block, while cytology samples are prepared in an artificial cell block or a smear.\(^8\) Joint guidelines from the College of American Pathologists, the International Association for the Study of Lung Cancer, and the Association for Molecular Pathology categorize EGFR, ALK, and ROS1 as the 3 “must-test” genes for which testing must be offered at all laboratories conducting molecular testing on lung cancer.\(^9\) The guidelines also assert that EGFR, ALK, and ROS1 testing should be performed on all patients with advanced-stage lung adenocarcinoma. Reverse transcription polymerase chain reaction or next-generation sequencing can be used to detect mutations in EGFR, ALK, and ROS1 genes. ALK or ROS1 rearrangements can also be detected by fluorescence in situ hybridization.

Despite the high cost of new targeted therapies, the molecular genetic testing and the TKI medications themselves have been shown to be a relatively cost-effective means of treatment. Narita et al\(^10\) found that EGFR-mutation testing and gefitinib therapy are more cost-effective than standard chemotherapy and no EGFR testing, with the EGFR testing/gefitinib strategy resulting in 0.036 quality-adjusted life years (QALY) per patient and an incremental cost-effectiveness ratio of US $32,500 per QALY gained. Holleman et al\(^11\) demonstrated that osimertinib is the most effective EGFR-TKI, with a mean of 2.01 QALY gained per patient at a cost of €128,343 (US $143,485) per QALY gained. Djalalov et al\(^12\) found that genetic testing for the EML4 (echinoderm microtubule-associated protein-like 4)-ALK fusion gene followed by the first-generation ALK-inhibitor crizotinib therapy resulted in a gain of 0.011 QALY, with an incremental cost-effectiveness ratio of CaD $255,970 (US $220,641) per QALY gained compared with standard chemotherapy. Compared with the first-generation ALK-inhibitor crizotinib, second-generation alectinib has resulted in a gain of 0.87 QALY, with an incremental cost-effectiveness ratio of US $39,312 per QALY.\(^13\) Of note, most studies on the cost-effectiveness of TKIs have been conducted in Canada, Europe, and Asia, indicating the need for additional studies that examine the cost-effectiveness of these drugs within the American health care system.

While a body of research on the development of targeted therapies for NSCLC brain metastases is rapidly accruing in the field of clinical oncology, much less has been written in the radiology literature. This article, therefore, aims to provide radiologists with a foundational overview of the TKIs that cross the BBB in NSCLC brain metastases and to provide a useful guide for radiologically assessing the effects of these novel drugs in the treatment of NSCLC brain metastases.

**Pathophysiology of the Blood-Brain Barrier in Brain Metastasis**

To appreciate the imaging features relevant to systemic treatment of brain metastases, an understanding of the pathophysiology of the BBB is critical. The BBB is a selectively permeable, nonfenestrated layer of endothelial cells that enclose the cerebral capillaries. This endothelial cell layer is surrounded by a supportive layer of extracellular matrix, pericytes, and astrocytes.\(^14\) Tight junctions connect the endothelial cells and allow the BBB to regulate the passage of molecules between cerebral capillaries and the brain parenchyma on the basis of the size and charge of these molecules.\(^15\) The intercellular tight junctions connecting the epithelial cells of the barrier are impermeable to many hydrophilic systemic chemotherapy agents, which explains the low intracranial efficacy of these agents.\(^16,17\) Efflux pumps unique to the BBB inhibit the entry of certain compounds into the BBB endothelial cells and also transport compounds out of the endothelial cells to return them to blood vessels, thereby preventing these compounds from reaching the brain parenchyma. Larger molecules are more likely than small molecules to be substrates of these efflux transporters. The small size of TKIs relative to cytotoxic chemotherapeutics may therefore provide a possible mechanism behind the greater BBB penetration with TKIs compared with cytotoxic chemotherapeutic agents.\(^18\) Additionally, the good cellular permeability of TKIs may also explain this efficacy of TKIs in treating intracranial metastases relative to traditional chemotherapeutics.

Metastasis begins when the metastatic tissue detaches from a primary tumor, enters the extracellular matrix, and migrates into a blood vessel.\(^15\) The circulating tumor cells migrate haptotropically to the site of metastasis, where they extravasate from the blood vessel into the parenchyma of the target organ. In the case of brain metastasis, this extravasation from cerebral capillaries into the brain parenchyma requires BBB penetration. Circulating tumor cells traverse the BBB by interacting with the barrier via mechanisms not yet fully elucidated, after which they establish a secondary tumor in the brain parenchyma.\(^14\) The expression of vascular endothelial growth factors by the metastatic cancer cells allows the metastatic tissue to establish a blood supply.\(^19\) This vascular endothelial growth factor secretion also initiates signaling cascades that cause the endothelium of the BBB to become hyperpermeable, thereby transforming the BBB into a blood-tumor barrier that has increased permeability to some chemotherapeutic agents.\(^19\)

Metastatic proliferation is further enhanced by the immune-privileged nature of the brain due to the BBB, as well as the lack of access of many chemotherapeutic drugs to the brain parenchyma, both of which allow NSCLC metastases to establish secondary tumors and thrive in the brain parenchyma. The recent development of targeted therapies that penetrate the BBB affords patients with brain metastases an increased survival that has not previously been achievable with systemic chemotherapeutics (Fig 1).

**EGFR-Tyrosine Kinase Inhibitors**

EGFR is a transmembrane glycoprotein and a member of the ErbB receptor tyrosine kinase family. When EGFR is bound to the extracellular ligand, its intracellular domain activates several intracellular pathways implicated in cell proliferation, growth, and survival. Activating mutations in EGFR result in the constitutive activation of downstream signaling cascades that lead to uncontrolled cell proliferation in NSCLC.
Tyrosine kinase inhibitors inhibit EGFR by competitively binding to the adenosine triphosphate binding site on the intracellular tyrosine kinase domain of the receptor. This prevents the tyrosine kinase domain from activating intracellular signaling cascades that would otherwise amplify cell proliferation and survival. A positive EGFR-mutation status has been associated with a higher incidence of brain metastases in patients with advanced nonsquamous NSCLC, so EGFR mutations are especially relevant to the treatment of patients with NSCLC brain metastasis. The most common EGFR mutations observed in patients with NSCLC are in-frame LREA exon 19 deletions and exon 21 L858R point mutations.

Three generations of TKIs have been developed to treat EGFR-mutant NSCLC. The first-generation EGFR-TKIs are erlotinib and gefitinib, both of which inhibit EGFR by binding reversibly to the kinase domain of the receptor. In a Phase II study investigating gefitinib and erlotinib as first-line treatments for patients diagnosed with EGFR-mutant NSCLC, a high response rate of 83% was observed in patients with brain metastases treated with either drug, with a median progression-free survival (PFS) of 6.6 months and an overall survival of 15.9 months. Erlotinib has been observed to have a BBB permeation rate of 4.4% in patients following 4 weeks of treatment, with the level of response being proportional to the concentration of the drug in the CSF. In a Phase III study of gefitinib versus carboplatin-paclitaxel in patients with advanced, EGFR-mutated NSCLC and no history of chemotherapy, the gefitinib group had a PFS of 10.8 months, compared with 5.4 months in the traditional chemotherapy group. The gefitinib group also had a higher response rate of 73.7% compared with 30.7% in the carboplatin-paclitaxel group.

Despite the frequently promising initial response observed in first-generation TKIs, CNS tumor recurrence is common, and patients with EGFR-mutated NSCLC typically develop resistance to erlotinib and gefitinib after a median time of 8–10 months on therapy. These issues led to the creation of second-generation EGFR-TKIs such as afatinib which, unlike its predecessors, binds irreversibly to the EGFR. In a compassionate use program involving patients with NSCLC brain metastases who had failed on platinum-based chemotherapy and had progressed on a first-generation TKI, afatinib produced an overall cerebral response rate of 35%, making afatinib a potentially viable option for patients with NSCLC with brain metastases who have progressed on other chemotherapeutic regimens.

The secondary T790M mutation in the adenosine triphosphate binding site of EGFR is the most common mutation conferring acquired resistance to both first- and second-generation TKIs. This mutation is, therefore, partially responsible for the resistance to first- or second-generation EGFR-TKIs typically developing after 9–12 months of therapy. Osimertinib is an approved third-generation EGFR-TKI that can overcome the
T790M mutation in patients who have acquired resistance to first- or second-generation EGFR inhibitors (Fig 2). Osimertinib irreversibly binds to EGFR and targets both the initial activating EGFR mutation as well as the T790M mutation, while sparing wild-type EGFR. Osimertinib and other experimental third-generation TKIs, unlike their first- and second-generation predecessors, are also mutant-selective in that they exert more potent EGFR inhibition on EGFR-mutant cells as opposed to EGFR wild type cells. In the Phase III AZD9291 Versus Gefitinib or Erlotinib in Patients with Locally Advanced or Metastatic Non-small Cell Lung Cancer (FLAURA) study, patients with EGFR-mutant, advanced NSCLC were assigned to receive either osimertinib or standard EGFR-TKI treatment, which consisted of either erlotinib or gefitinib. A subanalysis of the patients in the FLAURA study with asymptomatic or stable CNS metastases found a 91% CNS objective response rate in the osimertinib arm compared with 68% for the standard EGFR-TKI arm. The median PFS was longer with osimertinib versus standard EGFR-TKI therapy (18.9 versus 10.2 months, respectively). Due to its superior BBB permeability and intracranial efficacy compared with older TKIs, osimertinib is projected to become the new first-line standard of care in the treatment of NSCLC brain metastases.

ALK-Tyrosine Kinase Inhibitors

Oncogenic rearrangements involving the ALK gene are the second most common mutation in NSCLC and account for approximately 3% of lung adenocarcinoma cases. The ALK gene codes for a transmembrane receptor tyrosine kinase that activates intracellular signaling cascades to promote cell proliferation. The oncogenic mutation of ALK consists of the fusion of 2 genes, most frequently ALK and EML4, to create a fusion of the ALK receptor with mutations in the extracellular domain that allow it to be activated in the absence of ligand binding. Like EGFR inhibitors, ALK-TKIs inhibit the ALK receptor by binding to the intracellular tyrosine kinase domain to prevent downstream signal transduction. Up to 70% of patients with ALK rearrangement with NSCLC develop brain metastases, making ALK gene rearrangements a prime target for TKIs that can penetrate the blood-brain barrier.

The first TKI developed to target ALK-rearranged NSCLC was crizotinib. In a retrospective analysis of the patients with ALK-rearranged asymptomatic brain metastases included in the PROFILE 1005 and 1007 clinical trials of crizotinib, crizotinib treatment yielded an intracranial disease control rate of 56% in previously untreated patients with asymptomatic brain metastasis and 62% in previously treated patients. The median intracranial time-to-progression was 7 months for the previously untreated group and 13.2 months for the previously treated group. However, eventual progression of pre-existing intracranial lesions or the development of new intracranial lesions during crizotinib therapy was common; this finding aligns with the observation that resistance is typically acquired within 12–14 months of the initiation of crizotinib therapy.

Alectinib is a second-generation ALK-rearranged TKI that is highly selective for ALK and also blocks the L1196M mutation that confers resistance to crizotinib, therefore providing a treatment option to patients whose intracranial disease has progressed on crizotinib (Fig 3). In an analysis of the 122 patients with NSCLC brain metastasis included in the Phase III ALEX study, alectinib had a higher overall CNS response rate than crizotinib (85.7% versus 71.4%, respectively) and significantly delayed CNS progression compared with crizotinib in patients with advanced ALK-rearranged NSCLC. Another Phase III trial, while showing the superiority of alectinib over crizotinib in terms of increased PFS and a lower incidence of CNS progression, also found that
alectinib has a superior toxicity profile compared with crizotinib.36 The ALK-TKIs ceritinib, brigatinib, and lorlatinib have also shown efficacy in NSCLC brain metastases.10,22-24,37-39

**ROS1 Tyrosine Kinase Inhibitors**

ROS1 is present in 1%–2% of cases of NSCLC and represents another targetable driver mutation in NSCLC brain metastases. ROS1 is a tyrosine kinase that can fuse with several different genes to create multiple ROS1-fusion receptors that lead to overactivation of downstream pathways within the cell that lead to uncontrolled proliferation. Like EGFR and ALK, ROS1 can be inhibited by TKIs that bind to the tyrosine kinase portion within the intracellular domain of the ROS1 receptor. Owing to the sequence homology shared by the ALK and ROS1 proteins, several ALK inhibitors have demonstrated efficacy in ROS1-rearranged NSCLC as well.40 For instance, crizotinib has demonstrated intracranial efficacy in ROS1-rearranged NSCLC (Fig 4). A Phase II trial of crizotinib in advanced ROS1+ NSCLC reported an overall response rate of 73.9% and a PFS of 10.2 months in the 23 patients with measurable CNS lesions at baseline.41 The ALK/ROS1/pan-TRK (pan-tropomyosin receptor kinase) inhibitor entrectinib was also studied in patients with ROS1-inhibitor-naïve NSCLC, and the authors reported an objective response in 7 of 11 patients with CNS metastases, as well as a CNS response rate of 71% based on 7 patients with evaluable lesions.42

A Phase II study of the ALK inhibitor lorlatinib has also reported an intracranial overall response rate of 56.0% in 25 patients with brain metastases, including a mixture of patients whose lesions were crizotinib-naïve and crizotinib-refractory.43 The ALK inhibitor ceritinib was also examined in ROS1-rearranged NSCLC in a small Phase II study, which reported intracranial disease control in 5 of the 8 patients with CNS metastases enrolled in the study.44

**TKI Toxicities and Risks**

Overall, EGFR-TKIs are well-tolerated, with fewer than 10% of patients discontinuing therapy due to adverse events.45 Toxic
deaths with these 3 EGFR-TKIs are rare, with pneumonitis being the most frequent cause in these agents. ALK inhibitors have been associated with lung toxicity, with the most frequent serious adverse event on ALK-inhibitor therapy being pneumonia.46 Crizotinib has also been associated with cardiotoxities such as bradycardia, complete heart block, QT prolongation, and pericarditis.47

No major CNS toxicities have been described for EGFR-, ALK-, or ROS1-TKIs to date. However, radiation necrosis and pseudoprogression have been described in some TKIs following SRS. Pseudoprogression refers to a short-lived increase in lesion size following therapy commonly due to the inflammatory reaction associated with radiation necrosis. Radiation necrosis presenting as pseudoprogression has been reported with alectinib following SRS, as well as with ceritinib following WBRT.48,49 A risk with pseudoprogression is mistakenly identifying it as true progression, which could result in unnecessary radiation or surgical CNS interventions. A high index of suspicion is therefore necessary in identifying radiation necrosis and pseudoprogression in patients with NSCLC brain metastases who have undergone radiation therapy followed by a TKI.

Posterior reversible encephalopathy syndrome (PRES) is a reversible neurologic syndrome that manifests as T2-hyperintense vasogenic edema on MR imaging and presents clinically with acute hypertension, seizures, vision abnormalities, and altered mental status.50 PRES is most frequently associated with antiangiogenic TKIs that inhibit the vascular endothelial growth factor because PRES is thought to be related to altered cerebral blood flow regulation. However, cases of PRES have been reported with the EGFR inhibitors erlotinib, gefitinib, and osimertinib, as well as the ALK inhibitor crizotinib.50 Classically, PRES involves the posterior parietal and occipital lobes but can vary in its distribution, with the most common presentation patterns being a dominant parietal-occipital pattern, superior frontal sulcus pattern, or interventions.

FIG 4. A 50-year-old woman with NSCLC with ROS1-CD74 (Cluster of Differentiation 74) fusion. The patient was initially started on crizotinib. The primary lung tumor was responsive to therapy, but there was progressive disease in the brain. Given the numerous brain metastases, the patient elected to undergo local therapy with whole-brain radiation. The pretherapy images (A and B) demonstrate multiple peripherally enhancing lesions with surrounding hyperintense FLAIR signal consistent with surrounding vasogenic edema (arrows in A and B). The posttherapy images (D and E) demonstrate a decrease in the size and surrounding edema of the brain metastases (arrows in D and E). The patient was also able to remain on crizotinib for systemic therapy, given the favorable response in the chest (arrows in C and F).
a linear holohemispheric watershed pattern.\textsuperscript{51} As the name suggests, PRES is most often reversible on discontinuation of the offending agent but may progress to hemorrhage or infarction if not promptly treated.

Despite the rare occurrence of radiation necrosis/pseudoprogression or PRES with TKI therapy, the overall risk associated with TKIs remains low. Not only are targeted therapies extending the lives of patients, but they are also doing so with less drug-associated morbidity. Therefore, the minimal toxicities associated with TKIs explain why these medications constitute a powerful breakthrough in the field of thoracic oncology.

**Definitive Treatments for Brain Metastases**

SRS was originally approved for treatment of up to 3 brain metastases but is demonstrating success in up to 10 lesions at some institutions.\textsuperscript{52-54} WBRT is typically used in patients with more numerous brain metastases. However, the neurodegenerative effects of WBRT can cause substantial morbidity and impair quality of life while having only a moderately beneficial effect on overall survival.\textsuperscript{55} To minimize neurocognitive damage from WBRT, novel radiation techniques are currently being explored. Hippocampal-sparing WBRT reduces the radiation exposure of the hippocampus while preserving the therapeutic benefit on the grounds that the hippocampus has a low frequency of brain metastasis occurrences, while WBRT-induced hippocampal damage is frequently associated with neurotoxicity.\textsuperscript{56} The neurodegenerative effects of WBRT have also incited effort to reduce the exposure of the brain to radiation via the expansion of SRS to patients with greater numbers of brain metastases in place of WBRT.\textsuperscript{56}

Responses of NSCLC brain metastases to radiation therapy have been shown to improve when radiation therapy is combined with immune checkpoint inhibitors (ICIs). The primary ICIs investigated in NSCLC are the anti-programmed cell death protein 1 (anti-PD-1) monoclonal antibodies nivolumab and pembrolizumab, and the anti-cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) monoclonal antibody ipilimumab. These ICIs block tumor cell binding to PD-1, programmed death-ligand 1, or CTLA-4 on immune cells, respectively, to prevent tumor cells from interacting with immune cells and evading the host immune response. The anti-programmed death ligand 1 monoclonal antibody atezolizumab has also been studied in NSCLC. Atezolizumab functions by binding to the programmed death-ligand 1 on tumor cells rather than immune cells, which, nevertheless, prevents tumor cells from interacting with immune cells and relieves immune cell suppression as is also seen with the anti-PD-1 and anti-CTLA-4 checkpoint inhibitors.

In a cohort study comparing 34 patients with ICI-naïve NSCLC brain metastasis receiving SRS alone with 17 patients receiving SRS concurrently with either nivolumab, pembrolizumab, or atezolizumab, the concurrent SRS/ICI cohort achieved a higher rate of CNS complete response (50% versus 15.6%) and a shorter median time to regression of brain metastases (2.5 versus 3.1 months).\textsuperscript{57} A larger retrospective study analyzed the responses to SRS combined with ipilimumab, nivolumab, or pembrolizumab in 260 patients with NSCLC (n = 157), melanoma (n = 70), or renal cell carcinoma (n = 33) with brain metastases. This study found that concurrent SRS/ICI use was associated with a longer overall survival (24.7 months) compared with both SRS alone and SRS with nonconcurrent ICI (12.9 and 14.5 months, respectively).\textsuperscript{58} One explanation for this synergy between ICI and SRS is that SRS facilitates increased release of proinflammatory signals, which augment the efficacy of ICI therapy. Notably, there was no increased risk of neurotoxicity with concurrent SRS/ICI therapy compared with SRS alone in either study.

**Imaging Features of EGFR-, ALK-, and ROS1-Mutated NSCLC Brain Metastases.** Brain metastases from NSCLC with EGFR, ALK, or ROS1 mutations most often appear similar to brain metastases from other common primary malignancies. They typically appear round with well-demarcated borders, exhibit a homogeneous or mostly solid pattern of enhancement, and have surrounding vasogenic edema with mass effect. Several recent case reports and case series have noted development of unique, nearly entirely cystic brain metastases in NSCLC treated with TKIs.\textsuperscript{59-61} In one case series including patients with ALK-mutated NSCLC, patients undergoing treatment with crizotinib developed brain metastases characterized by a thin rim of malignant cells with a large mucinous center.\textsuperscript{59} These lesions exhibited no-to-minimal peripheral enhancement, no surrounding vasogenic edema, and variable central FLAIR signal, likely related to the amount of proteinaceous content within the central mucinous portion of the lesion.\textsuperscript{59} In all 3 cases, the patient was asymptomatic despite the relatively large size of the lesions.\textsuperscript{59} Another case report described a patient with NSCLC with an unknown genetic mutation status who was treated with erlotinib and developed entirely cystic brain metastases in the subarachnoid space without any associated enhancement.\textsuperscript{60} This patient’s lesions responded very well to WBRT, whereas the other cases of cystic metastases after treatment with TKIs did not change significantly after radiation therapy. Another patient with EGFR-mutated NSCLC with a cystic brain metastasis after gefitinib progressed to death from mass effect and subsequent brain herniation.\textsuperscript{61}

**Guidelines for Response to Therapy**

Due to the poor response of brain metastases to traditional systemic chemotherapy drugs and short life expectancies of these patients in the past, patients with brain metastases were usually excluded from clinical trials. On the rare occasion that these patients were included, they were almost always required to demonstrate stable or treated CNS disease to be eligible for a trial. At least in part due to this limitation, there were no consensus guidelines for monitoring the response of brain metastases to systemic therapy. Trials often extended the Response Evaluation Criteria in Solid Tumors (RECIST)\textsuperscript{62} to evaluate the CNS metastases. Brain metastases were combined with all non-CNS diseases in all other organs as a single compartment for assessment of response. Furthermore, CNS metastases were rarely chosen as target lesions for the assessment of response to therapy. Grouping both CNS and non-CNS metastases as a single compartment is problematic. Traditional systemic chemotherapeutic agents have minimal penetration across the BBB; therefore, CNS lesions may progress while the non-CNS lesions are responding well to therapy. This outcome could force a patient
to be removed from the clinical trial even in the setting of a strong response to therapy outside the CNS. If the CNS disease and non-CNS disease are treated as separate compartments, locally progressive CNS disease can be treated with local therapy (eg, SRS, WBRT, and so forth), and the patient can remain on systemic therapy if effective in the non-CNS compartment.

Now that TKIs have shown much greater success in treating NSCLC brain metastases, there is a need for consensus guidelines to assess this response. The Response Assessment in Neuro-Oncology (RANO) group set out to address this problem and to create guidelines for monitoring the CNS response, including guidelines for both brain metastases and leptomeningeal disease.63,64 While other criteria such as RECIST and RECIST 1.1 are sometimes used in clinical trials that include patients with brain metastases, the RANO group is the only group to establish criteria pertaining specifically to brain metastases. The remainder of this article will serve as a guide for radiologists to appropriately assess the response to TKIs and other concurrent therapies of brain metastases according to the guidelines established by the RANO group.63

Obtaining Adequate Imaging
The first role of the radiologist in the assessment of brain metastases is to ensure that imaging is obtained with adequate parameters to enable accurate assessment of the response to therapy. MR imaging without and with gadolinium contrast is the ideal imaging technique for assessment of the response to therapy in brain metastases. According to the RANO group, CT with contrast may be used as a substitute in patients with contraindications to MR imaging and/or gadolinium or in areas with limited access to MR imaging; however, MR imaging should be performed whenever possible.63 Initial baseline imaging for comparison with further studies should be performed within 4 weeks of the start of treatment.63 For contrast-enhanced sequences, the MR imaging study should include a postgadolinium T1-weighted sequence with a maximum section thickness of 5 mm and no gap between slices. Ideally, especially in patients with smaller brain metastases or in trials that allow target lesions <1 cm in maximal diameter, a section thickness of ≤1.5 mm should be obtained when feasible. The additional MR imaging sequences in the recommended protocol suggested by the RANO group can be found in their article.63

Identifying and Measuring Target Lesions
For a lesion to be eligible as a target lesion, it must fit the RANO definition of a measurable lesion. A measurable lesion, as defined by the RANO criteria, is one with an enhancing component that is at least 10 mm in the longest diameter and 5 mm in the perpendicular, short-axis diameter that is visible on at least 2 consecutive slices and can be reproducibly measured.63 Of note, if the lesion is cystic or mixed cystic and solid, the cystic component should not be included in the measurement (Fig 5). Other features, including cystic-only lesions, leptomeningeal disease, dural-based disease, or lesions with ill-defined borders, are considered nonmeasurable disease (Fig 6).
Up to 5 measurable lesions should be identified as target lesions. If the patient has >5 measurable lesions, then the target lesions should be selected while giving preference to the following factors:

- Largest size.
- Increasing size on consecutive imaging studies.
- No prior local treatment, including an operation, SRS, and so forth.

After the target lesions are selected, the longest diameter (including only the enhancing nodular component of the lesion) should be recorded for each of the target lesions (Fig 5). The sum of the longest diameter of all target lesions should also be included in the report. If there are no lesions that meet the criteria for a measurable target lesion, then the patient is said to have nonmeasurable disease. The RANO group suggests that studies that include the assessment of the CNS response to therapy as their primary end point should not include patients with nonmeasurable disease.63

Additional factors including a description of morphologic features at initial imaging and the location of brain metastases may prove important in the era of precision medicine and targeted therapies, including TKIs. One study of 144 patients with EGFR-mutated NSCLC brain metastases treated with a first- or second-generation EGFR-TKI found that erlotinib achieved a longer PFS than gefitinib or afatinib only in patients who demonstrated poor prognostic factors on MR imaging such as tumor necrosis, rim enhancement, or tumor location in the frontal lobe, putamen, or cerebellum.65 In patients without these poor prognostic imaging characteristics, the PFS was comparable among erlotinib, gefitinib, and afatinib. This finding suggests that the initial imaging features of brain metastases may be useful in predicting outcomes in patients treated with novel targeted therapies.

Nontarget Lesions

Although the nontarget brain metastases do not need to be measured, they should be identified and included in the baseline report. This is important so that on subsequent follow-up imaging these lesions can also be assessed for progression and new lesions can be easily identified.

Evaluation of Lesions on Follow-Up Imaging

On follow-up imaging, the previously identified target lesions should be found and the largest diameter of each lesion as well as the sum of these unidirectional measurements should be recorded. Next, the nontarget lesions should be compared with measurements on prior studies and results of the evaluation of whether these lesions are increased, decreased, or unchanged in size should be recorded. Even if the target lesions are unchanged or decreased in size, if there is unequivocal and substantial progression of the nontarget lesions, the patient is still considered to have progressive disease. Finally, evaluation for any new brain metastasis should be performed. A clear statement of whether any new lesion is larger than 5 mm in diameter; visible in axial, sagittal, and coronal planes; and definitively not due to artifacts should be given in the report. In the case of questionable new lesions, the patient and oncologist may choose to continue therapy and obtain short-term follow-up imaging to re-evaluate the lesion in question.

Response to Therapy

The RANO criteria for response to therapy in brain metastases take several clinical factors into account, including changes in the corticosteroid dose and Karnofsky Performance Scale. Any significant increase in the corticosteroid dose or worsening in performance status that cannot be returned to baseline is consistent with progressive disease. Thus, unless the radiologist has access to both of these factors in the clinical history, it would be sensible to not include the definitive RANO criteria response in a report before discussing the patient with the referring oncology team. However, follow-up imaging is one of the major criteria for determining response to therapy according to the RANO criteria. From an imaging standpoint, the following are the criteria for each possible response to therapy:

- Complete response: complete resolution of all target and nontarget metastases with no new enhancing lesions.
- Partial response: decrease of at least 30% in the sum of the longest diameter of the target lesions compared with baseline, no change or decrease in the nontarget lesions, and no new lesions.

FIG 6. Examples of nonmeasurable disease in a patient with NSCLC brain metastases. A and B, Leptomeningeal metastatic disease in the right parietal lobe (white arrows). C and D, A different patient with small-cell lung cancer demonstrates completely cystic metastases (white arrows) that do not have a measurable solidly enhancing component. Both leptomeningeal disease and completely cystic metastases are considered nonmeasurable disease by the RANO criteria.
• Progressive disease: either an increase of at least 20% in the sum of the longest diameter of the target lesions from the nadir, unequivocal progression of nontarget lesions, or any new lesion.

• Stable disease: change in the sum of the longest diameter of the target lesions between 30% decrease and 20% increase, stable or improved nontarget lesions, and no new lesions.

Finally, the RANO group suggests that when reporting the response to therapy in patients with brain metastases, the CNS disease should be separated from the non-CNS disease. For non-CNS disease, the criteria established in RECIST 1.1 are recommended. This approach gives the flexibility for a trial to calculate both local CNS progression-free survival, non-CNS progression-free survival, and bicompartamental progression-free survival. Any of these outcomes could then be used as a primary end point for the trial. This flexibility is potentially beneficial in situations in which patients have local CNS progression and a favorable response in the non-CNS compartment for undergoing concurrent local CNS therapy and being able to stay in the trial.

When patients with NSCLC with brain metastases are on immunotherapy or have received local therapy with WBRT or SRS, new lesions alone may not constitute progression of disease and instead may be due to pseudoprogression. Pseudoprogression refers to a transient increase in tumor size on imaging shortly after treatment due to edema, necrosis, or immune cell infiltration. This initial radiographic growth can mimic true progression, but unlike true progression, it is subsequently followed by a reduction in tumor burden. Pseudoprogression after radiation therapy, especially SRS, is a well-recognized phenomenon in patients with brain metastases and has been shown to occur in up to 25%-30% of patients radiographically after SRS. Studies have also shown that it may occur in at a higher rate in NSCLC-associated brain metastases and that ALK mutation may be an added risk factor. Whether TKI therapy after SRS and/or WBRT is an added risk remains to be determined, and there is conflicting evidence on the topic. Before the use of second-generation ALK inhibitors, there were no reports of radiation necrosis during crizotinib use, including when it was used after patients had undergone SRS. There have been reports of radiation necrosis in patients on alectinib for ALK-mutated NSCLC months to years after radiation therapy. The RANO criteria do give some guidance in these situations; and often short-term follow-up, advanced MR imaging or PET, or biopsy must be performed to find a more definitive answer. Because TKIs are the focus of this article and they have not been established as a risk factor for pseudoprogression, an in-depth discussion of the guidelines pertaining to pseudoprogression will not be included. It is, however, important to keep in mind the possibility of pseudoprogression in these patients in the correct clinical and imaging scenario.

Studies evaluating the best imaging parameters to evaluate the response to TKI therapy have not yet been performed; however, many believe that the unidirectional method proposed by the RANO group may not be the best approach. Some have suggested that volumetric measurements are a more accurate measurement of changes in the number of cancer cells and better account for irregular shapes and cystic changes that frequently occur in brain metastases after treatment. A study on the treatment of melanoma brain metastases with SRS demonstrated that volumetric measurements were a better predictor of overall survival and the need for salvage WBRT compared with linear measurements. The RANO group does admit that volumetric measurements may prove to be a better method; however, they believe that there are too little data, availability, and standardization of volumetric measurements to recommend this over unidirectional measurements at this time. Furthermore, until volumetric measurements become more automated, they require higher cost and more effort to provide these measurements.

Unlike metastases treated with systemic cytotoxic chemotherapy, which typically demonstrate response to therapy as a shrinkage in size and decreased surrounding edema, metastases treated with SRS and WBRT often undergo more complex morphologic changes after treatment. Kang et al described early-stage changes, including central and perilesional edema with an occasional transient increase in the size of the enhancing brain metastasis days after SRS. These changes are followed by central cavitation and necrosis, often with indistinct enhancing borders. Finally, in the chronic stage, the lesion remains as a peripherally enhancing glial scar without surrounding edema. Because of these dynamic changes in size with difficult-to-measure centrally necrotic lesions, SRS has been the subject of study for many other imaging biomarkers to predict a response to therapy, including MR diffusion, MR perfusion (especially dynamic contrast enhancement), and PET studies using 3-deoxy-3-[18F]fluorothymidine or 2-[18F]fluoro-2-deoxy-d-glucose. One study concluded that the response to SRS treatment of brain metastases could be predicted as soon as 7–10 days after treatment using MR diffusion-weighted imaging. However, after many studies in search of the best biomarkers, factors as simple as whether an increase or decrease in diffusivity predicts a good response to SRS still remain controversial and are likely related to the timing of the follow-up studies. Because of the complexity of the physiologic changes happening at the cellular level in brain metastases after different types of treatment, accepting new, advanced MR imaging or PET biomarkers as a means to assess treatment response should not be haphazard. Furthermore, the use of different imaging biomarkers for different types of therapy may not be the best means of evaluation. Well-designed prospective studies on imaging biomarkers are needed to fully understand the best biomarkers for assessment of the response to TKIs and other methods of treatment.

In our experience, morphologic changes after TKIs more closely follow the systemic cytotoxic chemotherapy pattern of response including decreased size and surrounding edema of mostly solidly enhancing lesions. However, further studies need to systematically evaluate the pattern of morphologic changes in brain metastases after treatment with TKIs, especially because bizarre imaging appearances, including nearly entirely cystic brain metastases, have been described in patients undergoing treatment with TKIs. Also, if size criteria prove to be the best means of measuring response to therapy, unidirectional, bidirectional, and volumetric measurements should be compared to see which method best predicts early response or progression and long-term patient outcomes.

CONCLUSIONS

In recent years, several systemic therapies have been developed that demonstrate efficacy in NSCLC metastatic to the brain. It is essential for radiologists to be informed about these novel therapeutic agents to evaluate the intracranial response to therapy on imaging in patients undergoing treatment with these drugs. As the systemic chemotherapeutic options for metastatic NSCLC evolve, the assessment of the responses of brain metastases on imaging is becoming ever more critical for evaluating and managing patients with metastatic NSCLC.

REFERENCES


40. Ou SH, Zhu VW. CNS metastasis in ROS1+ NSCLC: an urgent call to action, to understand, and to overcome. Lung Cancer 2019;130:201–07 CrossRef Medline


49. Song YP, Colaco RJ. Radiation necrosis: a growing problem in a case of brain metastases following whole brain radiotherapy and stereotactic radiosurgery. Cureus 2018;10:e2037 CrossRef Medline


Advanced Multicompartment Diffusion MRI Models and Their Application in Multiple Sclerosis

D.A. Lakhani, K.G. Schilling, J. Xu, and F. Bagnato

ABSTRACT

SUMMARY: Conventional MR imaging techniques are sensitive to pathologic changes of the brain and spinal cord seen in MS, but they lack specificity for underlying axonal and myelin integrity. By isolating the signal contribution from different tissue compartments, newly developed advanced multicompartment diffusion MR imaging models have the potential to detect specific tissue subtypes and associated injuries with increased pathologic specificity. These models include neurite orientation dispersion and density imaging, diffusion basis spectrum imaging, multicompartment microscopic diffusion MR imaging with the spherical mean technique, and models enabled through high-gradient diffusion MR imaging. In this review, we provide an appraisal of the current literature on the physics principles, histopathologic validation, and clinical applications of each of these techniques in both brains and spinal cords of patients with MS. We discuss limitations of each of the methods and directions that future research could take to provide additional validation of their roles as biomarkers of axonal and myelin injury in MS.

ABBREVIATIONS: AD = axial diffusivity; Dax = intra-axonal diffusivity; DBSI = diffusion basis spectrum imaging; FF = fiber fraction; IVF = isotropic volume fraction; NDI = neurite density index (also Vnf, Vnf, or \( f_{\text{nf}} \)); NODDI = neurite orientation dispersion and density imaging; ODI = orientation dispersion index; RD = radial diffusivity; SMT = spherical mean technique; Vax = intra-axonal volume fraction

M

S is an inflammatory-degenerative disease of the CNS affecting approximately 2.3 million people, nearly 900,000 of whom are in the United States.1

Transient and focal inflammation with variable degrees of demyelination and axonal injury is the pathologic hallmark of MS lesions. The acute inflammatory phase resolves within a few weeks.2,3 Thereafter, some lesions undergo repair of injured myelin and axons,4 while others may fail to repair and evolve into chronic plaques featured by variable degrees of chronic demyelination and axonal injury.4 Axons do not survive chronic demyelination, leading to axonal atrophy, transection, and subsequent neuronal death. These neurodegenerative processes may also extend outside focal plaques to areas of normal-appearing WM and normal-appearing GM.5

Conventional MR imaging (T1WI and T2WI) is highly sensitive to focal tissue injury in MS but lacks specificity to axonal content. Furthermore, T1WI and T2WI do not capture disease processes outside focal lesions; hence, there remains a pressing need to identify an MR imaging biometric that is both sensitive and specific to lesional and nonlesional axonal injury and loss.6

Toward this effort, several advanced multicompartment diffusion-based models have been recently developed and applied in MS (On-line Figure). The advantage of these methods over the more commonly used DTI (On-line Figure A) is the ability to isolate the signal contribution from different tissue compartments, thereby increasing the specificity for tissue subtypes and associated injuries. These models include neurite orientation dispersion and density imaging (NODDI), diffusion basis spectrum imaging (DBSI), multicompartment microscopic diffusion MR imaging with the spherical mean technique (SMT), and several models enabled with high-gradient diffusion MR imaging.

While some of these methods are very close to clinical translation and even available on clinical scanners, others lag behind. The need for high-gradient diffusion MR imaging and relatively...
long scanning times is still an important limiting factor to a widespread application of some of these models.

In this review, we appraise the current literature on the physics principles, histopathologic validation, and clinical applications of each of these techniques in both brain and spinal cord imaging of patients with MS. We also discuss limitations of each of the methods and future directions for translational research to provide additional validation for their role as a proxy for axonal injury in MS.

**Neurite Orientation Dispersion and Density Imaging**

The NODDI model was developed to provide a clinically feasible technique for in vivo mapping of neurite orientation, dispersion, and density.7 The advantage of NODDI is the requirement of a moderate number of diffusion-weighted images acquired with only 2 b-shells (b-values), in contrast to other multicompartment models that need large gradient strengths or extensive scan times.7

**Principles of Physics.** NODDI is a multicompartmental model that distinguishes 3 microstructural environments: intracellular (or intra-axonal), extracellular (or extra-axonal), and CSF compartments (On-line Figure B). By affecting diffusion in a unique way, each compartment results in a separate MR imaging signal.

The intracellular component is represented by the intradendrites and intra-axonal spaces. Due to the highly restricted nature of diffusion in this space, the intracellular component is modeled as a set of geometric “sticks,” or infinitely thin cylinders, with diffusion completely restricted perpendicular to these sticks but unhindered along them. The diffusivity along a stick is assumed to be fixed at 1.7 × 10⁻³ mm²/s. This set of sticks can be arranged in many distributions of orientations, e.g., all coherently aligned in a single direction or more dispersed and spread over a range of orientations. The distribution of neurites (i.e., sticks) is described by an average direction and a concentration parameter, which describes the orientation dispersion around the main direction. The extracellular component represents the space around the neurites or axons, which is occupied by glial cells and neuronal cell bodies. In contrast to the restricted intracellular space, the diffusion in this space is “hindered” by neurites and is modeled as a simple anisotropic Gaussian distribution (as in DTI). The parallel and perpendicular diffusivities of the tensor in extracellular space are physically determined by the neurite morphology itself (i.e., the intrinsic diffusivity and the neurite dispersion).

The CSF compartment is modeled as isotropic Gaussian diffusion with a fixed isotropic diffusivity (3.0 × 10⁻³ mm²/s).

Thus, the NODDI model results in metrics describing orientation, shape, diffusivities, and fractions of the different compartments. As reported in On-line Table 1, NODDI produces a number of clinically relevant indices: 1) the intracellular volume fraction, represented by the neurite density index (referred to in the literature as NDI, \( V_{ic} \), \( V_{icw} \), or \( V_{icf} \)), that ranges from 0 to 1, denoting complete loss or full preservation of axons, respectively; 2) the neurite orientation dispersion index (ODI), which ranges from 0, representing perfectly coherently oriented WM structures, to 1, representing isotropically dispersed neurites; and 3) the isotropic volume fraction (often abbreviated as IVF, \( V_{iso} \) or \( f_{iso} \)), which can be interpreted as the voxel volume fraction of free water (i.e., CSF).

Fixing the intrinsic and isotropic diffusivities and deriving the extracellular diffusivities from the dispersion index reduces the number of free parameters and simplifies the fitting of the model.8-10

**Histopathologic Validation.** Grussu et al11 used the thoracic spine tissue of a patient with primary-progressive MS and a lumbar spine specimen of a patient with secondary-progressive MS, as well as of 2 healthy controls. The authors used circular variance as a histologic marker of neurite orientation variability, and the myelin staining fraction, astrocyte staining fraction, microglia staining fraction, and neurofilament staining fraction as markers of myelin, astrocytes, microglia, and axon contents, respectively. Correlations were seen between ODI and circular variance in the normal-appearing WM of both healthy controls (\( r = 0.84, P < .001 \)) and patients with MS (\( r = 0.60, P = .001 \)), as well as between ODI and the myelin staining fraction in MS samples (\( r = 0.40, P < .05 \)). Negative correlations were seen between NDI and circular variance (\( r = -0.78, P < .001 \)) in healthy controls and between NDI and the myelin staining fraction (\( r = 0.74, P < .001 \)), neurofilament staining fraction (\( r = 0.56, P < .01 \)), and the microglia staining fraction (\( r = 0.49, P < .01 \)) in patients with MS. No significant correlations were observed between IVF and histology markers. The authors found similar but less specific results with DTI-derived indices, in that several DTI-derived metrics correlated with histopathologic counterparts of myelin injury (like NODDI) as well as astrocyte/microglia infiltrates. In addition, none of the DTI indices showed a correlation with circular variance. These key findings highlight the ability of NODDI to show a different layer of pathology not captured by conventional DTI, e.g., changes in the complexity of neurite arborizations.

**Clinical Applications.** Several cross-sectional studies have been performed to test the clinical feasibility and the ability of NODDI to differentiate tissue injury in MS and its associations with patients’ disabilities. NODDI acquisition has been reported as ~22 minutes for whole-brain imaging (0.4 × 0.4 × 2 mm³ resolution, 2 b-shells, and 90 directions)12 and ~18 minutes for a single section of the cervical spinal cord (1.25 × 1.25 mm × 10 mm³, 2 b-shells, and 96 directions).13

We detail significant clinical findings reported in all published clinical studies in On-line Table 2. In summary, it was noted that when reaching a statistically significant difference, NDI/IVF values were always lower in lesions compared with normal-appearing WM,12,14-16 in normal-appearing WM compared with the normal WM of healthy controls (both in brains14,16,17 and spinal cord15,17), and in the spinal cord normal-appearing GM compared with the GM of healthy controls.13 These results were consistent across patients with relapsing-remitting12-17 or secondary-progressive MS.12,16,17 However, thus far, only values of NDI measured in both the cortex18 and the spinal cord17 were found to be associated with clinical disability measured with the Expanded Disability Status...
Scale\textsuperscript{19} in patients with both relapsing-remitting\textsuperscript{17,18} and secondary-progressive\textsuperscript{17,18} MS.

On the contrary, ODI measurements yielded more contradictory results because ODI measurements were higher in lesions compared with normal-appearing WM in some studies focused on patients with relapsing-remitting MS\textsuperscript{13,14,18} and lower in others focused on either relapsing-remitting or mixed MS populations.\textsuperscript{12,15} Similarly, while Schneider et al\textsuperscript{15} found lower levels of ODI in the brain normal-appearing WM of patients with relapsing-remitting MS; By et al\textsuperscript{13} reported opposite trends in the spinal cord of a similar population of patients. Only cortical ODI measures were associated with clinical decline in patients with both relapsing-remitting\textsuperscript{14,18} and secondary-progressive\textsuperscript{18} MS.

**Summary of Advantages and Limitations.** NODDI appears to offer a clinically feasible technique that adds a layer of pathologic sensitivity and specificity to the detection of MS disease. A high intersubject variability is observed with ODI, thus making its interpretation complex. This finding is not surprising because by depicting the degree of tissue organization, ODI is a more granular, and therefore variable, measure of disease variability. It is proposed that ODI is more informative if used to assess longitudinal changes of individual patients rather than to assess comparisons between patients. On the contrary, although not always sensitive to disease, NDI/Vax appears to capture some degree of change in axonal content, which is also measurable in perilesional tissue.\textsuperscript{16}

A few technical limitations need to be acknowledged regarding this model: 1) NODDI simplifications limit the model to simple fiber geometries, 2) the fixed diffusivities may not be true in altered physiology of the disease, and may bias measurements, and 3) debate remains on the relationship (and absolute values) of the intrinsic intra- and extracellular diffusivities.\textsuperscript{8-10}

Thus, further work is needed to untangle these technical challenges, and to investigate the efficacy of NODDI-derived metrics as biomarkers of neurodegeneration, by assessing the sensitivity to tissue injury and clinical measures in larger cohorts of patients followed longitudinally over time.

**Multicompartment Microscopic Diffusion MR Imaging with the Spherical Mean Technique.**

The SMT estimates microscopic features specific to the intra- and extraneurite compartments in the CNS. The use of the spherical mean technique is the most relevant advantage of this method, in that the SMT minimizes the confounding effects derived from axonal fiber crossings, curving, and orientation dispersion. SMT is suitable for clinical applications that require information on axonal volume fraction as well as axonal directions if one obtains spherical deconvolution.

**Principles of Physics.** SMT requires $\geq 2$ b-shells (each with multiple diffusion directions) and assumes diffusion-weighted signals arising from intra- and extra-axonal spaces (On-line Figure C), resulting in the apparent intra-axonal volume fraction ($V_{\text{ax}}$) and the apparent extra-axonal volume fraction. Instead of approximating the whole intra- and extra-axonal spaces using single diffusion tensors, SMT assumes that signals arise from many "per-axon" diffusion tensors in arbitrary directions.\textsuperscript{20} By averaging the same b-value diffusion MR imaging signals over all directions, one can remove the orientation dependence of signals. Moreover, SMT assumes the following: 1) the intra-axonal radial diffusivity (RD) is zero due to the very small axon sizes in the CNS, and 2) the axial diffusivity (AD) is the same in both intra- and extra-axonal spaces due to the lack of restriction per axon. By such means, SMT can separate intra- and extra-axonal signals and 2 independent microstructural parameters, ie, the $V_{\text{ax}}$ and the apparent intra-axonal diffusivity ($D_{\text{ax}}$) can be fit from the data. Other microstructure parameters, such as extra-axonal RD, can be calculated from $V_{\text{ax}}$ and $D_{\text{ax}}$ (On-line Table 1).

**Histopathologic Validation.** There are no studies in either humans with MS or animal models of MS that validated SMT-derived metrics against histopathologic counterparts; however, there is histopathologic validation of SMT in other models of diseases. Kaden et al\textsuperscript{21} studied the mouse model of the tuberous sclerosis complex and validated microscopic features specific to intra- and extraneurite compartments. The tuberous sclerosis animal model is more suitable for detecting CNS axonal injury than the MS model of experimental autoimmune encephalomyelitis because it is free from the inflammatory component that contaminates data from experimental autoimmune encephalomyelitis. $V_{\text{ax}}$ and $D_{\text{ax}}$ were measured in several WM tracts of the brain, including the genu, midbody, splenium of the corpus callosum, and anterior commissure, in both tuberous sclerosis and control mice. $V_{\text{ax}}$ values were lower ($P<.05$) and $D_{\text{ax}}$ values were higher ($P<.05$) in the WM of animals with tuberous sclerosis compared with controls. Furthermore, a decrease in the myelinated axon fraction (with the myelin fraction excluded, $P<.001$) was mirrored by a reduction of the MR imaging–based $V_{\text{ax}}$. Although obtained with a different disease model, the provided validation against axonal histology is fundamental. It is applicable to any condition affecting myelin and axonal integrity and supports the ability of SMT to quantify axonal content without artifactual effects from fiber-crossing and orientation dispersion.

**Clinical Application.** Only 2 studies investigating the feasibility and applicability of SMT in the brains\textsuperscript{22} and spinal cords\textsuperscript{23} of patients with MS have been performed thus far. The authors reported acquisition times between $\sim 18$ minutes\textsuperscript{23} (single-section C-spine, $1.25 \times 1.25 \times 10$ mm resolution, 2 b-shells, with 32 and 64 directions) and $\sim 22$ minutes (full brain coverage, $0.4 \times 0.4 \times 2$ mm resolution, 90 directions).\textsuperscript{22} The implementation of a recently developed multiband technique can remarkably reduce the scan time while keeping similar image quality. For example, a multiband factor of 3 can reduce the total scan time of SMT from $\sim 22$ minutes to $\sim 10$ minutes.

In brains\textsuperscript{22} and spinal cords\textsuperscript{23} of patients with MS, $V_{\text{ax}}$ and $D_{\text{ax}}$ (brain only\textsuperscript{23}) differed between lesions and normal-appearing WM. Decreased values of $V_{\text{ax}}$ were also observed in normal-appearing WM compared with normal WM of healthy controls (spinal cord only\textsuperscript{23}). Reduction in acquisition schemes by 50% did not affect the sensitivity of SMT in distinguishing all 3 tissue subtypes.\textsuperscript{23}

Grussu et al\textsuperscript{24} studied the diffusion time-dependence of IVF, ODI, $V_{\text{ax}}$, and $D_{\text{ax}}$ in 3 healthy controls and quoted percentage
changes of these parameters in the diffusion time range of 29–76 ms. Increases in IVF (4.1%) and $V_{ax}$ (2%) and decreases in
the ODI (−16%) and $D_{ax}$ (−5.4%) as a function of the diffusion
time were noted in the motor columns of the spinal cord.
Furthermore, increases in the IVF (3.7%) and decreases in the
ODI (−15%), $V_{ax}$ (−5.1%), and $D_{ax}$ (−11.5%) were measured
in the WM of the sensory columns of the spinal cord. The
results suggest a diffusion time-dependence of NODDI- and
SMT-derived metrics, which will provide new opportunities to
optimize these methodologies for spinal cord imaging.

Summary of Advantages and Limitations. By taking the av-
erages of diffusion signals over all directions and making no
assumptions about the orientation of fibers, SMT removes the
orientation dependence of diffusion MR imaging signals.
These elements simplify data analysis and improve the SNR
because they remarkably reduce the interscan variability.
Histopathologic evidence supports the ability of SMT to isolate
signal contributions from axons. In clinical studies, both $V_{ax}$
and $D_{ax}$ have shown the potential to differentiate tissue sub-
types in the brains and spinal cords of patients with MS. Even
by reducing the acquisition time by 50%, the SMT metrics
were found to be equally sensitive in differentiating these tissue
subtypes.25 Both metrics, however, have failed to show correla-
tion with the clinical measure of disease. This issue may be
related simply to the relatively small sample size of cross-section-
tal studies performed thus far. Larger studies with longitudi-
unal designs are warranted to elucidate the role of SMT-
derived metrics as biomarkers of neurodegeneration in MS.

Diffusion Basis Spectrum Imaging

DBSI was proposed to resolve multiple tensor-like populations of
water that may arise not only in healthy tissue but also from axon
injury, inflammation, and demyelination.25 Despite solid histopa-thologic validation studies, clinical applications of this technique
are limited.

Principles of Physics. DBSI models diffusion as a combination of
multiple discrete anisotropic tensors and a spectrum of isotropic dif-
fusion tensors (On-line Figure D). The discrete anisotropic tensors
are intended to represent myelinated and unmyelinated axons ori-
ented in varying directions. The isotropic tensors, then, represent
the integration of multiple pools of water, typically separated into
the restricted spectrum from 0 ≤ apparent diffusion coefficient ≤
$0.3 \times 10^{-3}$ mm$^2$/s, which may reflect cellularity. The nonrestricted
isotropic diffusion spectrum from $0.3 \times 10^{-3}$ mm$^2$/s < apparent
diffusion coefficient < $3.0 \times 10^{-3}$ mm$^2$/s reflects extracellular
edema and CSF. Thus, in DBSI, the diffusion signal is modeled as a
summation of many diffusion tensors and an integration of free
water with varying diffusivities, involves fitting many free variables,
and uses a complex signal-fitting approach.

DBSI results in estimates of several anisotropic diffusion com-
ponents (representing intra-axonal water molecules), each
weighted by signal intensity fractions (referred to as their fiber
ratios), oriented in the direction of the fiber population, and
associated with their own parallel and perpendicular diffusiv-
ity.25 Thus, just like DTI, DBSI results in metrics of RD, AD,
fractional anisotropy, and the fraction of anisotropic compo-
nents (called the fiber fraction [FF]) (On-line Table 1). In addi-
tion, the fraction of water molecules of highly restricted
isotropic structures can be measured (referred to as the re-
stricted isotropic diffusion fraction or as the cell ratio in the
original implementation), as well as that of the less restricted
isotropic water (referred to as the water ratio).

Histopathologic Validation. Shirani et al26 studied histopa-thologic correlates of DBSI in comparison with DTI in a single tume-
factive lesion of a patient with newly diagnosed MS. DTI-derived
AD and RD and DBSI-derived RD were increased, while DTI-
and DBSI-derived fractional anisotropy were decreased in the
lesion compared with the normal WM. DBSI-derived FF was also
minimally decreased in the lesional area. Histopathologic analy-
ses showed features of demyelination with increased inflamma-
tion and cellularity, but little axonal damage. Overall, the authors
concluded that DBSI more selectively depicted demyelination
and cellular increase and was not affected by the presence of
edema, which, instead, biased the DTI-derived AD values.

Wang et al27 studied postmortem examinations of the cervical
spinal cords of 2 patients with relapsing-remitting MS and 1 with
primary-progressive MS. The authors found correlations between
FF and silver stain ($r = 0.7–0.8$, $P < .001$, in 3/3 samples); DBS-
derived RD and Luxol fast blue stain ($r = −0.4–−0.84$, $P < .05$, in
3/3 samples); and restricted fraction and hematoxylin stain,
reflecting the number of nucleated cells ($r = 0.3–0.8$, $P < .05$, in
2/3 samples). No significant correlations were seen between
DBSI-derived AD and silver stain.

To assess the performance of DBSI and DTI metrics in the
presence of CSF contamination, Wang et al27 compared values
derived from regions in the center of the corpus callosum (repre-
senting pure axonal fibers) and from areas of the WM bordering
the lateral ventricles (representing CSF contamination) in 5
healthy controls. DBSI-derived AD and RD for each of these vox-
els were indistinguishable, while CSF contamination substantially
altered DTI-derived AD and RD.

Furthermore, DBSI and DTI values were derived from regions at
the crossing between the corona radiata and the corpus callosum;
and within the corona radiata and the corpus callosum in 5 healthy
controls, to assess the effect of fiber-crossing. DBSI AD and RD val-
ues were similar, while DTI AD and RD values differed between the
corpus callosum and regions at the crossing between the corpus cal-
losum and the corona radiata. DBSI AD values did not differ
between the pure corona radiata and corona radiata fibers extracted
from the region of crossing with the corpus callosum.

In conclusion, contrary to DTI, DBSI-derived AD and RD values
were not influenced by CSF contamination and fiber-crossing.

Clinical Application. Shirani et al28 applied DBSI in patients with
relapsing-remitting MS ($n = 22$, disease duration = 2.0–18.8 years),
secondary-progressive MS ($n = 16$, disease duration = 15.4–39.0 years),
and primary-progressive MS ($n = 17$, disease duration = 6.5–
18.3 years). They used recursive partitioning, a nonparametric decision
tree–based regression and classification, to assess the ability of DBSI-
derived metrics to classify patients with different disease subtypes. FF
and restricted isotropic fraction classified 35/55 patients (63%) into the
correct MS subtype. Similarly, FF, nonrestricted isotropic fraction and RD measured in the normal-appearing WM of the corpus callosum correctly classified 37/55 patients (67%). No differences were seen in the FF, isotropic restricted fraction, isotropic nonrestricted fraction, RD and AD between WM lesions and normal-appearing WM among patients with relapsing-remitting, secondary-progressive, and primary-progressive MS.

Summary of Advantages and Limitations. The performance of DBSI has been extensively validated against histopathology in humans, and DBSI was shown to perform better than conventional DTI in the presence of edema. Accordingly, DBSI has shown promising results for assessing axonal quantity (DBSI-AD and -FF) while correcting for signal impurity derived from fluid contamination and fiber-crossing. Clinical applications of this methodology are in their infancy. Initial studies support the notion that DBSI correctly classifies the phenotype of about two-thirds of patients with MS. Additional studies are necessary to further elucidate its role in vivo. Similarly, sensitivity and specificity validations are lacking in simulated and physical phantom environments, and further investigation of precision and bias is necessary.

High-Gradient Diffusion MR Imaging

Principles of Physics. This method is similar to the previously developed AxCaliber method, which maps the mean axon diameter and water fractions of axons and CSF. It assumes diffusion signals arising from 3 nonexchanging compartments: restricted diffusion in the intra-axonal space, hindered diffusion in the extra-axonal space, and free diffusion in CSF (On-line Figure E). Unlike NODDI and SMT, this method assumes that axons have nonzero diameters and typically require diffusion data with multiple diffusion times and multiple b-values. By assuming that axons are cylinders, analytic models can be fit to diffusion data to extract quantitative microstructural information, such as mean axon diameter and apparent intra-axonal volume fraction (On-line Table 1). Note that the obtained mean axon diameter is volume-weighted rather than an arithmetical mean value, the latter predominantly used previously in the pathology literature. It is challenging to fit comprehensive analytic equations to diffusion-weighted signal data from a voxel consisting of axonal fibers in >1 direction (such as fiber-crossing, curving, and dispersion). Therefore, this method was initially used only in regions with 1 dominant axonal fiber direction, such as the corpus callosum or the spinal cord.

A recent report, however, has extended the application of this method to different WM tracts in the brain. This extended method assumes that axonal fibers are oriented in a principal direction, which makes it possible to average DWI signals perpendicular to this principal fiber orientation. By such a means, a volume-weighted axon diameter index can be obtained in all white matter tracts in the brain, as shown in On-line Figure E.

There are existing debates on the ability of this method to fit the axon diameter. Early studies of AxCaliber in vivo were found to overestimate axon diameters in the corpus callosum in healthy subjects are still overestimated compared with the histologic findings. Moreover, a recent study showed that the extra-axonal diffusion dominates the overall diffusion signals in the typical diffusion time range; thus, it should be used with caution to interpret the fitted axon diameter in vivo. The high-gradient diffusion method requires special gradient coils with gradient strengths up to 300 mT/m, which is much stronger than the 30–80 mT/m gradient strength on typical commercial MR imaging systems. This requirement is an important obstacle to clinical translation.

Clinical Application. Huang et al studied the clinical feasibility and applicability of high-gradient diffusion MR imaging in MS. The authors reported a scanning time of ~51 minutes for a whole-brain acquisition at a 2-mm isotropic resolution. Six patients with relapsing-remitting MS (disease duration, 1–11 years) and 6 healthy controls were initially imaged with the aim of characterizing axonal disease in MS lesions and normal-appearing WM. Increased axonal diameter, decreased restricted fraction, increased free water fraction, and reduced mean axon density were described in lesions compared with normal-appearing WM, but not between normal-appearing WM and normal WM.

Recently, the same group reported data from a larger patient population, illustrated in On-line Table 2. Similarly, Santis et al found increased axonal diameters and decreased restricted volume fraction/axonal density in lesions compared with the normal-appearing WM of the corpus callosum, in normal-appearing WM compared with normal WM, as well as in the normal-appearing WM of patients with secondary-progressive MS compared with that of patients with relapsing-remitting MS.

Summary of Advantages and Limitations. Clinical studies show the potential for high-gradient diffusion MR imaging to differentiate tissues with different degrees of axonal pathology. The physics underlying the use of high-gradient diffusion MR imaging limit its ability to quantify axon diameter in the presence of fiber-crossing. Furthermore, no histopathologic validation has been performed, a factor that currently hampers extrapolating the full significance from the in vivo data. It is interesting to note that the axonal diameter measured using this technique were larger in patients as compared to that of healthy controls. The finding may reflect the well-known pathologic course of axon death. Axons undergo remodeling with initial swelling before atrophy and cell death. Thus, the time when patients are imaged may determine the direction of group differences one can find.

CONCLUSIONS

There is an urgent need to discover a biometric of neurodegeneration and repair, which can be used to untangle MS disease progression before it manifests as an irreversible disability. Advanced, multi-shell-based diffusion methods have the potential to add pathologic specificity to imaging findings, and hence serve as biomarker of neurodegeneration and neuroregeneration. Additional work is required, however, to translate these methods to clinical application.

Most of these techniques have 2 main technical advantages. First, they use conventional, off-the-shelf diffusion MR imaging pulse sequences, which are widely available on clinical MR imaging.
scanners. This use removes any obstacle toward their translation to routine clinical imaging. Second, most of these diffusion methods have their data analysis codes open to the public, making it easy to implement independent studies. However, technical improvements based on the implementation of multibands are warranted to enhance clinical reliability. These changes will aim for faster acquisitions with higher signal-to-noise ratio. Effort will also be devoted to the implementation of postprocessing techniques for the spinal cord and the development of atlases that can more accurately couple these advanced imaging metrics with anatomic landmarks.

From a histopathologic validation standpoint, the provided evidence is solid but remains limited because it is based on a small number of samples or animal studies using a different model of neurodegeneration. Ideally, postmortem studies on human brains are the closest ones for studying human disease. These studies, however, have the limitations of formalin fixation, which substantially alters diffusivity properties and shortens the T1 and T2 relaxation times of tissues compared with their in vivo state. Furthermore, there is a time-dependent effect of formalin fixation on diffusivity. Data show that scanning different brains at a consistent postfixation interval leads to different absolute measurements between samples, which could, in turn, lead to erroneous conclusions.

Histopathologic validation of lesions requiring biopsy is also an efficient method but has the caveat that lesions undergoing biopsy are usually atypically aggressive. While not resembling, in full, MS pathology, these evaluations still offer a solid validation method for the imaging-pathologic correlates. Similarly, animal studies, while again offering a disease model that may not resemble, in full, that of MS, do offer opportunities to validate histologic counterparts of imaging data. It is unlikely, however, that individual centers will be able to offer a full spectrum of validation; and to this end, collaborative efforts are certainly crucial to delineate the solidity and reproducibility of individual findings.

From a clinical standpoint, it appears that metrics derived from each of the discussed techniques are sensitive to different degrees of pathology in MS for both the brain and spinal cord. A common criticism of those methods is that changes measured in different tissue types using metrics derived from advanced diffusion models may, at times, be like those measured using conventional DTI or other quantitative techniques. Even if the criticism is valid, given the physics principles underpinning each of those techniques, advanced multicompartment diffusion MR imaging models offer a novel and superior way to visualize and quantify pathology. Thus, even if the effect size is similar, all these models add a substantial layer of pathologic specificity to disease detection.

Whether this increased pathologic specificity corresponds to an improved ability to explain disability remains to be assessed. While data thus far do not seem to necessarily suggest so, none of the presented studies have been powered to address this question. It is also likely that across-time changes of these proposed measures, rather than single-time-point assessments, are reflective of across-time disease progression. Additional larger clinical studies to address this question are the next step. It is indeed likely, and most relevant clinically, that individual patient changes across time, rather than between-subject cross-sectional comparisons, are going to be both more informative and explicative of the pathologic and clinical evolution of MS. This consideration is particularly relevant in lieu of the fact that axonal size and volume change with time as a function of the pathologic processes. Swelling and enlargement often at times precede shrinkage and transection, and this feature must be taken into consideration when comparing data from different patient cohorts or even when comparing patients with healthy controls in cross-sectional evaluations.

Last, but no less important, intercenter reliability measures must be estimated. Lack of intercenter and interscanner consistency is an important drawback and will consequently hinder clinical translation of any biomarker.

Future research based on cooperative effort aiming to fill the gap in the histopathologic, clinical, and technical knowledge as detailed above is most warranted.

ACKNOWLEDGMENTS

The authors would like to thank Dr Seth Smith for the design and conceptualization of the On-line Figure and Dr Karina Ciccone for thoroughly reviewing the manuscript for grammar and appropriateness of the delivered message.

Disclosures: Francesca Bagnato—UNRELATED: Speaker honoraria from EMD-Serono and Novartis. She serves/as sit of multicenter studies sponsored by EMD- Serono and Novartis. Paid to the individual.

REFERENCES

Methylphenidate Effects on Cortical Thickness in Children and Adults with Attention-Deficit/Hyperactivity Disorder: A Randomized Clinical Trial


ABSTRACT

BACKGROUND AND PURPOSE: Although methylphenidate is frequently used to treat children with attention-deficit/hyperactivity disorder, it is currently unknown how methylphenidate affects brain development. In a randomized controlled trial, we investigated whether the cortical effects of methylphenidate are modulated by age.

MATERIALS AND METHODS: Between June 1, 2011, and June 15, 2015, we conducted a randomized, double-blind, placebo-controlled trial (Effects of Psychotropic Drugs on Developing Brain-Methylphenidate) in 99 males with attention-deficit/hyperactivity disorder (according to Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, criteria) from referral centers in the greater Amsterdam area in the Netherlands. The trial was registered on March 24, 2011 (identifier NL34509.000.10) and subsequently at the Netherlands National Trial Register (Identifier NTR3103). Participants (first enrolled October 13, 2011) were 10–12 years or 23–40 years of age and randomized to treatment with either methylphenidate or a placebo for 16 weeks. Our main outcome was a change in cortical thickness in predefined ROIs as measured by MR imaging pre- and posttreatment.

RESULTS: We observed a time × medication × age interaction (F[1,88.825] = 4.316, P < .05) for the right medial cortex ROI, where methylphenidate treatment yielded less cortical thinning in children, but not in adults or the placebo groups.

CONCLUSIONS: Our finding that the effects of methylphenidate on right medial cortical thickness differ between children and adults infers that the drug affects gray matter development in this brain region. This warrants replication in larger groups with longer follow-up to determine whether this effect can also be observed in other cortical brain regions and whether it may have long-term consequences.

ABBREVIATIONS: ADHD = attention-deficit/hyperactivity disorder; MNI = Montreal Neurological Institute; RCT = randomized controlled trial.

The psychostimulant methylphenidate is among the drugs most frequently prescribed to children for treatment of attention-deficit/hyperactivity disorder (ADHD). Randomized controlled trials (RCTs) have shown that methylphenidate suppresses children’s physical growth during treatment, but they have not included measures of brain development. However, longitudinal brain maturation studies using MR imaging indicate that children with ADHD, on average, deviate from controls in their brain developmental trajectories, including regionally thinner cortices and lower basal ganglia volumes. Recently, highly powered cross-sectional studies on subcortical volumes and cortical thickness and surface area have shown ADHD case-control differences, with higher effect sizes in children than adults. In these observational cross-sectional studies, the authors found no influence of psychostimulant medication on subcortical volumes or cortical thickness of the areas showing group differences, but in smaller surface areas of regions of the frontal cortex in children taking the medication. Prospective studies on children with ADHD using-versus-not using psychostimulants have yielded

Received February 1, 2020; accepted after revision April 8.

From the Department of Psychology (K.B.W., I.A., D.A.R., A.B., A.M.F.) and Diagnostic Physics (I.G., A.B.), Oslo University Hospital, Oslo, Norway; Departments of Radiology and Nuclear Medicine (K.B.W., A.M.F.) and Diagnostic Physics (I.G., A.B.), Oslo University Hospital, Oslo, Norway; and Department of Radiology and Nuclear Medicine (A.S., A.B., L.R.), Amsterdam UMC, Academic Medical Center, Amsterdam, the Netherlands.

This study was funded by faculty resources of the Academic Medical Center, University of Amsterdam, and grant 11.32050.26 from the European Research Area Network Priority Medicines for Children (Sixth Framework Programme).

Clinical Trial registry name and registration number: Central Committee on Research Involving Human Subjects (an independent registry, identifier NL34509.000.10) and the Netherlands National Trial Register (identifier NTR3103).

Please address correspondence to Liesbeth Reneman, MD, PhD, Amsterdam UMC, Academic Medical Center, Department of Radiology and Nuclear Medicine, G1-222 Meibergdreef 9 1105 AZ Amsterdam, the Netherlands; e-mail: l.reneman@amsterdamumc.nl

Indicates article with supplemental on-line photo.

Indicates article with supplemental on-line protocol.

Indicates article with supplemental on-line appendix, photo, and trial protocol.

http://dx.doi.org/10.3174/ajnr.A6560
FIG 1. ROIs investigated. Cortical ROIs were determined on the basis of the MNI coordinates of the vertices corresponding to peak group differences in the observational prospective study of psychostimulant treatment and the developing cortex by Shaw et al. Labels were created at the corresponding vertices on the FreeSurfer average surface. The labels were dilated 15 times, resulting in hexagon-shaped ROIs, covering approximately 550 mm² when transformed to the vertices corresponding to peak group differences in the observational prospective study of psychostimulant treatment and the developing cortex by Shaw et al. The ROIs chosen for investigation were the left frontal cortex ROI in lateral (A) and frontal (B) view; 2) the right medial cortex ROI in the medial view; and 3) the right posterior cortex ROI in the lateral view.

some evidence for medication-associated differences for cortical thickness changes, but an absence of such evidence for basal ganglia trajectories. However, such observational studies are limited by possible selection effects. Moreover, the cortical differences observed have been in terms of greater thinning of the cortex in the right medial frontal motor region, the left middle/inferior frontal gyrus, and the right posterior parieto-occipital region in children with ADHD not taking psychostimulants, and different effects of psychostimulant medication can be expected during development, when the brain changes at a higher rate than in adulthood.

We recently reported age-dependent effects of methylphenidate on changes in cerebral blood flow of the dopaminergic system (our predefined primary aim), as measured by pharmacologic MR imaging, in young-versus-adult patients with ADHD. Because our RCT is the first study that examines the effects of methylphenidate on brain development, this means that there are only limited and indirect data available to perform a sample-size calculation. Our goal for this research was to be able to detect differences in the age-dependent effect of methylphenidate on the brain if these differences are in the magnitude of a standardized effect size of 1.25 (see also On-line Research Protocol, page 18). A secondary aim of the Effects of Psychotropic Drugs on Developing Brain (ePOD) study is to assess the effects of methylphenidate treatment on cortical changes in children and adults with ADHD in a double-blind RCT. Due to ongoing development of the dopaminergic system during childhood and adolescence, we hypothesized that methylphenidate treatment would be associated with reduced cortical thinning (rate of change) in the left frontal, right posterior, and right medial cortices and that this would be observed in children only, not in adults. In addition to testing this hypothesis, we also explored possible cortical effects in a vertex-wise whole-brain analysis and possible relations between cortical change and change in ADHD symptom severity.

MATERIALS AND METHODS

Trial Design

The ePOD–Methylphenidate RCT was a 16-week double-blind, randomized, placebo-controlled, multicenter trial of the use of methylphenidate and a blinded end point evaluation in psychostimulant treatment–naïve patients with ADHD. The length of the trial was dictated by the waiting list for treatment. In the Netherlands at the time of the execution of this trial, a waiting list of 17 weeks was very common before (child and adolescent) psychiatric evaluation could take place and treatment could commence. It is in this timeframe that the current study was executed, to ensure only a very small or no delay in the placebo group before active treatment could take place. The timeframe was thus chosen so that patients treated with a placebo would not be harmed by study participation (for further details, see the On-line Research Protocol). The primary outcome measure of the ePOD–Methylphenidate trial was to report on the effect modification by age of the methylphenidate treatment on the outgrowth of the dopaminergic system.

Here, we report our secondary outcome. A prior article dealt with the effect of methylphenidate on white matter development, whereas this article assesses the effects on the rate of change in cortical thickness of selected ROIs assessed using structural MR imaging in exactly the same sample of children and adults with ADHD, randomly assigned to either placebo or active treatment with methylphenidate, pre- and posttreatment. The selection of ROIs was based on prior results from an observational prospective study of psychostimulant treatment and the developing cortex by Shaw et al. The ROIs chosen for investigation are shown in Fig 1. The study protocol applied the code of medical ethics and was registered by the Central Committee on Research Involving Human Subjects (an independent registry) on March 24, 2011 (identifier NL34509.000.10) and subsequently at the Netherlands National Trial Register (identifier NTR3103), with enrollment of the first patient on October 13, 2011. In addition, the institutional review board of the Academic Medical Center approved the study. The full protocol is described elsewhere. The trial ended on June 15, 2015, and was monitored by the Clinical Research Unit of the Academic Medical Center, University of Amsterdam, the Netherlands. All patients and parents or legal representatives of the children provided written informed consent.

Participants

Participants were psychostimulant treatment–naïve, 50 boys (10–12 years of age) and 49 men (23–40 years of age), diagnosed as having ADHD and recruited through clinical programs at the Department of Child and Adolescent Psychiatry at Triversum (Alkmaar, the Netherlands), De Bascule Academic Center for Child and Adolescent Psychiatry (Amsterdam), and PsyQ mental health facility (The Hague). As laid out in the protocol, the cutoff point of 10–12 years of age for the pediatric population was chosen because the peak prevalence of ADHD has been found to be
10 years of age, and also because several MR imaging parameters show greater age differences in early than later childhood.

One hundred thirty-one individuals were screened for eligibility. All included met strict criteria for ADHD (all subtypes) according to the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) and were diagnosed by an experienced psychiatrist. Diagnoses were confirmed with diagnostic interviews. More detailed inclusion and exclusion criteria are available in the On-line Appendix and On-line Research Protocol (see also Schrantee et al.).

**Intervention, Randomization, and Blinding**

Patients were randomly assigned to either methylphenidate or placebo treatment. The treating physicians prescribed the study medication under double-blind conditions on clinical guidance in accord with Dutch treatment guidelines. Medication dosages were titrated per individual, under supervision of a child and adolescent psychiatrist. Adult participants received coaching sessions, and parents of children received psychoeducation. Adherence to the study medication was monitored, and the compliance rate (0.0–1.0) was calculated (On-line Appendix).

**Outcomes**

**Primary Outcome Measure of ROI Cortical Thickness Change.** Imaging was performed using 1 of two 3T MR imaging systems (scanner 1, Intera; scanner 2, Achieva; Philips Healthcare). All participants were scanned on the same scanner pre- and postintervention. Anatomic 3D T1-weighted fast-field echo sequences were acquired and used for cortical surface analyses. Cortical thickness measures were derived using FreeSurfer (http://surfer.nmr.mgh.harvard.edu), Version 5.3. Cortical thickness was estimated vertex-wise across the brain surface using an automated approach. Scans were rated for the presence of motion and were excluded unless rated (1 = no sign of motion, or 2 = minor signs of motion, but no major distortion and acceptable reconstruction). An ROI approach was considered the viable option, given the necessarily restricted participant numbers of this RCT. Montreal Neurological Institute (MNI) coordinates of the vertices corresponding to peak group differences in the observational study were converted to Talairach coordinates, and labels were created at the corresponding vertices on the FreeSurfer average surface. The MNI coordinates for peak group differences forming the basis for the cortical ROIs were as follows: left frontal: −26, 61, 6; right medial frontal: −23, 40; right posterior: 30, −93, 12. The labels were dilated 15 times, resulting in hexagon-shaped ROIs covering approximately 550 mm² when transformed to the individual participant’s brain surface (Fig 1). Surface measures were extracted from the individual participants (Fig 1 and On-line Appendix).

**Statistical Analyses**

Analyses of ROI values were performed using SPSS, Version 24 (IBM), per reviewer request; analyses on handedness were later added, performed using SPSS, Version 26 (https://www.ibm.com/analytics/data-science/predictive-analytics/spss-statistical-software). One-way ANOVA was used to test differences in baseline brain characteristics, scan interval, scan movement ratings, compliance rates, and age for children and adults separately. Linear mixed models with age group (children, adults) × treatment group (placebo, methylphenidate) × time (pre-, posttreatment), as well as the main effects and interactions of each, were set up to test interaction effects on rates of change, using maximum likelihood estimation, with the covariance structure set to compound symmetry. For further analysis of medication effects per age group, we split the sample by age and set up similar linear mixed models with treatment group (placebo, methylphenidate) × time (pre-, posttreatment), as well as the main effects and interactions of each, to test interaction effects on rates of change within each age group for all ROIs separately. The significance level was set at $P < .05$.

While the ROI analysis was the primary outcome measure given the small number, an exploratory vertex-wise analysis on the whole cortical mantle was also performed (see On-line Appendix for details). In addition, cortical ROI change scores were correlated with the change in symptom severity pre- and posttreatment to explore possible relations.

**RESULTS**

**Randomization and Baseline Characteristics**

A total of 99 patients with ADHD were randomized to methylphenidate or placebo treatment. After randomization, 1 adult disclosed previous methylphenidate treatment and was therefore excluded from the analyses. On inspection of the scans, it was found in the child placebo group that one subject had a relatively large posterior cyst causing possible displacement of brain ROIs and another had an abnormality in the right frontal lobe that seemed to affect reconstructions of surfaces in the lateral cortex (by neuroradiologic follow-up thought to possibly be a benign tumor). Both of these individuals were therefore excluded from the current analyses. Forty-eight children and 47 adults were included in the primary analysis (CONSORT flow diagram, based on a template from http://www.consort-statement.org/consort-statement/flow-diagram, Fig 2). Of these, 1 child and 2 adults had invalid baseline data, and 1 child had invalid follow-up data due to scan motion or poor scan quality, leading to flawed image reconstruction. Note that 1 of the adults whose baseline data did not pass quality control was also among the ones who discontinued intervention. Treatment medication adherence data were missing for 1 child in the placebo group and for 3 children in the methylphenidate treatment group. No serious adverse events were noted in any of the individuals studied.

Sample characteristics are given in the Table (for in-depth clinical descriptives and behavioral measures used, see the On-line Appendix and Schrantee et al., Schmand et al., Rosler et al., Oosterlaan et al., and Wechsler). There were no significant differences in baseline age, ADHD symptom severity, reported handedness (proportion right-handed), estimated intelligence quotient, or ROI brain characteristics or scan interval or movement ratings between the placebo and medication groups. For the children, no tendency ($P < .10$) toward differences between treatment groups was observed. However, for the adults, there was a tendency toward a thicker cortex in the right posterior ROI ($P = .072$) at baseline in the placebo group and a greater compliance rate in the treatment group ($P = .096$). Because outcome ROI analyses were
for longitudinal cortical thickness change only, we chose not to correct for intracranial volume in the analyses.

**Main Outcome of the ROI Cortical Thickness Change**

Cortical thickness values for the groups pre- and posttreatment are shown in Fig 3, and individual data are shown in On-line Fig 1. Analyses using linear mixed models with scan time as a repeat measure × age × medication group showed, for the right medial cortex ROI a main effect of age ($F[1,95.606] = 199.713, P < .001$), likely reflecting expected age-related thinning, and an interaction of medication group × age group × time ($F[1,88.825] = 4.316, P = .041$). As shown in Fig 3, in the children, the methylphenidate group showed increasing thickness, while the placebo group showed cortical thinning. Although there were no significant baseline differences, slight differences were seen at baseline throughout the ROIs across age and medication groups, which also contributed to the identified interaction. The individual data plots did, however, not reveal striking outliers at any time point. Relatively lesser thinning in children in the methylphenidate treatment group was also seen in the right posterior cortex ROI, but not in the left frontal cortex ROI. However, no significant 3-way interaction effects of medication group × age group × time were found, except for the right medial cortex ROI. Other significant fixed effects in these analyses were limited to a main effect of age on the left frontal cortex ($F[1,96.308] = 46.788, P < .001$) and the right posterior cortex ($F[1,96.289] = 51.361, P < .001$), for which an interaction of age group × medication group was also found ($F[1,96.289] = 3.973, P = .049$). This effect appeared to be due to the child medication group having, on average, thicker cortices with the adult medication group having, on average, thinner cortices than the same-age placebo groups.

The analyses per age group confirmed an interaction of medication group × time ($F[1,44.893] = 4.793, P = .034$) in the young group only for the right medial cortex ROI. In the per-age group analyses, no other significant interaction effects were found, only a main effect of the medication group in the adults for the right posterior cortex ROI ($F[1,47.274] = 4.483, P = .040$), reflecting a thicker cortex in the adult placebo group at both time points. However, there were no trends toward the interaction of medication group × time for this ROI ($P = .986$). For comparison with previous results from the longer observational study from which the ROIs were derived, the present change rate for the right medial cortex ROI was converted to a rate of change per year, shown in Fig 4. The mean rates of change in cortical thickness for this region were quite similar across studies, but confidence intervals were larger in the present study.

The exploratory vertex-wise whole-brain analyses did not reveal any significant interaction between group and treatment on the thickness rate of change in any region when corrected for multiple comparisons (false discovery rate, $P < .05$). One-way ANOVA confirmed an effect of group on symptom change in the present sample, with post-, minus pre-treatment scores being lower for the methylphenidate than for the placebo groups ($F[1,42] = 5.551, P = .023$, and $F[1,41] = 5.857, P = .020$ for children and adults, respectively). Mean posttreatment ADHD symptom-severity scores in children were $30.2 \pm 10.1 (n = 22)$ versus $21.7 \pm 10.5 (n = 22)$ for the placebo and methylphenidate groups.
It may still be more likely that groups self-selecting baseline characteristics in terms of age, ADHD symptom severity, were randomized. While there were no significant differences in rates for those included in the analyses, divided by age and medication group.

**DISCUSSION**

The present results yield partial support for the hypothesis that methylphenidate treatment in development is associated with reduced cortical thinning, as identified here in the right medial cortex. The 3-way interaction of treatment group × age group × time for the right medial cortex resemble those found across longer time intervals in the observational study. Third, the MR imaging measures used in the previous observational study and the present one vary in terms of scanner platforms and field strengths (1.5T versus 3T), sequence parameters, and cortical reconstructions and segmentations. The observational study used CIVET (https://sourceforge.net/projects/civet/) software for cortical reconstruction and thickness estimation, whereas FreeSurfer was used in the present study. The measurement of cortical thickness requires that the analysis be performed at the nodes of a 3D polygonal mesh capturing the distance between the WM surface and the intersection between the outer cortical boundary and the CSF. This is in accordance with a geometric definition, which can vary across methods, so the way cortical thickness is quantified is of importance for the end results. Of note, the normal maturational cortical trajectories measured by CIVET and FreeSurfer can be discrepant. Moreover, the exact re-creation of the cortical effect sites identified in the observational study was not possible; therefore, smaller ROIs were drawn around the peak effects. Methodologic issues aside, the present results yield an indication toward thicker regional cortices with methylphenidate treatment in childhood specifically, as shown in the 3-way interaction of treatment group × age group × time for the right medial cortex, with absence of any treatment group × time interactions in respective. Mean posttreatment ADHD symptom-severity scores in adults were 26.5 ± 7.9 (n = 21) versus 20.1 ± 6.7 (n = 22) for the placebo and methylphenidate groups, respectively. However, there were no significant relationships between symptom-severity change and the estimated rate of ROI cortical change per year in either children or adults.
adults. This finding corresponds to evidence from animal models, indicating that psychotropic drugs affect the brain in a differential manner depending on the age of exposure.\textsuperscript{35,36} The results also align with findings of age-dependent effects of methylphenidate on cerebral perfusion and the blood flow response of the dopaminergic system in the same sample studied here.\textsuperscript{37} Hence, the present data on cortical thickness changes are in line with the neurochemical imprinting hypothesis,\textsuperscript{36} to some extent also being at play at a macrostructural level.

To the extent that brain differences among ADHD case-controls are predominantly found in children rather than adults, as was recently reported,\textsuperscript{7,8} it is also given that such differences can only be influenced when they are present in childhood. In large recent cross-sectional studies on observational cohorts, no effect of psychostimulant medication on subcortical volumes\textsuperscript{8} or the cortical thickness of regions showing patient-control group differences (fusiform gyrus, temporal pole)\textsuperscript{7} were found. There are, as mentioned above, multiple methodologic challenges to studying the effects of medication in a pure observational setting. The absence of such observational cross-sectional medication effects on cortical thickness could mean that such effects are not present in the studied regions or that any such effects are transient. It could, however, also be that medication may cause persistent changes, even if they can only be detected at the level of brain structure for a limited time.

FIG 3. Differences in rates of cortical thickness change of the ROIs in children and adults taking methylphenidate or placebo medication. For each ROI, at the y-axis, cortical thickness in millimeters (in left frontal, right posterior, and right medial cortices) is shown at baseline (x-axis time point 1) and at the end point (x-axis time point 2) of the trial. Note that as cortical thickness differs across regions and age groups, different origin values for cortical thickness are given across panels to enable visualization, but the scale is otherwise the same, and increments on the y-axes invariably represent 0.1 mm. Mean rates of change are shown with dotted lines for the placebo group and solid lines for the methylphenidate group. Error bars represent the 95% confidence intervals.

FIG 4. Mean rate of change in cortical thickness estimated per year for the present RCT and the previous observational study for the right medial cortex region. For children receiving methylphenidate (blue) and placebo (red), cortical change values for the right medial ROI across the duration of the present RCT were converted to an estimated mean rate of change in cortical thickness per year (left panel). This was done for comparison with previous results (right panel) from the observational study from which the ROI was derived (Shaw et al\textsuperscript{9}), reporting on the mean rate of cortical change per year across a period of about 4 years.
Limitations

There are several limitations to the present study. First, the sample was too small for detection of effects in a whole-brain analysis. Thus, although there were no significant baseline differences, slight differences were seen throughout the ROIs across age and medication groups, and baseline differences also contributed to the identified interaction. The vast normal variation in brain anatomy, including cortical measures, unfortunately makes it unlikely that relatively small groups will have completely identical values. Second, because we did not detect significant relationships between changes in cortical ROIs and changes in symptom severity, the functional significance remains uncertain. ADHD symptom severity was not related to subcortical differences in a recent cross-sectional mega-analysis either, so direct morphometric–functional relationships may not necessarily be expected. However, it would be of interest in future larger studies to investigate whether both baseline and longitudinal differences in cortical thickness are associated with a differential response to medication. Third, the follow-up interval was relatively short, and while changes in cortical thickness can be reliably detected across time at intervals even shorter than the present ones, medication effects may develop differently across a longer time. It is unknown why the right medial frontal cortical region, rather than the left frontal and left posterior cortical ROIs, showed the hypothesized effect in the present study. There was no reason to expect a differential strength of effects based on the previous observational study. We do not believe it is necessary to speculate about possible causes for this in this relatively small sample in which we also do not find correlations with symptom severity in any of the ROIs. Furthermore, it is beyond the scope of the present article to address functional imaging changes, but changes in cortical thickness with medication may likely also relate to and possibly interact with functional imaging changes, such as blood oxygen level–dependent functional signal variability.

Finally, macrostructural MR imaging measures cannot pinpoint the neural changes underlying the observed differences in thickness. Neurogenesis does not take place in the cortex in these age ranges, but a number of other processes may be at play, inclusive of differences in growth and regressive changes, astrocytes, dendritic branching, synaptic pruning, intracortical myelination, dendritic spine formation, and vascular changes. Preferably, animal models should combine high-resolution MR imaging and histologic methods to elucidate the neural fundament. The present results warrant replication in larger groups and possibly with longer follow-up, to determine whether the effects are permanent and whether they have long-term functional correlates.

CONCLUSIONS

The results of this RCT yield partial support for the hypothesis that methylphenidate treatment in development is associated with reduced cortical thinning in childhood, but not adulthood. This was identified here in the right medial cortex, but not in 2 other cortical ROIs previously indicated. Because significant relationships between changes in cortical ROIs and changes in symptom severity were not observed, the functional significance remains uncertain. Replication of the effects in prolonged studies in larger groups would be required to determine both possible temporal dynamics and functional significance of cortical effects.

Disclosures: Inge Amlien—RELATED: Grant: UNINETT Sigma2 NS90845, Comments: computer resources (storage/processing). Atle Bjørnerud—UNRELATED: Employment: Oslo University Hospital and University of Oslo.

REFERENCES


V.M. Zohrabian, L.H. Staib, M. Castillo, and L. Wang

ABSTRACT

BACKGROUND AND PURPOSE: Scientific collaboration is traditionally acknowledged through coauthorship. Studies on this topic are few in the neuroimaging literature. This study is a bibliometric analysis of the American Journal of Neuroradiology (AJNR) between 1980 and 2018, with the primary aim of evaluating changes in article collaboration.

MATERIALS AND METHODS: Full bibliographic records from 1980 to 2018 were retrieved. Yearly metrics calculated included the number of articles published, the average number of authors, and the average number of affiliations per article. The levels of evidence of 160 random articles were determined. Geographic characteristics of author affiliations were analyzed. Changes across time were evaluated using linear regression, while Spearman rank-order correlation was used to determine relationships between level of evidence and time, number of authors, and number of affiliations.

RESULTS: There was a steady linear growth in the number of articles ($R^2 = 0.70, P < 1\times10^{-10}$) from 1980 to 2018. There were clear linear increases in the average number of authors ($R^2 = 0.91, P < 1\times10^{-15}$) and affiliations ($R^2 = 0.90, P < 1\times10^{-15}$) per article. There was a significant correlation between level of evidence and time period (Spearman $\rho = -0.42, P < 1\times10^{-7}$), indicating that articles trended toward better methodologic quality or strength of results over time. A significant correlation existed between the level of evidence and the number of authors (Spearman $\rho = -0.39, P < 1\times10^{-6}$). There were linear increases in the average number of different geographic locales of authors per article by country/region ($R^2 = 0.80, P < 1\times10^{-13}$), state/province ($R^2 = 0.88, P < 1\times10^{-15}$), and locality/city/town ($R^2 = 0.86, P < 1\times10^{-15}$).

CONCLUSIONS: From 1980 to 2018, as the quantity of articles published in the AJNR increased, their level of evidence improved, while an increasing number of authors with different affiliations and from different geographic locales collaborated on these articles.

ABBREVIATION: WOS = Web of Science

Bibliometric analyses continue to gain traction in the scientific literature as they use quantitative approaches to evaluate and compare research across countries, institutions, investigators, and journals. In more than a decade, only a very small number of studies in the American Journal of Neuroradiology (AJNR), the official journal of the American Society of Neuroradiology, have centered on bibliometrics. Whereas most of these have highlighted citation numbers or Impact Factors, few have underscored authorships or affiliations. Furthermore, even though a number of these inquiries have demonstrated continually increasing international contributions to the AJNR, none have focused on collaborations among researchers from different geographic areas. One prior study found high interdisciplinary cooperation in neurointerventional research published in the AJNR between 2003 and 2012, yet it did not examine other types of collaboration. Scientific collaboration, so valuable to the dissemination of knowledge and production of high-quality research, is traditionally acknowledged through coauthorship and, therefore, can be measured by the number of authors and affiliations. In this study, we hypothesized that collaborative efforts among researchers with different affiliations and from different geographic locales, both domestic and international, have grown with time. We conducted a systematic, in-depth analysis of the characteristics and trends of articles published in the AJNR between 1980 and 2018, with the primary aim of evaluating changes in collaboration during 4 decades.

MATERIALS AND METHODS

Institutional review board approval was not required because human subjects were not involved. The Web of Science (WOS) (Clarivate Analytics) was determined to be the optimal...
browsable bibliographic data source for this study. Full bibliographic records of all items (n = 14,758) published in the AJNR from January 1, 1980, to December 31, 2018, were retrieved from the WOS and saved as tab delimited data files. A custom script was written in Ruby, an open-source programming language, to extract a subset of the bibliographic records to include only those with a document type of “Article” (n = 10,789), defined in WOS as “reports of research on original works,” which include “research papers, features, brief communications, case reports, technical notes, chronology, and full papers that were published in a journal and/or presented at a symposium or conference.”

This process excluded any comments, editorials, replies, letters, and reviews from our dataset. This subset of 10,789 bibliographic records served as the basis of all statistical analyses in this study.

The script also extracted all unique author affiliation addresses (n = 29,577) from the 10,789 bibliographic records and sent them individually to Google’s Geocoding API, which returned standardized geographic locations of those addresses at the locality/city/town, state/province, and country/region levels. The geocoding results were checked for accuracy at the country/region level, and inaccuracies (n = 1293) were manually corrected.

The script grouped records by publication year and calculated the following metrics for each year: number of articles published, average number of authors per article, average number of affiliations per article, average number of affiliation countries/regions per article, average number of affiliation states/provinces per article, and average number of affiliation localities/cities/towns per article.

The script also generated a random sample of 20 records for each 5-year period between 1980 and 2014 and an additional 20 random records for the 4-year period between 2015 and 2018, totaling 160 records. Two trained individuals retrieved the full text of each of the articles and independently assigned a level of evidence (levels 1–5) to each according to the evidence-based medicine ratings guidelines published on the AJNR Web site, as in Table 1. After independent review, any disagreements were resolved by consensus.

The articles were also stratified by country/region of origin or the geographic affiliation of the primary author for the study. The countries/regions of origin were determined for most of the articles (n = 10,230, 95%). The affiliation country/region of the last/senior author, when available, was established as the article origin. For the purposes of this study, we presumed that the last/senior author’s institution was the focal point of the work and where the work originated. For articles without such information, an origin was recognized if only a single affiliation country/region was listed for the article.

Visual representations of international collaboration across time were achieved by plotting network diagrams using Cytoscape (Cytoscape Consortium, Version 3.7.2, 2019; https://cytoscape.org/download.html). The “nodes” in the diagrams represented countries/regions with articles that had at least 1 coauthor from a different country/region, and the size of the nodes represented the number of such articles. The nodes were linked by “edges,” the thickness of which represented the number of articles with coauthors from the pair of connected countries/regions. The network diagrams were scaled using Cytoscape for better readability.

All data were tabulated and statistical analyses performed using R statistical and computing software (Version 3.3.2, 2016; http://www.r-project.org/). Changes with time were evaluated using linear regression. The Spearman rank-order correlation was used to assess relationships between level of evidence and time, number of authors, and number of affiliations.

RESULTS

Number of Articles
A total of 10,789 articles published in AJNR between 1980 and 2018 met the inclusion criteria for this bibliometric study. As shown in On-line Fig 1, there was a steady linear growth in the number of articles published yearly in the AJNR since 1980 (R² = 0.70, P < 1e–10), from 85 articles in 1980 to 321 articles in 2018. Article counts increased, on average, by 6.4 each year.

Number of Authors and Affiliations
On-line Fig 2 shows linear increases in the average number of authors (R² = 0.91, P < 1e–15) and affiliations (R² = 0.90, P < 1e–15) per article. The average number of authors per article increased by 0.10 per year. The average number of affiliations per article increased by 0.08 per year. The average number of authors per article was 3.4 in 1980, 5.3 in 2000, and 8.0 in 2016. The average number of affiliations represented in each article was 1.6 in 1980, 2.5 in 2000, and 4.4 in 2016.

Level of Evidence
There was a significant correlation between the level of evidence and time period (Spearman ρ = −0.42, P < 1e–7), indicating that articles trended toward better methodologic quality or strength of results across time, as shown in On-line Fig 3. For example, the proportion of level 5 articles decreased, while level 3 and 4 articles increased during the 4 decades.

Level of Evidence and Number of Authors and Affiliations
A significant correlation was found between the level of evidence and the number of authors (Spearman ρ = −0.39, P < 1e–6), as shown in On-line Fig 4. The proportion of level 5 articles decreased as the number of authors increased. However, no significant relationship was found between the level of evidence and the number of affiliations (Spearman ρ = −0.15, P = .054, not significant).

Geographic Locales of Author Affiliations
Authors from the United States were, by far, the largest contributors to AJNR, appearing in 6270 (58%) articles. The United States

Table 1: Evidence-based medicine ratings, adapted from the AJNR Web site

<table>
<thead>
<tr>
<th>Level</th>
<th>Study Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Randomized controlled trial</td>
</tr>
<tr>
<td>2</td>
<td>Prospective cohort study</td>
</tr>
<tr>
<td>3</td>
<td>Retrospective cohort study, case-control study</td>
</tr>
<tr>
<td>4</td>
<td>Case series</td>
</tr>
<tr>
<td>5</td>
<td>Mechanism-based reasoning</td>
</tr>
</tbody>
</table>
was followed by Japan (9%), Germany (7%), and Canada (5%) as the next largest contributing countries.

As in On-line Fig 5, more than half (n = 5658, 55%) of AJNR articles published between 1980 and 2018 (whose origins could be determined) originated from the United States. The United States was followed by Japan (8%), Germany (5%), and Canada (4%) as the next largest origin countries.

A detailed assessment of geographic locales of author affiliations in each article was performed. On-line Fig 6 demonstrates linear increases in the average number of different geographic locales of author affiliations per article from 1980 to 2018 by country/region ($R^2 = 0.80$, $P < 1e-13$), state/province ($R^2 = 0.88$, $P < 1e-15$), and locality/city/town ($R^2 = 0.86$, $P < 1e-15$). The average number of country/regions represented in each article increased by 0.01 per year. The average number of states/provinces per article increased by 0.03 per year. The average number of localities/cities/towns per article increased by 0.04 per year. Of the 10,789 articles analyzed, 1456 (14%) were the result of collaborations between at least 2 countries/regions, and 250 (2%), among at least 3. These numbers trended upward across the decades, as shown in Table 2. Furthermore, as shown in Table 3, the percentages of AJNR articles by authors from the same country/region, state/province, and locality/city/town decreased during the 4 time periods, with the largest decrease occurring at the locality/city/town level. For example, 74% of AJNR articles were by authors from the same locality/city/town during 1980–1989. That percentage dropped to 37% during 2010–2018. At the state/province level, the percentage dropped from 82% to 43%, while at the country/region level, it dropped from 96% to 75%.

Nearly all countries/regions were increasingly involved in international collaborations across time. On-line Fig 7 reveals that on the basis of the change between the last 2 decades, China had the largest gain in international collaborations (606%), followed by Norway (340%), Spain (338%), and France (319%). The United States (147%), Germany (145%), and Japan (108%), on the other hand, demonstrated more modest gains during the past 2 decades. Table 4 demonstrates that the United States held a decreasing share of the “international collaboration pie” across time because the percentage of such articles involving the United States decreased from 1980–1989 (74%) to 2010–2018 (62%).

The complexity of the network diagrams increased during the 4 time periods, 1980–1989 (On-line Fig 8A), 1990–1999 (On-line Fig 8B), 2000–2009 (On-line Fig 8C), and 2010–2018 (On-line Fig 8D). Between 1980 and 1989, only 24 countries/regions (in 39 pairs) published articles in the AJNR with international authors. These numbers increased to 57 and 368, respectively, between 2010 and 2018. The top 5 countries/regions that collaborated internationally in the AJNR, and the top 3 collaboration pairs overall (1980–2018) are presented in Tables 5 and 6.

**DISCUSSION**

Collaboration enhances the quality of scientific research. Collaboration allows sharing of knowledge, facilitates acquisition of specialized and new expertise, fosters creativity, and promotes innovation. While macro-level collaborations are typically large initiatives by collaborating nations or international organizations, meso-level collaborations are usually among academic institutions, sometimes across national boundaries. Micro-level collaborations, on the other hand, involve individual researchers or small groups of individual researchers working together. The formal acknowledgment of scientific collaboration is traditionally through coauthorship and can be measured by the number of authors and affiliations. In this study, we performed in-depth bibliometric analyses of the characteristics and trends of AJNR articles published during the past 4 decades, from 1980 to 2018. We
examined article counts, level of evidence, number of authors, and the number and geography of author affiliations and, in doing so, proved that collaboration has increased across the years.

In examining collaborations, we established that the average number of authors per article increased by 0.10 per year, from 3.4 per article in 1980 to 8.0 per article in 2018. Multiple prior studies have likewise demonstrated a trend toward a greater number of authors per article, which may, in part, be due to increasing scientific complexity and specialization. Our manual analysis of a small random sample of 160 articles revealed that the level of evidence has improved since 1980, trending toward better methodologic quality or strength of results. Whereas studies published in 1980 were all level 4 or 5, studies in 2015 were mostly level 2 or 3. Furthermore, we demonstrated a significant association between the level of evidence and the number of authors so that more methodologically robust studies tended to have a larger number of authors and vice-versa. However, this finding is a correlation, and we have not proved that the number of authors and the level of evidence are causally related. A prior study showed that the number of authors per article was larger for original research than for case reports and review articles. On the other hand, increasing authorship may be merely because publications are an increasing central part of recruitment, promotion, and retention of faculty in academic medicine. In fact, increasing authorship in the biomedical literature has been subject to debate, especially in the radiology literature, because several authors have noted increasing trends of inappropriate authorship during several decades.

A recently published analysis of the impact of the recommendations of the International Committee of Medical Journal Editors on 49 clinical radiology journals between 1946 and 2013 found that overall authorship increased with time, independent of the journal of publication, country of publication, publication type, and language of publication. Moreover, of the top 10 radiology journals by volume, AJNR ranked above the average for change in authorship with time. Therefore, authors are reminded to strictly adhere to the guidelines of the International Committee of Medical Journal Editors, which are specified on the AJNR Web site, to ensure proper attribution of authorship.

In our examination, we also established that the average number of affiliations per article increased by 0.08 per year. However, no significant association was found between the level of evidence and the number of affiliations. We speculate that access to intra-institutional diversity of expertise has improved with time and has helped in conducting studies that required more methodologic rigor.

Authors from the United States were, by far, the largest contributors to the AJNR, appearing in 6270 (58%) articles. The United States was followed by Japan (9%), Germany (7%), and Canada (5%). During the 4 decades, however, the percentage of articles with authors from the United States declined (-30%), from 80% in 1980–1989 to 50% in 2010–2018, while the percentages of articles with authors from other countries/regions increased, with China and Germany demonstrating the largest increases (+7%). The United States was the origin country for more than half (n = 5658, 55%) of all AJNR articles published between 1980 and 2018 whose origins could be determined, which is not surprising because AJNR is the official journal of the American Society of Neuroradiology. The United States was followed by Japan (8%), Germany (5%), and Canada (4%) as the next largest origin countries. However, the percentage of articles originating from the United States declined (-36%) with time, from 80% in 1980–1989 to 44% in 2010–2018, while South Korea (+6%) and China (+5%) demonstrated the largest gains.

Our data are consistent with a previously published work that has shown that although the United States is the largest contributor of articles to selected journals in clinical radiology between 1991 and 2000 (43.2%), the relative shares of the United States, United Kingdom, and Canada have decreased, whereas those of Japan, France, Germany, Italy, South Korea, Spain, Switzerland, Turkey, Austria, and China have increased. Additionally, a study in the American Journal of Roentgenology found that from 1980 to 1982, ten percent of the articles originated from institutions outside the United States, but between 1990 and 1992, international articles accounted for 25% of published articles. A follow-up of the study found that while the total number of published articles from most nations increased from 1980 to 2002, Korea, Japan, and Germany made the largest advances. Moreover, our study confirmed a prior work in the AJNR, which noted a relative decline in contributions from the United States since 1992 and proposed that articles from China, India, and Eastern Europe would continue to increase as the economies of these countries expanded and became more integrated with the rest of the world through globalization.

Most interesting, beyond merely observing increasing international contributions in the AJNR, when the geography of authors’ affiliations was examined, we confirmed a trend toward increasing national and international collaborations on articles during 4 decades. This is demonstrated by an increase in the average number of geographic locales per article by year in On-line Fig 6 and is visually portrayed by the increasing complexity and interconnectivity of the network diagrams in On-line Fig 8A–D. Internationally, the United States and Germany (n = 149), the United States and Canada (n = 147), and the United States and Japan (n = 109) collaborated on the greatest number of articles between 1980 and 2018, though nearly all countries/regions demonstrated increasing international collaboration. China demonstrated the largest gain, with a 606% increase between 2000 and 2009 and 2010 and 2018, followed by Norway (340%), Spain (338%), and France (319%). On the other hand, the United States demonstrated a more modest gain of 147% during these last 2 decades, followed by Germany (145%) and Japan (108%).

As the volume of international collaboration increased across time, the relative percentage of each country/region declined. For example, although the United States, both as an origin and collaborating country, contributed the largest number of AJNR articles during the 4 decades, its relative share in international collaboration

---

Table 6: Top 3 pairs overall collaborating internationally in AJNR, 1980–2018

<table>
<thead>
<tr>
<th>Country 1</th>
<th>Country 2</th>
<th>Pair Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>United States</td>
<td>149</td>
</tr>
<tr>
<td>Canada</td>
<td>United States</td>
<td>147</td>
</tr>
<tr>
<td>Japan</td>
<td>United States</td>
<td>109</td>
</tr>
</tbody>
</table>

---

decreased with time, as shown by the declining percentages of international collaboration articles involving the United States in Table 4. A trend of increasing collaboration across academic borders in science and engineering and the social sciences has been noted on a much larger scale by Jones et al.29 In examining 4.2 million research articles published between 1975 and 2005, the researchers demonstrated that multi-university collaborations grew from 10% of articles in 1975 to 30%–35% of articles in 2005, while single-author articles became increasingly rare during the same time period.29 However, the researchers also found that the average distance between collaborators grew only slightly, from 725–750 miles in 1975 to 800 miles in 2005.29 This latter finding is also substantiated in our study because we found that the percentage of articles by authors from the same country/region decreased to a much lesser extent during the 4 decades than the percentages of those by authors from the same state/province, and especially the same locality/city/town, as shown in Table 3.

The digital age has lowered the transaction costs for international collaboration by breaking down barriers related to physical distance and variations in technology.30 Technologic advances have made it much easier to communicate, and a multitude of Web sites and apps have enabled users to create and share content in the virtual space. Social media and social collaborative networks, which saw a revolution in the 21st century, have permeated nearly every aspect of personal and professional lives. The number of social media users worldwide has more than doubled in less than a decade, from 0.97 billion in 2010 to 2.62 billion in 2018.31 As artificial intelligence and machine learning continue to transform social media platforms, they will likely become even more pervasive. While earlier studies have shown a low frequency of social media use among scientists,32-34 the landscape is quickly changing, with a recent survey in Nature revealing that more than 95% of respondents have used some form of social media or social collaborative network for professional purposes, that 50% of professional users access Facebook on a daily basis, and that most respondents use social media for discovering and/or reading scientific content.35 Most interesting, in our study, we noted that the most rapid change in the average number of geographic locales per article took place since 2006, in the midst of the social media revolution, as shown in On-line Fig 6. However, this trend, albeit more modest, is also observed in the 1980s to 1990s, long before the Internet and social media boom, implying that additional factors are responsible for growth in collaboration.

There are several limitations to our study, most of which pertain to data extraction. At the time of this study, we were not able to dissect author affiliations down to institutional or departmental levels because such information in a structured, consistent format was not immediately available for retrieval. Instead, we had to rely on geographic location data parsed from author affiliation addresses by Google’s Geocoding API. Although fairly accurate, we found that results returned from the API had some inaccuracies at the country/region level at a rate of 4.37% (1293 of 29,577), affecting 9.39% (1013 of 10,789) of all records. These inaccuracies were manually corrected, but only at the country/region level. Due to resource constraints, it was not practical to check and correct for inaccuracies at the state/province and locality/city/town levels; therefore, geocoding inaccuracy rates at these levels are not known. However, on the basis of our cursory review of records and the low inaccuracy rate for country/region, these inaccuracies are likely minor.

Additionally, we realize that in some instances, authors may have recorded multiple affiliations. Unfortunately, data from WOS do not specify primary or secondary affiliations. In our study, authors and affiliations were analyzed separately most of the time. In those cases, all authors and affiliations were considered. The linkage between authors and affiliations was only relevant when determining country/region of origin for each article. In that case, we had to choose 1 (and only 1) affiliation for the last author and use the country/region for that affiliation as the country/region of origin for the article. For this specific analysis, affiliations were considered only at the country/region level, and very few articles (n = 26) had last/senior author affiliations spanning multiple countries/regions, representing only 0.25% of articles whose countries/regions of origin were determined (n = 10,230) and 0.24% of all articles in our dataset (n = 10,789). Given the very small number of such articles, any potential inaccuracies caused by multiple affiliations were likely minimal.

Furthermore, in assigning country of origin, we presumed that the last author’s institution was the focal point of the work that was done and geographically where the work originated. In the radiology literature, the convention is that the first author is the greatest contributor to the work and is its principal writer, whereas the last author is the more senior and experienced one responsible for its content. Because scholarly output is an increasingly important part of promotion, younger researchers need to publish as first authors early in their academic careers when they may be still be in training, change specialties, or switch institutions or even countries altogether. Because senior authors tend to be higher on the academic ladder, their affiliations are less likely to change. However, this latter point is an inference based on anecdotal evidence and may not be entirely consistent. Moreover, the records we analyzed were of those published between 1980 and 2018. When we compared metrics by decades, the last decade (2010–2018) was 1 year shorter than the other decades. Consequently, when we generated a smaller sample for level-of-evidence grading by 5-year periods, the last period (2015–2018) was shorter by 1 year. However, because we were interested in examining larger trends across decades, inclusion of data from 1 additional year was not likely to considerably alter our findings.

Finally, our study uses the number of authors and the geography of affiliations as proxies for collaboration. As discussed earlier, inappropriate authorship is a problem facing not only radiology, but all of medicine. To what extent inappropriate authorship was represented in our bibliographic sample and how this may have influenced our findings are not known. Future studies might use sophisticated metrics for collaboration, such as a collaborative index, degree of collaboration, collaborative coefficient, modified collaborative coefficient, and degree of centrality, as well as examine relationships of these metrics with research productivity, journal Impact Factor, and number of citations. Larger analyses might also be stratified by article topic or category, for example, brain, spine, head and
neck, and interventional procedures, to ascertain differences in levels of evidence as well as to quantify levels of evidence for countries of origin and correlate these findings with changes in the number of articles from these countries across time.

CONCLUSIONS

Scientific collaboration is essential for the dissemination of knowledge and production of high-quality research. From 1980 to 2018, as the quantity of articles published in the AJNR increased, their level of evidence improved, while an increasing number of authors with different affiliations and from different geographic locales collaborated on these articles.

REFERENCES

15. Chew FS. Coauthorship in radiology journals. AJR Am J Roentgenol 1988;150:23–26 CrossRef Medline
Medicare for All: Considerations for Neuroradiologists


ABSTRACT
SUMMARY: The year 2019 featured extensive debates on transforming the United States multipayer health care system into a single-payer system. At a time when reimbursement structures are in flux and potential changes in government may affect health care, it is important for neuroradiologists to remain informed on how emerging policies may impact their practices. The purpose of this article is to examine potential ramifications for neuroradiologist reimbursement with the Medicare for All legislative proposals. An institution-specific analysis is presented to illustrate general Medicare for All principles in discussing issues applicable to practices nationwide.

ABBREVIATIONS: MFA = Medicare for All; PC = professional component

Senator Bernie Sanders originally introduced the Medicare for All (MFA) Act in 2017, which aimed to establish a single national health insurance program. At that time, the bill was supported by only 16 other senators and eventually died in the 115th Congress. Fast forward 2 years: Several 2020 presidential candidates now emphatically endorse MFA, aligned with the fact that during the past decade, the support by the public for a national health plan has grown.1 In April 2019, Senator Sanders reintroduced his MFA bill with expanded covered services.2 This came on the heels of an additional MFA bill proposed in the House of Representatives in February 2019 by Representative Pramila Jayapal.3 Other politicians have proposed offering public health insurance while preserving private options.4 With health care reform at the center of national debate during this election year, it is important for physicians to be abreast of potential landmark health care legislation. While varying plans have been proposed on how to reform US health care, only the 2 aforementioned MFA Acts have been formally introduced in Congress at the time of this writing. The fundamental objectives common to both bills are summarized in Table 1. In this article, we use the term “MFA” specifically in reference to both Medicare for All Acts of 2019.

The financial impact of MFA implementation on individual radiology practices would depend on the case mix of services of each practice, the payer mix for each patient service, and reimbursement rates currently negotiated with private payers. The current payment models of the practice (eg, fee-for-service, bundled payments, or global budget) are also determinants of the potential impact of MFA on that particular practice. The MFA bills do not specifically address payment models; therefore, it is unclear how these would be affected. Thus, it is uncertain what would happen with reimbursement penalties and bonuses under the Merit-Based Incentive Payment System and Medicare Access and CHIP Reauthorization Act of 2015.

The ambiguity in MFA makes it difficult to predict the effect of MFA on a practice. However, the salient feature in the MFA legislative proposals is essentially transitioning the US multipayer health care system to a single-payer system. The goal of this Practice Perspectives is to discuss issues neuroradiologists may encounter in such a transition. We also provide an institution-specific analysis to illustrate the degree of impact this shift can have on a practice. The analysis is based on the case mix of the primary author’s institution and reimbursement figures for its locale. Medicare global, professional component (PC), and technical component fees were determined using the Medicare Physician Fee Schedule for 2019.5 Private reimbursement rates for noncapitated services provided by in-network providers to patients with primary residence in Louisiana or Mississippi (the Data Use Agreement prohibited us from reporting state-level metrics for Louisiana only) were extracted from the IBM MarketScan Commercial Claims and Encounters Database for the year 2017. All payments were adjusted for inflation using the seasonally-adjusted Consumer Price Index for Medical Care obtained from the Federal Reserve...
Bank of St. Louis and expressed in 2019 US dollars. Medicaid fees in this study represent hospital-based Medicaid PC fee averages in Louisiana for the respective examinations and were provided by the radiology finance department of our lead study site. Although the analysis is institution-specific, it serves as a real-world example to illustrate general issues that apply to practices nationwide.

**Financial Implications for Providers**

In the United States, approximately 49% of the general population receives health insurance through their employers; 20%, through Medicaid; and 14%, through Medicare. About 9% of the population is uninsured. Other coverage comprises individual/nongroup insurance (6%) and military (1%). Under MFA, private health insurers would be prohibited from offering benefits included in the public program. The transition from a multipayer to single-payer system could have negative financial consequences for many radiologists because private insurer fees typically exceed those of Medicare. With regard to neuroradiology, brain MR imaging payments from private insurers were shown to be more than twice as high as Medicare fee-for-service payments in a 2018 study published by the Congressional Budget Office.

The difference between Medicare and private insurance fees is affected by factors including the negotiating power of a practice with private insurers, case mix, and the Geographic Practice Cost Index. As an example of variation between Medicare and private insurance reimbursement for a neuroradiology practice, Table 2 shows a brain examination case mix from the primary author’s radiology department with fees from Medicare and estimated average reimbursements from private insurers.

The shift to MFA could be especially disruptive for practices that currently treat a large portion of privately insured patients because private insurer fees are typically higher than those of Medicare throughout the country. Conversely, MFA could increase revenue for some practices by reducing their uncompensated care for uninsured patients as well as by increasing their reimbursement for care of patients currently covered by lower-paying Medicaid. Previous research found the average national Medicaid-to-Medicare fee index is 0.72, though inpatient Medicaid fees vary on the basis of specific hospital cost-to-charge ratios. Some also argue that MFA would reduce costs for providers by decreasing their administrative burdens.

Furthermore, practices might benefit from reduced time between when services are rendered and reimbursed. A current advantage of Medicare over private insurers is that Medicare often reimburses sooner.

### Table 1: Fundamental objectives of the Medicare for All Acts of 2019

<table>
<thead>
<tr>
<th>Objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Establishment of a national health insurance program administered by the government that would cover all US residents</td>
</tr>
<tr>
<td>Creation of a nation health budget</td>
</tr>
<tr>
<td>Elimination of cost-sharing (deductibles, coinsurance, copayments) and balance billing for covered services</td>
</tr>
<tr>
<td>Interdiction against duplicate benefits from private health insurers or employers</td>
</tr>
</tbody>
</table>

### Table 2: Example of neuroradiology case mix and 2019 Medicare fees compared with private insurance

<table>
<thead>
<tr>
<th>Examination (CPT)</th>
<th>No. of Examinations (2019)</th>
<th>Case Mix (%)</th>
<th>Medicare Global</th>
<th>PI Global</th>
<th>Medicare PC</th>
<th>PI PC</th>
<th>Medicare TC</th>
<th>PI TC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT head (70450)</td>
<td>44,300</td>
<td>69.5</td>
<td>$115.20</td>
<td>$369.82</td>
<td>$43.72</td>
<td>$74.54</td>
<td>$71.48</td>
<td>$295.28</td>
</tr>
<tr>
<td>CT head with contrast (70460)</td>
<td>50</td>
<td>0.1</td>
<td>$162.76</td>
<td>$420.63</td>
<td>$58.56</td>
<td>$95.97</td>
<td>$104.20</td>
<td>$324.66</td>
</tr>
<tr>
<td>CT head without and with contrast (70470)</td>
<td>564</td>
<td>0.9</td>
<td>$190.09</td>
<td>$437.57</td>
<td>$65.34</td>
<td>$106.08</td>
<td>$124.74</td>
<td>$331.49</td>
</tr>
<tr>
<td>MR imaging brain (70551)</td>
<td>9400</td>
<td>14.7</td>
<td>$224.92</td>
<td>$550.98</td>
<td>$76.16</td>
<td>$131.26</td>
<td>$148.77</td>
<td>$419.72</td>
</tr>
<tr>
<td>MR imaging brain with contrast (70552)</td>
<td>244</td>
<td>0.4</td>
<td>$311.96</td>
<td>$486.19</td>
<td>$91.71</td>
<td>$158.49</td>
<td>$220.24</td>
<td>$327.70</td>
</tr>
<tr>
<td>MR imaging brain without and with contrast (70553)</td>
<td>9225</td>
<td>14.5</td>
<td>$368.27</td>
<td>$916.65</td>
<td>$117.39</td>
<td>$217.90</td>
<td>$250.88</td>
<td>$693.75</td>
</tr>
<tr>
<td>Total</td>
<td>63,783</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:**—NR indicates nonreimbursed; PI, private insurance; TC, technical component; CPT, Current Procedural Terminology.

<table>
<thead>
<tr>
<th>Examination (CPT)</th>
<th>No. of Examinations (2019)</th>
<th>Case Mix (%)</th>
<th>Medicare Global</th>
<th>PI Global</th>
<th>Medicare PC</th>
<th>PI PC</th>
<th>Medicare TC</th>
<th>PI TC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT head (70450)</td>
<td>44,300</td>
<td>69.5</td>
<td>$115.20</td>
<td>$369.82</td>
<td>$43.72</td>
<td>$74.54</td>
<td>$71.48</td>
<td>$295.28</td>
</tr>
<tr>
<td>CT head with contrast (70460)</td>
<td>50</td>
<td>0.1</td>
<td>$162.76</td>
<td>$420.63</td>
<td>$58.56</td>
<td>$95.97</td>
<td>$104.20</td>
<td>$324.66</td>
</tr>
<tr>
<td>CT head without and with contrast (70470)</td>
<td>564</td>
<td>0.9</td>
<td>$190.09</td>
<td>$437.57</td>
<td>$65.34</td>
<td>$106.08</td>
<td>$124.74</td>
<td>$331.49</td>
</tr>
<tr>
<td>MR imaging brain (70551)</td>
<td>9400</td>
<td>14.7</td>
<td>$224.92</td>
<td>$550.98</td>
<td>$76.16</td>
<td>$131.26</td>
<td>$148.77</td>
<td>$419.72</td>
</tr>
<tr>
<td>MR imaging brain with contrast (70552)</td>
<td>244</td>
<td>0.4</td>
<td>$311.96</td>
<td>$486.19</td>
<td>$91.71</td>
<td>$158.49</td>
<td>$220.24</td>
<td>$327.70</td>
</tr>
<tr>
<td>MR imaging brain without and with contrast (70553)</td>
<td>9225</td>
<td>14.5</td>
<td>$368.27</td>
<td>$916.65</td>
<td>$117.39</td>
<td>$217.90</td>
<td>$250.88</td>
<td>$693.75</td>
</tr>
<tr>
<td>Total</td>
<td>63,783</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:**—NR indicates nonreimbursed; PI, private insurance; TC, technical component; CPT, Current Procedural Terminology.

<table>
<thead>
<tr>
<th>Examination (CPT)</th>
<th>No. of Examinations (2019)</th>
<th>Case Mix (%)</th>
<th>Medicare Global</th>
<th>PI Global</th>
<th>Medicare PC</th>
<th>PI PC</th>
<th>Medicare TC</th>
<th>PI TC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR imaging brain WO contrast</td>
<td>70551</td>
<td>15%</td>
<td>$0</td>
<td>$74.64</td>
<td>$72.08</td>
<td>$131.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MR imaging brain WO/W contrast</td>
<td>74178</td>
<td>15%</td>
<td>$0</td>
<td>$102.53</td>
<td>$88.15</td>
<td>$173.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US Abd complete</td>
<td>76770</td>
<td>10%</td>
<td>$0</td>
<td>$37.14</td>
<td>$37.86</td>
<td>$63.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screening mammogram</td>
<td>77067</td>
<td>20%</td>
<td>$0</td>
<td>$10.25</td>
<td>$8.52</td>
<td>$18.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XR chest single</td>
<td>71046</td>
<td>40%</td>
<td>$0</td>
<td>$38.79</td>
<td>$38.46</td>
<td>$77.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>100%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:**—WO indicates without; WO, without; Abd, abdomen; US, ultrasound; XR, x-ray.

**a**Medicaid fees represent hospital-based Medicaid PC fee averages in Louisiana for the respective examinations and were provided by the Radiology Finance Department of our lead study site.

To illustrate the impact on reimbursement in transitioning to a single-payer system, we present in Table 3, a simplified hypothetic general radiology case mix of 5 examinations, including MR imaging of the brain. These data are based on the case mix from the primary author’s radiology department and reimbursement figures for this locale to provide as realistic an example as possible of radiology reimbursement in the US multipayer system. Current Procedural Terminology codes were used to obtain volumes of the corresponding examinations performed in the first half of 2019. The finance department of the institution providing these data advised against publishing department-specific reimbursement data. With this condition, the analysis includes commonly performed examinations outside neuroradiology because high-volume studies typically make up a large proportion of overall reimbursements. Furthermore, a variety of imaging modalities more accurately reflects the impact on neuroradiologists who tend to practice in broader radiology environments.

In the general Louisiana population, approximately 13.1% of individuals are enrolled in Medicare, while about half have private health insurance. Typically, Medicare enrollees use more health services than privately insured individuals, which is reflected in the inpatient population payer mix (Medicare accounts for 25.8%) derived from national data. Table 4 presents these examples of payer mixes. The payer mix of a typical radiology practice in Louisiana likely lies between these 2 examples if the practice performs both inpatient and outpatient imaging.

If we assume no change in the use of health services, if all patients in Louisiana were transitioned to Medicare at current Medicare PC fees, overall PC reimbursements for all 5 examinations in this analysis would decrease—with MR imaging of the brain leading that decline. The estimated revenue reduction for our simplified radiology case mix is 14%–20% (Table 5). Given the larger decrease for MR imaging payments, practices with higher percentages of MR imaging brain examinations could be more affected. The change in reimbursement is smaller in the inpatient scenario because Medicare and Medicaid currently make up larger proportions of payers in this setting.

### Table 4: Examples of payer mix

<table>
<thead>
<tr>
<th>Payer Type</th>
<th>Louisiana General Population Payer Mix</th>
<th>Inpatient Population Payer Mix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private insurance</td>
<td>50.5%</td>
<td>35.6%</td>
</tr>
<tr>
<td>Medicare</td>
<td>13.1%</td>
<td>25.8%</td>
</tr>
<tr>
<td>Medicaid</td>
<td>28.3%</td>
<td>31.9%</td>
</tr>
<tr>
<td>Uninsured</td>
<td>8.1%</td>
<td>6.7%</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

*Based on data from the Kaiser Family Foundation 2017. The source data included an additional payer type “Other Public,” which comprised military and the Veterans Health Administration. Our payer mix percentages were weighted to exclude this additional group, which made up 1% of the original mix.

*Derived from data from the Health Care Cost and Utilization Project and U.S. Census Bureau 2017.

### Table 5: Change in professional component reimbursement under Medicare for All at a radiology practice in Louisiana

<table>
<thead>
<tr>
<th>Examination</th>
<th>Payer Type Scenario</th>
<th>100% Medicare</th>
<th>Current General Population Payer Mix</th>
<th>% Difference</th>
<th>Current Inpatient Payer Mix</th>
<th>% Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR imaging brain WO contrast</td>
<td></td>
<td>$74.64</td>
<td>$96.46</td>
<td>–23%</td>
<td>$89.14</td>
<td>–16%</td>
</tr>
<tr>
<td>CT Abd/pelvis WO/W contrast</td>
<td></td>
<td>$102.53</td>
<td>$125.85</td>
<td>–19%</td>
<td>$117.11</td>
<td>–12%</td>
</tr>
<tr>
<td>US Abd complete</td>
<td></td>
<td>$37.14</td>
<td>$46.07</td>
<td>–19%</td>
<td>$42.80</td>
<td>–13%</td>
</tr>
<tr>
<td>Screening mammogram</td>
<td></td>
<td>$38.79</td>
<td>$47.95</td>
<td>–19%</td>
<td>$44.81</td>
<td>–13%</td>
</tr>
<tr>
<td>XR chest single</td>
<td></td>
<td>$10.25</td>
<td>$13.20</td>
<td>–21%</td>
<td>$12.13</td>
<td>–14%</td>
</tr>
<tr>
<td>% Change if all uninsured, Medicaid, and privately insured patients are transitioned to Medicare</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*W indicates with; WO, without; Abd, abdomen; US, ultrasound; XR, x-ray.

### Considerations for Neuroradiology Practices

The 14%–20% figure in Table 5 pertains to the primary author’s institution and is meant to serve as an example in discussing generalizable topics under MFA. A separate study by Galvani et al. projected a 7.4% reduction if all physician and clinical services were reimbursed at Medicare rates. For practices that own equipment, technical fees contribute largely to the total cost of an imaging study. Louisiana Medicaid does not reimburse the technical component separately but rather reimburses for an imaging examination either by full service if the provider owns the equipment or by the PC alone. Other states may have nuanced differences in the way Medicaid reimburses the technical component.

The size of a practice is an important determinant of reimbursement. Because larger practices tend to negotiate higher prices with private health plans than smaller practices, smaller groups may experience less drastic reimbursement decreases under MFA. There is also a smaller reduction in the inpatient payer scenario because Medicare and Medicaid make up larger proportions of payers in this setting. This is discordant with the hypothesis that practices with more privately insured patients will experience more drastic reimbursement reductions compared with those that treat a high percentage of Medicare, Medicaid, and uninsured patients. Louisiana has a high proportion of Medicaid patients and around an average proportion of uninsured patients compared with the national US averages. In accounting for uncompensated care and lower PC fees from Medicaid, this analysis still showed an overall decrease in PC reimbursement in a Medicare single-payer scenario. Neuroradiologists in states with higher proportions of privately insured patients may experience a more drastic impact on reimbursements under MFA.

The analysis does not factor in fluctuations in imaging volume that may occur in a single-payer system. The volume of imaging examinations would likely increase due to decreased...
cost-sharing for patients. Further research is also needed to examine realistic cost-savings opportunities (eg, administrative costs) in evaluating the overall financial impact of MFA on radiology.

The estimated decrease in PC reimbursement in this analysis is based on 2019 Medicare PC fees. The Medicare Physician Fee Schedule for 2020 entailed cuts in reimbursement for radiology to boost reimbursement for care-related evaluation and management services. It is estimated this change will cost radiologists $450 million in 1 year alone, and $5.6 billion during the next decade. We also note that with reimbursement changes following the Patient Protection and Affordable Care Act and Medicare Access and CHIP Reauthorization Act of 2015, it is also possible that fee-for-service may be substantially or completely replaced by alternative payment models (eg, bundled or episode-based payments, capitated payments, global budgets, and government-salaried providers).

Is Medicare for All Better for Our Patients?

Compared with patients in other high-income countries, Americans are more likely to report financial barriers to health care, with a third of US adults reporting difficulty paying medical bills in 2018. This situation is often sensationalized in the media, creating a perception that the US health care industry takes advantage of sick people. Such a belief can undermine public trust in our health care system. Below, we discuss concerns commonly brought up by American patients and concerns regarding MFA.

In 2018, approximately 30.4 million persons in the United States were uninsured. MFA would certainly extend coverage to more patients, and proponents of this system argue that increasing access to preventative care will reduce overall health care costs (eg, by decreasing emergency department use or hospital admissions for acute exacerbations of chronic illnesses). Prior research found that extending Medicaid coverage resulted in more health care use, including emergency department visits, without immediate improvement in health measures such as blood pressure or cardiovascular risk. It did, however, reduce the prevalence of depression and medical debt.

In addition to expanding coverage, proponents argue that MFA would provide more stable coverage. Currently, many Americans receive health insurance through their employers, and changing jobs can result in coverage fluctuations. MFA does not address how the government would feasibly absorb this cost from employers and risks encountering similar issues as other countries if unprepared. Several countries with highly regarded universal health care systems now struggle to finance their health care systems. Since 2010, the National Health Service budget of the United Kingdom has fallen short of its health care costs, resulting in longer wait times for patients, reduced payments to providers, and calls for tax increases. Canada and Australia have faced similar issues of unsustainable rising health care costs, with consequent restrictions on services and increased reliance on private insurance.

The above is not a critique of the principle of universal health insurance but rather of how health care costs have become unsustainable despite a variety of coverage designs. In reforming health care, current politics seems to focus more on who will pay for care rather than on strategies and resources to promote healthier lifestyles. Modern challenges, such as increases in the prevalence of chronic diseases, risk overwhelming the health care system. Although MFA aims to increase access to care, any such legislation will likely have a minimal effect on the root causes of contemporary issues such as sedentary lifestyles and poor diet. While we have focused mainly on financial implications, we acknowledge that there are innumerable other factors to consider in evaluating health care systems.

In summary, implementation of a single-payer system might solve some problems but would likely introduce several others. As key stakeholders in health care, neuroradiologists should scrutinize how emerging legislation will affect their practices and patients.


REFERENCES

12. Medicaid-to-Medicare Fee Index; State Health Facts. 2016. https://www.kff.org/medicaid/state-indicator/medicaid-to-medicare-fee-index/?currentTimeframe=0&selectedRows=%7B%22wrapups%22:%7B%22united-states%22:%7B%7D%7D%7D&sortModel=%7B%22collid %22:%22Location%22,%22sort%22:%22asc%22%7D. Accessed August 15, 2019


26. Torjesen I. Taxes must rise to fund 5% budget increase the NHS needs. BMJ 2018;361:k2337 CrossRef Medline


Direct In Vivo MRI Discrimination of Brain Stem Nuclei and Pathways

T.M. Shepherd, B. Ades-Aron, M. Bruno, H.M. Schambra, and M.J. Hoch

ABSTRACT

BACKGROUND AND PURPOSE: The brain stem is a complex configuration of small nuclei and pathways for motor, sensory, and autonomic control that are essential for life, yet internal brain stem anatomy is difficult to characterize in living subjects. We hypothesized that the 3D fast gray matter acquisition T1 inversion recovery sequence, which uses a short inversion time to suppress signal from white matter, could improve contrast resolution of brain stem pathways and nuclei with 3T MR imaging.

MATERIALS AND METHODS: After preliminary optimization for contrast resolution, the fast gray matter acquisition T1 inversion recovery sequence was performed in 10 healthy subjects (5 women; mean age, 28.8 ± 4.8 years) with the following parameters: TR/TE/TI = 3000/2.55/410 ms, flip angle = 4°, isotropic resolution = 0.8 mm, with 4 averages (acquired separately and averaged outside k-space to reduce motion; total scan time = 58 minutes). One subject returned for an additional 5-average study that was combined with a previous session to create a highest quality atlas for anatomic assignments. A 1-mm isotropic resolution, 12-minute version, proved successful in a patient with a prior infarct.

RESULTS: The fast gray matter acquisition T1 inversion recovery sequence generated excellent contrast resolution of small brain stem pathways in all 3 planes for all 10 subjects. Several nuclei could be resolved directly by image contrast alone or indirectly located due to bordering visualized structures (eg, locus coeruleus and pedunculopontine nucleus).

CONCLUSIONS: The fast gray matter acquisition T1 inversion recovery sequence has the potential to provide imaging correlates to clinical conditions that affect the brain stem, improve neurosurgical navigation, validate diffusion tractography of the brain stem, and generate a 3D atlas for automatic parcellation of specific brain stem structures.

ABBREVIATION: FGATIR = fast gray matter acquisition T1 inversion recovery

The human brain stem represents the complex interdigitation of compact anatomic pathways and nuclei that control or modulate motor, sensory, autonomic, and cognitive functions. Many brain stem structures are essential for survival. Patients can experience profound symptoms and disability even from small focal brain stem lesions due to ischemic stroke, multiple sclerosis, infections, vascular malformations, or tumors. Histopathologic changes in specific brain stem structures can lead to neurodegeneration associated with movement disorders, such as Parkinson disease, progressive supranuclear palsy, and multiple system atrophy. Even with state-of-the-art clinical MR imaging, it remains challenging to confidently localize the spatial locations of specific brain stem structures or to detect early pathologic changes to these structures in individual patients before postmortem examination.

To estimate the location of a particular brain stem structure in current practice, clinicians triangulate on the basis of surface landmarks, limited MR imaging–visible internal features, and mental representations of histology-stained sections from the same superior-inferior level. Systematic dissections of normal cadaver brains may improve the accuracy of this qualitative approach, particularly
to better identify safe entry zones for surgical procedures of the brain stem.\textsuperscript{12,13} However, this overall approach does not account for anatomic distortion that frequently occurs with pathology (ie, no one performs surgery on the normal brain stem). Unfortunately, most stereotactic imaging–based brain atlases have emphasized resolution of cortex, white matter, or specific functional neurosurgery targets within the diencephalon.\textsuperscript{14–16} Detailed image-based parcellation for internal brain stem anatomy remains scant.\textsuperscript{19,20} The widely used FreeSurfer (http://surfer.nmr.mgh.harvard.edu) parcellation provides a single atlas label for the entire brain stem, and the more recent brain stem substructures algorithm only divides the brain stem into “midbrain,” “pons,” and “medulla.”\textsuperscript{21–23}

High-field MR imaging of postmortem, isolated brain stem specimens reveals the potential for MR imaging contrast resolution of internal brain stem anatomy comparable with histology,\textsuperscript{24–27} though external validity is limited because this requires 8+ hours of scanning–dissected, coil–optimized individual specimens that may have altered MR imaging properties due to fixation and the postmortem interval.\textsuperscript{28,29} Improved internal brain stem contrast also can be obtained with in vivo MR imaging using $7T$ systems.\textsuperscript{30–33} Susceptibility mapping,\textsuperscript{34,35} relaxometry,\textsuperscript{30,35,36} or various representations/models of diffusion contrast.\textsuperscript{37–39} Diffusion tractography approaches can putatively identify many internal pathways and features of the brain stem.\textsuperscript{36,39} However, diffusion tractography is prone to spurious tract generation and true-positive segmentations are challenging to validate.\textsuperscript{40} Furthermore, diffusion contrast is often altered by underlying pathology (unfortunately when identifying brain stem structures is most needed clinically).

Several groups have previously reported novel MR imaging sequences that manipulate $T_1$ relaxation to suppress signal from white matter and maximize internal contrast within the thalamus and basal ganglia for direct targeting during functional neurosurgery.\textsuperscript{41–46} These sequences have various acronyms, but perhaps the most relevant was fast gray matter acquisition $T_1$ inversion recovery (FGATIR).\textsuperscript{42} Optimization of image contrast by manipulation of MR imaging sequence parameters was described in detail previously—\textsuperscript{43} the key concept is suppression of signal from voxels that contain myelin, even if the tissue does not contain classic white matter bundles by histology (eg, individual thalamic nuclei). Here, we demonstrate the tremendous advantages of also applying this MR imaging contrast mechanism for the direct identification of many internal brain stem structures in individual living subjects.

**MATERIALS AND METHODS**

This study was compliant with the New York University institutional review board and requirements of the Health Insurance Portability and Accountability Act. The data that support the findings of this study are available from the corresponding author on reasonable request. Healthy adult volunteers were imaged using a 3T Magnetom Prisma scanner with a 64-channel head and neck coil (Siemens). The sequence used here differs minimally from the originally reported FGATIR MR imaging sequence,\textsuperscript{42} or the more recent fluid and white matter suppression (FLAWS) component\textsuperscript{43} of the updated MP2RAGE sequence.\textsuperscript{44} For simplicity, we refer to the sequence used as FGATIR throughout the remainder of this report. For all experiments described below, parallel imaging was not used, the radiofrequency mode was “fast,” and “prescan normalize” was used (per manufacturer). Complete protocol details will be provided to interested readers. We empirically developed a protocol to obtain excellent contrast and spatial resolution to discriminate brain stem structures as determined by consensus between 2 board-certified diagnostic neuroradiologists with expertise in brain stem neuroanatomy.

We first explored how to use signal averaging during image reconstruction to improve contrast resolution of brain stem structures using a 1-mm isotropic resolution sequence with the following parameters: $TR/TE/TI = 3000/2.4/410\ ms$, nonselective $180^\circ$ inversion pulse, flip angle $= 6^\circ$, $256 \times 256$ matrix, $256$-mm-square FOV, $160 \times 1$ mm sagittal slices, bandwidth $= 120$ Hz/pixel, time $= 12$ minutes 9 seconds per 1 average. Due to the relatively long scan time per signal average and potential for subtle-but-compounding motion artifacts, we compared obtaining 4 averages in $k$-space before Fourier image transformation (the standard approach to signal averaging) versus when 4 averages were obtained independently, reconstructed to image space, spatially coregistered using a 6-$df$ rigid-body transform with FMRIB Linear Image Registration Tool (FLIRT; http://www.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT), and finally averaged together. While there is a signal-to-noise penalty for not obtaining multiple averages in $k$-space before Fourier image transformation,\textsuperscript{47} this approach (4 separate 12-minute 9-second scans versus a continuous scan time of 48 minutes 36 seconds) resulted in less image degradation from suble head motion, even in cooperative, tolerant, and experienced volunteers.

Next, we compared contrast resolution between 1- and 0.8-mm isotropic resolution in which the latter sequence parameters were the following: $TR/TE/TI = 3000/2.11/410\ ms$, nonselective $180^\circ$ inversion pulse, flip angle $= 6^\circ$, matrix $= 288 \times 288$, FOV $= 230$ mm square, $192 \times 0.8$ mm sagittal slices, bandwidth $= 460$ Hz/pixel, time $= 14$ minutes 26 seconds per 1 average. The image quality and fine anatomic detail of the 0.8-mm isotropic resolution were preferred by both neuroradiologists. Then, a single subject underwent the 0.8-mm isotropic protocol in 2 separate sessions with 4 and 5 individual averages (57 minutes 44 seconds and 72 minutes 10 seconds, respectively). Within-session averages were coregistered using 6-$df$ rigid-body transforms, and data between the 2 scan sessions were aligned using a 12-$df$ affine transform. Individual averages were then combined sequentially to assess image-quality improvement with additional averages. The 9-average dataset was used for annotating the relevant brain stem anatomy in this report; however, 4 averages were more practical for multiple subjects and appeared to provide sufficient contrast resolution for most individual brain stem structures (per both neuroradiologists). Ten right-handed subjects (5 women, 28.8 ± 4.8 years of age, mean Edinburgh Handedness Inventory score $= 86 \pm 19$) then were each imaged using 4 individual averages of the 0.8-mm FGATIR protocol to assess inter-subject variability in brain stem anatomy. Anatomic assignment of structures was determined by consensus between the 2 neuroradiologists using commonly available references and standard nomenclature.\textsuperscript{48–50} Note that axial images were created parallel to the commissural plane, while coronal images were parallel to the rhomboid fossa (which is not orthogonal to the commissural plane) (On-line Fig 1). For all figures, please see the Table for the specific labeled structures.
Key summary of labels for multiplanar anatomic images of the brain stem (Figs 1–6 and On-line Fig 2)

<table>
<thead>
<tr>
<th>Label</th>
<th>Structure</th>
<th>Label</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mammillary body</td>
<td>32</td>
<td>Principal inferior olivary nucleus</td>
</tr>
<tr>
<td>2</td>
<td>Cerebral peduncle</td>
<td>33</td>
<td>Glossopharyngeal nerve</td>
</tr>
<tr>
<td>3</td>
<td>Pallido- and corticonigral tracts</td>
<td>34</td>
<td>Medullary pyramid</td>
</tr>
<tr>
<td>4</td>
<td>Substantia nigra, pars reticularis</td>
<td>35</td>
<td>Olivocerebellar fibers</td>
</tr>
<tr>
<td>5</td>
<td>Substantia nigra, pars compacta</td>
<td>36</td>
<td>Area postrema</td>
</tr>
<tr>
<td>6</td>
<td>Oculomotor nerve</td>
<td>37</td>
<td>Spinal trigeminal nucleus</td>
</tr>
<tr>
<td>7</td>
<td>Red nucleus</td>
<td>38</td>
<td>Cuneate nucleus</td>
</tr>
<tr>
<td>8</td>
<td>Medial lemniscus</td>
<td>39</td>
<td>Central gray matter</td>
</tr>
<tr>
<td>9</td>
<td>Spinothalamic tract</td>
<td>40</td>
<td>Internal arcuate fibers and sensory decussation</td>
</tr>
<tr>
<td>10</td>
<td>Central tegmental tract</td>
<td>41</td>
<td>Spinocerebellar tracts</td>
</tr>
<tr>
<td>11</td>
<td>Periaqueductal gray</td>
<td>42</td>
<td>Cuneate fasciculus</td>
</tr>
<tr>
<td>12</td>
<td>Medial longitudinal fasciculus</td>
<td>43</td>
<td>Gracile fasciculus</td>
</tr>
<tr>
<td>13</td>
<td>Inferior colliculus</td>
<td>44</td>
<td>Pyramidal decussation</td>
</tr>
<tr>
<td>14</td>
<td>Lateral lemniscus</td>
<td>45</td>
<td>Subthalamic nucleus</td>
</tr>
<tr>
<td>15</td>
<td>Decussation of superior cerebellar peduncles</td>
<td>46</td>
<td>Dentatorubrothalamic tract</td>
</tr>
<tr>
<td>16</td>
<td>Superior cerebellar peduncles</td>
<td>47</td>
<td>Posterior perforated substance</td>
</tr>
<tr>
<td>17</td>
<td>Rubrospinal tract</td>
<td>48</td>
<td>Habenulopeduncular tract</td>
</tr>
<tr>
<td>18</td>
<td>Corticospinal tract</td>
<td>49</td>
<td>Gracile nucleus</td>
</tr>
<tr>
<td>19</td>
<td>Pontocerebellar fibers</td>
<td>50</td>
<td>Lateral geniculate nucleus</td>
</tr>
<tr>
<td>20</td>
<td>Pontine nuclei</td>
<td>51</td>
<td>Medial geniculate nucleus</td>
</tr>
<tr>
<td>21</td>
<td>Trigeminal nerve</td>
<td>52</td>
<td>Superior colliculus</td>
</tr>
<tr>
<td>22</td>
<td>Sensory and motor trigeminal nuclei</td>
<td>53</td>
<td>Posterior commissure</td>
</tr>
<tr>
<td>23</td>
<td>Internal genu of facial nerve</td>
<td>54</td>
<td>Commissure of superior colliculus</td>
</tr>
<tr>
<td>24</td>
<td>Superior olivary nucleus</td>
<td>55</td>
<td>Brachium of the inferior colliculus</td>
</tr>
<tr>
<td>25</td>
<td>Middle cerebellar peduncle</td>
<td>56</td>
<td>Obex</td>
</tr>
<tr>
<td>26</td>
<td>Vestibular nuclear complex</td>
<td>57</td>
<td>Interpeduncular nucleus</td>
</tr>
<tr>
<td>27</td>
<td>Abducens nucleus</td>
<td>58</td>
<td>Tectospinal tract</td>
</tr>
<tr>
<td>28</td>
<td>Facial nerve</td>
<td>59</td>
<td>Trochlear nucleus</td>
</tr>
<tr>
<td>29</td>
<td>Vestibulocochlear nerve</td>
<td>60</td>
<td>Mesencephalic reticular formation</td>
</tr>
<tr>
<td>30</td>
<td>Trapezoid body</td>
<td>61</td>
<td>Ventral trigeminomotor tract</td>
</tr>
<tr>
<td>31</td>
<td>Inferior cerebellar peduncle</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FIG 1.** Superior-to-inferior axial in vivo MR images parallel to the intercommisural plane for the brain stem (number denotes the position below the plane), including the superior midbrain (A, 8.8 mm), inferior midbrain (B, 15.2 mm), midbrain-pons junction (C, 18.4 mm), superior midpons (D, 25.5 mm), inferior midpons (E, 29.5 mm), inferior pons (F, 35.9 mm), superior or “open” medulla (G, 43.1 mm), closed medulla at the sensory decussation (H, 47.1 mm) and motor decussation (I, 55.1 mm). A 26-year-old healthy control female subject, 0.8-mm isotropic resolution, 9 individual datasets coregistered and averaged, 2 imaging sessions with ~2-hour 15-minute total scan time (scale bar = 5 mm). Please see the Table and Results section for explanation of labeled structures.
Finally, to demonstrate clinical feasibility, we performed a single average, 1-mm isotropic resolution FGATIR sequence (~12 minutes) in a patient with hemiplegia after ischemic stroke involving the corticospinal tract.

RESULTS
The FGATIR sequence provided novel MR imaging–derived contrast throughout the in vivo brain stem and allowed the direct identification of many small functionally important structures. Structure identification by 2 neuroradiologists was facilitated by 3-plane cross-referencing. Selected serial axial images illustrate contrast resolution of internal brain stem anatomy at canonical levels (Fig 1). Selected coronal and sagittal images also reveal excellent contrast resolution of internal brain stem anatomy (Figs 2 and 3, respectively). Such longitudinal images demonstrate the superior-inferior course of several brain stem pathways and the potential for the 0.8-mm isotropic FGATIR to enhance our understanding of the 3D organization of the brain stem. Videos of axial, coronal, and sagittal images through the brain stem are provided to interested readers on-line (On-line Videos 1–3).

Within the brain stem, the signal intensity or brightness of a specific structure appeared inversely correlated with the intensity of myelin staining in postmortem histology atlases. Blood in the vertebral arteries and dural venous sinuses appeared the brightest. The next most hyperintense structures were CSF and adjacent gray matter such as the periaqueductal gray and area postrema, the latter a circumventricular organ with an altered extracellular matrix and vascular structure. The third most hyperintense structures were more central gray matter within the brain stem such as the substantia nigra and pontine nuclei. Structures with intermediate hyperintensity were less compact or less densely myelinated structures like the inferior colliculi, central tegmental tract, and corticonigral pathways. Dark structures were classic myelinated brain stem pathways such as the medullary longitudi- nal fasciculus, medial lemniscus, and pontocerebellar fibers. The darkest structures were densely myelinated pathways, including the medullary pyramids, superior/middle cerebellar peduncles, and larger cranial nerves. Cortical bone in the central skull base and the tectorial membrane also appeared very dark. Of note, the signal intensity (and shape) of the corticospinal tract varied along its course, appearing dark within the cerebral peduncles and medullary pyramids, but more diffuse and relatively brighter within the basis pontis. This appearance may represent interdigitiation of pontine nuclei and pontocerebellar projections between white matter fascicles of the corticospinal tract.

Labeling in the figures emphasizes unambiguous structures that can be identified well on the images. Most of these labeled structures are well-known to clinicians familiar with brain stem anatomy (eg, the medial
FIG 4. Direct, unambiguous identification of many myelinated structures in the inferior midbrain (A) also helps improve the spatial accuracy for indirect localization of bordering brain stem nuclei and pathways that are less densely myelinated at this level (B). The ventral trigeminothalamic tract is located between the decussation of the superior cerebellar peduncles and medial lemniscus. Subtle dark bumps on the anterior and posterior aspects of the decussation are consistent with the rubrospinal and tectospinal tracts, respectively. The bright midline structure anterolateral to the medial longitudinal fasciculus should be the substantia nigra pars compacta. The reticular formation is the bright region posterolateral to the medial longitudinal fasciculus and central tegmental tract and anterolateral to the periaqueductal gray. This region will also contain the locus coeruleus. The trochlear nucleus should be along the posterior border of the medial longitudinal fasciculus. The interpeduncular nucleus is off-midline anterior to the decussation of the superior cerebellar peduncles. Similar indirect localization can be derived at other brain stem levels, but is beyond the scope of this initial report.

FIG 5. Axial in vivo MR images of the pontomedullary junction (superior to Fig 1F) selected from 4 different healthy adult controls (2 men) suggest reproducibility of overall image contrast across individuals and sexes (0.8-mm isotropic resolution, 4 averages, 58-minute total scan time, scale bar = 5 mm). The putative location of the pontomedullary reticular formation, a structure of interest for predicting limb function recovery after stroke, is denoted with an asterisk.

longitudinal fasciculus). Directly labeled visible structures also can be used to generate more exact indirect localization for other bordering internal brain stem structures. Figure 4 provides a highly detailed example of this approach for the inferior midbrain where the likely positions of multiple additional structures can be estimated relative to the borders and anatomic spaces formed between the medial longitudinal fasciculus, medial lemniscus, and decussation of the superior cerebellar peduncles. Annotation at this level of detail with descriptions of functional relevance is possible throughout the brain stem, but is beyond the scope of this initial report. Other selected examples of indirect localization that may be of general interest include the oculomotor complex (bright region posteromedial to the central tegmental tract in Fig 1A), pedunculopontine nucleus (bright region medial to the medial lemniscus and lateral to the decussation of the superior cerebellar peduncles, in Fig 1B), locus coeruleus (bright region posterior to the central tegmental tract and medial to the superior cerebellar peduncle in Fig 1C), and facial nucleus (bright region posterolateral to the central tegmental tract and medial to the cerebellar peduncles in Fig 1F). An inferior notch along the medial aspect of the left middle cerebellar peduncle with gray matter signal intensity in Fig 2D is the cochlear nucleus.

There also are subtle features in the images that are difficult to assign with certainty (and sometimes only visualized on 1 side) that likely correspond to specific structures better resolved with postmortem MR microscopy and histology atlases. Limited visualization may be from partial volume effects due to the small size of the structures relative to 0.8-mm isotropic voxels and/or less contrast relative to surrounding nervous tissue. On the right side of Fig 1B, there is an arc-like medially concave dark structure oriented anterior-posterior, consistent with the expected location of the left oculomotor nerve. In Fig 1C, -D, the tectospinal tract appears as a faint dark dotlike structure anterior to the medial lemniscus, particularly on the left side. In Fig 1E, the left tectospinal tract is more obvious, but there is marked asymmetry compared with the subject’s right side. Also, in Fig 1E, the indistinct anterior margin of the medial lemniscus may represent the anteromedial continuation of the trapezoid body. In Fig 2D, an intermediate signal intensity line oriented inferomedial to superolateral, dividing the area postrema (on the right side of the image), may be the sulcus limitans. Some structures also are better recognized when the imaging plane is transverse to the structure. For example, the facial nerve in the lateral midpons is difficult to appreciate in the axial plane (Fig 1F), but better resolved in the coronal plane (Fig 2D).

The FGATIR images did not discriminate all brain stem structures well. It was not possible to distinguish the corticospinal tract from other white matter in the cerebral peduncle (Fig 1A). Boundaries of the medial lemniscus, tectospinal tract, and medial longitudinal fasciculus, which are adjacent to one another in an anterior-posterior orientation within the medulla, remain indistinct (Fig 1G). The inverse relationship of MR imaging signal intensity to white matter content described above also was not entirely consistent. For example, the red nucleus appeared isointense to enveloping cerebellorubral and cerebellothalamic fibers on MR imaging, but these structures are more distinct with classic white matter staining on histology.

Contrast discrimination of structures improved with more signal averages for the FGATIR sequence with 0.8-mm isotropic resolution, but subjectively yielded less improvement beyond 4 averages (Online Fig 2). Thus, a 4-average protocol was used to image 10 healthy volunteers (~1-hour scan time) to assess individual variability in contrast (Fig 5). We observed subtle differences in morphology (eg, compare the medullary pyramids in A and C) or contrast resolution (eg, compare sensory decussation in A and D), but the overall contrast and the structures observed appeared consistent across individuals. In some subjects, the
sulc) and consistent with trans-synaptic degeneration of the tract. The site of ischemic stroke (right basal ganglia and internal capsulomedullary pyramid. These findings were spatially remote from the normal right corticospinal tract signal intensity in the cerebral hemisphere involving the posterior frontoinsular cortex and subjacent basal ganglia (not shown). A, The axial planes for B–D. There is asymmetric absence of the compact dark corticospinal tract (red arrows) in the middle third of the right cerebral peduncle of the midbrain (B), right central basis pontis (C), and right medullary pyramid (D). These images appear very similar to myelin-stained histology changes of the brain stem after chronic right-sided infarct published in a popular neuroanatomy atlas."33 This clinical example illustrates the feasibility of detecting brain stem pathology in patients using the faster, but lower 1-mm isotropic spatial resolution.

corticospinal tract signal intensity appeared asymmetrically darker on the left within the basis pontis (eg, Fig 1E). A similar observation has been reported in postmortem T2-weighted MR imaging microscopy of brain stems.34

To demonstrate clinical feasibility, we used the 1-average, 1-mm isotropic FGATIR (theoretically increasing available intravoxel signal by ~95%) to obtain images in a consented patient with prior right-sided infarct (Fig 6). Here we observed loss of the normal right corticospinal tract signal intensity in the cerebral peduncle and basis pontis, with complete absence of the right medullary pyramid. These findings were spatially remote from the site of ischemic stroke (right basal ganglia and internal capsule) and consistent with trans-synaptic degeneration of the tract.

DISCUSSION

Better discrimination of internal brain stem anatomy is critical for building knowledge about brain stem structure-function relationships in health, normal aging, and disease.32,33 In current clinical practice, however, brain stem structure location can only be inferred indirectly on axial MR images on the basis of a priori knowledge of anatomy derived from histology (or postmortem MR microscopy), the superior-inferior position and surface landmarks, and limited internal contrast on clinical MR imaging. Current research atlases for imaging data like FreeSurfer32,33 provide limited detail for the internal brain stem anatomy. Here, we identified brain stem anatomy using the FGATIR sequence, which uses a short inversion time to suppress signal from myelin. Previous groups using FGATIR (or similar contrast) emphasized applications for whole-brain segmentation or identification of diencephalic and spinal cord structures.41–46 FGATIR not only enables the direct visualization of brain stem structures, but also the brain stem structures identified can be used for more precise indirect localization of additional bordering brain stem structures (eg, Fig 4). Internal brain stem contrast can be generated directly using FGATIR without the need for postprocessing, modeling assumptions, or ultra-high-field MR imaging systems. Contrast was consistent across multiple subjects and was feasible to obtain in patients with stroke.

Other MR imaging methods have provided some contrast resolution of internal brain stem anatomy, but are clinically or technically limited. Submillimeter isotropic resolution MR microscopy images can be obtained from isolated postmortem brain stem samples;44–47 however, these protocols required 8–200 hours of scanning using ultra-high-field magnets and radiofrequency coils smaller than the human head, making this approach impractical for living subjects. Maps of diffusion parameters and tractography provide improved contrast resolution of internal brain stem anatomy.36,39 However, diffusion MR imaging resolution of submillimeter brain stem pathways is challenging in vivo due to signal-to-noise limitations. Reliance on high-resolution diffusion MR imaging for anatomic localization can also be problematic due to inherent geometric distortion, particularly in structures near the skull base, and results in spatial misregistration with volumetric sequences. Unlike FGATIR, diffusion MR imaging contrast depends on complex mathematic representations or models that have proved challenging to validate in vivo.40 FGATIR may provide complementary contrast for validating brain stem diffusion tractography or for independently extracting diffusion values from brain stem structures/regions. Joint reconstruction of multiple MR imaging contrasts, including FGATIR, may provide the best contrast resolution of brain stem anatomy.55

The 1-mm isotropic FGATIR acquisition with 1 average (~12-minute acquisition) is a relatively long acquisition for clinical MR imaging protocols, but appears feasible in patients (Fig 6), particularly in specific clinical circumstances that justify such a long scan. For example, the FGATIR sequence was originally developed for functional neurosurgery to improve deep brain stimulator electrode placement within the subthalamic nucleus.42 We plan to implement FGATIR in clinical patients to better localize specific brain stem structures affected by focal lesions (eg, infarcts, demyelinating, or inflammatory lesions). This sequence also may prove helpful for planning surgical biopsies or anatomic corridors for tumor resection.12,13 The 800-μm isotropic data require 3–4 averages for useful contrast resolution, which may be susceptible to motion-related image degradation in noncompliant individuals. This concern may be addressed by applying several recently developed methods to accelerate FGATIR acquisitions, including sliding window reconstructions of k-space,56 compressed sensing image reconstruction of undersampled data,57 and deep learning approaches to improve suboptimal image contrast produced by shortened scan times.58,59 Clinical feasibility also can be supported through the use of a diffeomorphic template for coregistration of limited contrast...
FGATIR images from individual patients or by training a neural network to recognize small brain stem structures with limited FGATIR contrast.

Using FGATIR to depict brain stem anatomy has other potential limitations. Not all brain stem structures were directly resolved by FGATIR. We identified brain stem structures on the basis of correspondences to labeled myelin-stained histology images in published atlases, but direct structure-function correlations in the same brain would be helpful. Like diffusion MR imaging, FGATIR image contrast in the brain stem also could change with aging or disease. Such changes occur in other non-diffusion-weighted scans of the brain stem such as neuromelanin MR imaging in subjects with Parkinson disease. In our ongoing work, FGATIR-based contrast in older patients with stroke and essential tremor appears similar to that in age-matched healthy controls and younger subjects (unpublished data). However, the effects of aging and disease on FGATIR contrast require systematic investigation.

To date, the teaching, clinical targeting, and research of brain stem structures have largely relied on neuroanatomic atlases that have limitations. Most brain stem plates in the Haines atlas are from a single individual with trans-synaptic degeneration of the corticospinal tract due to internal capsule infarction. The Schaltenbrand brain atlas, widely used for functional neurosurgery in the United States, relies predominately on 3 thalami of different sizes and limited spatial overlap. This atlas also provided annotated images of only 1 brain stem in the axial plane. MR imaging data with FGATIR contrast have some advantages over traditional neuroanatomy atlases that use postmortem histology. First, the 3D imaging datasets enable simultaneous localization and visualization of specific structures in multiple planes. Depiction of brain stem anatomy is not limited to the plane of section like histologic sampling, and specific oblique image reconstructions can be obtained to emphasize certain key spatial relationships. Second, collections of MR imaging data are less expensive and less time-intensive, providing greater opportunity to characterize brain stem structure variations across multiple individual brains. We intend to collect FGATIR data from a large cohort of subjects to create a diffeomorphic template with atlas labels for the human brain stem. These data could be used to investigate size and shape differences in the brain stem and as an independent method for extracting quantitative data from MR imaging and PET data in future investigations of small brain stem structure changes associated with aging and different diseases.

**CONCLUSIONS**

The FGATIR sequence provided excellent novel contrast resolution of internal brain stem pathways and nuclei at 3T in healthy living subjects. The 1-mm isotropic resolution protocol was clinically feasible (~12-minute acquisition) and may immediately help with structure-function correlations or surgical corridor planning in the brain stem for individual patients. The FGATIR sequence also provided image contrast that may be helpful to complement and/or validate other advanced MR imaging methods for brain stem anatomy such as diffusion tractography.


**REFERENCES**

1. Carpenter MB, Strong OS, Truex RC. *Human Neuroanatomy* (Formerly Strong and Elwyn’s *Human Neuroanatomy*). 7th ed. Lippincott Williams & Wilkins; 1976


Acetazolamide-Loaded Dynamic 7T MR Quantitative Susceptibility Mapping in Major Cerebral Artery Steno-Occlusive Disease: Comparison with PET


ABSTRACT

BACKGROUND AND PURPOSE: Dynamic changes in cerebrovascular reactivity after acetazolamide administration vary markedly among patients with major cerebral arterial steno-occlusive disease. MR quantitative susceptibility mapping can dynamically quantify the cerebral magnetic susceptibility. The purpose of this study was to determine whether dynamic changes in susceptibility after administration of acetazolamide on 7T quantitative susceptibility mapping are associated with pre-existing states of CBV and the cerebral metabolic rate of oxygen in the cerebral hemispheres with major cerebral arterial steno-occlusive disease.

MATERIALS AND METHODS: Sixty-five patients underwent 7T MR imaging at baseline and at 5, 10, 15, and 20 minutes after acetazolamide administration. Differences between the susceptibility of venous structures and surrounding brain tissue were calculated in the quantitative susceptibility mapping images. Susceptibility differences at 5, 10, 15, and 20 minutes after acetazolamide administration relative to baseline were calculated in 97 cerebral hemispheres with major cerebral arterial steno-occlusive disease. CBV and the cerebral metabolic rate of oxygen were also calculated using 15O-gas PET in the resting state.

RESULTS: Dynamic changes of susceptibility after acetazolamide administration were classified into 3 patterns: abnormally increasing 5 or 10 minutes after acetazolamide administration; abnormally decreasing within 20 minutes after acetazolamide administration; and remaining unchanged after acetazolamide administration. CBV was significantly greater in the first pattern than in the latter 2. The cerebral metabolic rate of oxygen differed significantly in descending order from the first to middle to last pattern.

CONCLUSIONS: Dynamic changes of susceptibility after acetazolamide administration on 7T MR quantitative susceptibility mapping are associated with pre-existing states of CBV and the cerebral metabolic rate of oxygen in major cerebral arterial steno-occlusive disease.

ABBREVIATIONS: ACZ = acetazolamide; BOLD = blood oxygen level-dependent; CMRO2 = cerebral metabolic rate of oxygen; CVR = cerebrovascular reactivity; OEF = oxygen extraction fraction; QSM = quantitative susceptibility mapping; RS = relative susceptibility

When the chronic progression of major cerebral arterial steno-occlusive disease reduces cerebral perfusion pressure, cerebrovascular autoregulatory mechanisms attempt to preserve CBF through the dilation of precapillary resistance vessels; this is referred to as stage 1 ischemia.1-3 In stage 1 ischemia, cerebrovascular reactivity (CVR) to acetazolamide (ACZ), which reflects the degree of cerebrovascular autoregulatory vasodilation, starts to decrease.1-3 However, autoregulatory capacity cannot compensate for further reductions in cerebral perfusion pressure, which result in a decline in CBF; this is referred to as stage 2 ischemia, or misery perfusion.1-3 CVR to ACZ is exhausted in this stage.1-3 In the clinical setting, CVR to ACZ is measured using perfusion SPECT,4 dynamic susceptibility contrast perfusion MR imaging,5 and transcranial Doppler ultrasonography.6

MR imaging is capable of quantifying the blood oxygen level-dependent (BOLD) changes in venous structures and/or brain parenchyma induced by differences in magnetic susceptibility between oxy- and deoxyhemoglobin.7-9 Wu et al10 continuously measured CVR to ACZ using BOLD imaging with 3T MR imaging in patients with steno-occlusive disease of major cerebral arteries and demonstrated that overall, BOLD signal began...
increasing immediately after ACZ administration, approaching a plateau at ~8.5 minutes after administration. However, in some patients, BOLD signal showed a more severely reduced relative reactivity at 6–7 minutes after the initiation of ACZ infusion and progressive partial recovery toward a pre-ACZ administration value at 10–11 minutes. These findings suggested that dynamic changes to susceptibility after ACZ administration vary markedly among such patients. CVR to ACZ itself may depend on the pre-existing state of vasodilation, which can be assessed using CBV, and pre-existing dilated vessels show little or no response to ACZ. We hypothesized that a pre-existing state of cerebral metabolism is another condition impacting CVR to ACZ; ACZ acts via enzyme reaction, and a reduction in this enzyme reaction causes a slow, decreased CVR to ACZ. Cerebral metabolism, which can be assessed using the cerebral metabolic rate of oxygen (CMRO₂), may be reduced in such a state.

A postprocessing technique known as quantitative susceptibility mapping (QSM) on MR imaging can quantify the magnetic susceptibility of venous structures and/or brain parenchyma from T₂*-weighted magnitude/phase images, which are easily obtained using a commercial scanner. Several investigators have measured CVR to ACZ using QSM on 3T MR imaging and suggested that QSM is sensitive to dynamic modulation of the oxygen extraction fraction (OEF) during hemodynamic augmentation. A recent study using QSM with a 7T scanner reported that it offered profound susceptibility effects and optimized postprocessing techniques. The scan time for this mapping is relatively short (3 minutes 25 seconds), and it may also enable imaging of dynamic changes in blood oxygenation levels in venous structures and/or brain parenchyma.

Therefore, the purpose of the present study was to determine whether dynamic changes in susceptibility after ACZ administration on 7T QSM are associated with pre-existing CBV and CMRO₂ on PET in patients with ICA or MCA steno-occlusive disease.

MATERIALS AND METHODS

All procedures performed in studies involving human participants were in accordance with the institutional ethics committee, and written, informed consent was obtained from all subjects or their next of kin before patient participation.

Healthy Subjects

To obtain healthy control values from brain ¹⁵O-gas PET studies, we enrolled 10 healthy adult men (mean age, 47 ± 8 years; age range, 35–60 years) who had no history of diabetes mellitus, dyslipidemia, or hypertension and had an absence of asymptomatic lacunar infarction and leukoaraiosis on conventional brain MR imaging.

Patient Inclusion Criteria

We prospectively selected patients who provided written informed consent to participate and who met the following basic inclusion criteria: presence of clinical symptoms suggesting ischemic episodes in the MCA or ICA territory at 1–6 months before visiting our institution; useful residual function (0, 1, or 2 on the mRS); presence of uni- or bilateral MCA (M1 portion) or extra- or intracranial ICA stenosis (>70% or 50% for the ICA and MCA, respectively) or occlusion on angiography; and absence of infarcts in the entire cortical area supplied by the M4 branch of the MCA on T₁WI and T₂WI, and DWI.

MR Imaging Protocol and Generation of QSM

A 7T MR scanner (Discovery MR950; GE Healthcare) with quadrature transmission and 32-channel receive head coils was used. QSM source data were obtained using a 3D spoiled gradient recalled acquisition technique with flow compensation (TR, 30 ms; TE, 15 ms; flip angle, 20°; FOV, 256 mm; acquisition matrix size, 512 × 256; section thickness, 2 mm; number of slices, 160; reconstruction voxel size after zero-fill interpolation, 0.5 mm³; and scan time, 3 minutes 25 seconds). Magnitude and real/imaginary phase images were regenerated from this acquisition.

First, each patient underwent the above-mentioned MR imaging at baseline. Next, ACZ (1000 mg; range, 13–19 mg/kg body weight) was dissolved in physiologic saline (20 mL); this solution was then administered intravenously for 1 minute. The MR imaging was performed in the same fashion 4 times so that each midscan time was 5, 10, 15, and 20 minutes after the end of intravenous administration of ACZ.

Before the start of this ACZ challenge study, 30 other cerebral hemispheres with ICA or MCA steno-occlusive disease underwent 5 MR imaging scans with the same timing without ACZ administration to assess the minute-to-minute test-retest variability of susceptibility as measured using QSM.

QSM images were generated from the source images using an in-house program with a multiple dipole-inversion combination with k-space segmentation and regularization-enabled sophisticated harmonic artifact reduction for phase data methods, as described previously. A 2D Gaussian low-pass filter with a kernel size of 60% of the total image power in each section was applied to extract iron deposition in deep nuclei, hemosiderin deposition, dural sinuses, and large venous structures, and a 2D Gaussian high-pass filter of 2% was applied to extract small venous structures and surrounding brain tissue was calculated in each hemisphere with ICA or MCA steno-occlusive disease under- went 5 MR imaging scans with the same timing without ACZ administration to assess the minute-to-minute test-retest variability of susceptibility as measured using QSM.

Brain ¹⁵O-Gas PET Study

Patients who were scheduled to undergo MR imaging studies with the ACZ challenge underwent PET studies in a resting state without ACZ challenge. PET studies were performed using a SET-3000GCT/M scanner (PET/CT; Shimadzu) (full width at half maximum for in-plane and axial spatial resolutions of 3.5 and 4.2 mm, respectively) within 5 days before MR imaging. Before the PET scans, a 3-minute transmission scan with a
cesium 137 ($^{137}$Cs) point source was conducted using a bismuth germanate transmission detector ring that was coaxially attached to the gadolinium silica oxide emission detector ring. Next, CBF was measured while the patient inhaled C$^{15}$O$_2$ continuously through a mask. The CMRO$_2$ and OEF were measured during continuous inhalation of C$^{15}$O$_2$; these data were collected for 5 minutes. CBV was measured during a single breath of C$^{15}$O. Finally, CBF, OEF, and CMRO$_2$ were calculated using the steady-state method and then corrected on the basis of CBV.

**Image Data Analyses**

Image data analyses were performed according to previously reported methods. By means of SPM, Version 12 (http://www.filion.ucl.ac.uk/spm/software/spm12), PET images that were coregistered to QSM source images, as well as QSM images after Gaussian smoothing ($\sigma = 10$ pixel), were warped to Montreal Neurological Institute coordinates. Next, 318 constant ROIs were placed in both the cerebral and cerebellar hemispheres automatically using a 3D stereotaxic ROI template with SPM. The ROIs were then grouped into a total of 10 segments—callosomarginal, pericallosal, precentral, central, parietal, angular, temporal, posterior, hippocampal, and cerebellar—in each hemisphere according to the arterial supply. Five of these 10 segments (precentral, central, parietal, angular, and temporal) perfused by the MCA were then combined and defined as an MCA ROI (Fig 1).

Mean susceptibility differences on QSM images and CBF, CBV, OEF, and CMRO$_2$ on PET images were measured using image-analysis software (ITK-SNAP; www.itksnap.org) in the MCA ROI in the cerebral hemisphere with ICA or MCA stenooclusive disease. In each MCA ROI of each cerebral hemisphere with ACZ challenge, relative susceptibility (RS, %) was calculated as follows: 100 × susceptibility difference at each time (5, 10, 15, or 20 minutes) after ACZ administration on QSM images/susceptibility difference in the baseline on QSM images. In each MCA ROI of each cerebral hemisphere without ACZ challenge, RS was also calculated as follows: 100 × susceptibility difference on QSM images on the basis of the presence or absence of an abnormal increase or decrease in RS at any time point.

In each MCA ROI of each patient, when CBV was greater than the upper limit of the 95% CI of control values obtained from 20 cerebral hemispheres of 10 healthy subjects, this ROI was defined as having abnormally elevated CBV; when the CMRO$_2$ was lower than the lower limit of the 95% CI of control values, this ROI was defined as showing abnormally reduced CMRO$_2$.

**Statistical Analysis**

Data are expressed as the mean ± SD. Pearson correlation coefficients were used to assess correlations between baseline QSM-OEF and pre-existing PET-OEF. A difference in RS at each time point after ACZ administration was evaluated using the Friedman test. When this difference was significant, differences among RS values at each time after ACZ administration were compared using the Wilcoxon signed rank test followed by the Holm method. A difference in CBF, CBV, OEF, or CMRO$_2$ among ≥3 groups was evaluated using the Kruskal-Wallis test. When this difference was significant, differences among CBF, CBV, OEF, or CMRO$_2$ in each group were compared using the Mann-Whitney U test followed by the Holm method. A difference in CBV or CMRO$_2$ among 2 groups was evaluated using the Mann-Whitney U test. $P$ values < .05 were considered statistically significant for all assessments.

**RESULTS**

During the 30-month study period, a total of 89 patients successfully underwent 77 MR imaging without severe adverse effects. In 4 of these patients, data sufficient to generate QSM were not obtained because of motion artifacts; these 4 patients were excluded from the analysis. Therefore, 85 patients (65 and 20 with and without ACZ challenge, respectively) were finally enrolled in the study.
Among these 85 patients (46 men, 39 women; mean age, 55 ± 12 years; age range, 30–76 years), 60 had hypertension, 32 had diabetes mellitus, and 55 had dyslipidemia. Thirty-seven, 27, 3, and 17 patients had bilateral ICA (bilateral extracranial diseases in 27; bilateral intracranial disease in 5; unilateral extracranial disease and contralateral intracranial disease in 5), unilateral ICA (extracranial disease in 22; intracranial disease in 5), the M1 portion diseases of bilateral MCA, and unilateral M1 portion steno-occlusive disease, respectively. The remaining patient had unilateral extracranial ICA and contralateral M1 portion steno-occlusive diseases. Therefore, 127 MCA ROIs of the 127 cerebral hemispheres with MCA or ICA steno-occlusive disease (97 and 30 with and without ACZ challenge, respectively) were eventually analyzed.

The 95% CI of RS values in the second, third, fourth, and fifth MR imaging in 30 MCA ROIs without ACZ challenge were 91.85%–108.55%, 90.81%–109.33%, 90.59%–109.49%, and 90.70%–109.22%, respectively. No difference was observed among each RS.

The 95% CIs of CBV and CMRO2 obtained from 20 MCA ROIs of 10 healthy subjects were 2.45–5.01 mL/100 g and 2.31–4.70 mL/100 g/min, respectively.

A significant correlation was observed between baseline QSM-OEF and pre-existing PET-OEF \( (r = 0.865; P \leq .0001) \) (Fig 2).

Figure 3 shows chronological changes in RS after ACZ administration for the 97 MCA ROIs. Patterns of these changes were classified into the following 3 groups: RS was abnormally increased at ≥1 time point in 22 MCA ROIs; RS was abnormally decreased at ≥1 time point in 47 MCA ROIs; and RS remained within the 95% CI of controls at all time points (defined as unchanged RS) in 28 MCA ROIs. None of the MCA ROIs exhibited an abnormally increased RS and an abnormally decreased RS at different time points.

Of the 22 MCA ROIs with abnormally increased RS, 2 exhibited abnormally increased RS at all time points; 1 exhibited abnormally increased RS5, RS10, and RS15; eleven exhibited abnormally increased RS5 and RS10; one exhibited abnormally increased RS5, and RS20; five exhibited abnormally increased RS5 alone; and 2 exhibited abnormally increased RS10 alone.

Of the 47 MCA ROIs with abnormally decreased RS, 18 exhibited abnormally decreased RS at all time points; 3 exhibited abnormally decreased RS5, RS10, and RS15; three exhibited abnormally decreased RS5, RS10, and RS20; two exhibited abnormally decreased RS5, RS15, and RS20; one exhibited abnormally decreased RS10, RS15, and RS20; one exhibited abnormally decreased RS10, RS15, and RS20; one exhibited abnormally decreased RS5 and RS10; one exhibited abnormally decreased RS15 and RS20; nine exhibited abnormally decreased RS5 alone; 3 exhibited abnormally decreased RS10 alone; 2 exhibited abnormally decreased RS15 alone; and 4 exhibited abnormally decreased RS20 alone.

In the 22 MCA ROIs with abnormally increased RS, RS differed significantly among 4 time points: RS5 \((113.84 \pm 6.28\%); P = .0081\) for RS10, \(P < .0001\) for RS15, \(P < .0001\) for RS20, RS10 \((111.58 \pm 5.06\%); P = .0001\) for RS15, \(P < .0001\) for RS20, RS15 \((104.72 \pm 5.43\%); P = .0019\) for RS20, and RS20 \((100.93 \pm 5.55\%)\) in descending order (Fig 3). In the 47 MCA ROIs with abnormally decreased RS, no significant differences were observed among RS5 \((88.69 \pm 5.56\%); RS10 \((88.55 \pm 5.98\%); RS15 \((88.71 \pm 5.87\%); and RS20 \((89.00 \pm 5.83\%). In the 28 MCA ROIs with unchanged RS, no significant differences were observed among RS5 \((97.89 \pm 4.89\%); RS10 \((99.06 \pm 4.47\%); RS15 \((97.85 \pm 3.74\%); and RS20 \((97.80 \pm 3.78\%).

The On-line Table shows comparisons of pre-existing CBV and CMRO2 among MCA ROIs with abnormally increased RS, abnormally decreased RS, unchanged RS, and MCA ROIs in healthy subjects as controls. CBV was significantly greater in

![FIG 2. Correlations between baseline QSM-OEF before ACZ administration and pre-existing PET-OEF in 97 MCA ROIs.](image)

![FIG 3. Chronological changes in RS after ACZ administration in 97 MCA ROIs. Left, middle, and right graphs show the 22 MCA ROIs with abnormally increased RS at ≥1 time point, the 47 MCA ROIs with abnormally decreased RS at ≥1 time point, and the 28 MCA ROIs without abnormally increased or decreased RS at any time point, respectively. Red, blue, and black lines denote ROIs with abnormally elevated CBV, an abnormally decreased cerebral metabolic rate of oxygen, and neither, respectively. Upper and lower dotted horizontal lines denote the upper and lower limits of the 95% CI of controls without ACZ challenge, respectively.](image)
MCA ROIs with abnormally increased RS than in those with abnormally decreased RS, unchanged RS, or controls. No significant differences in CBV were identified in other any comparisons. Of the 22 MCA ROIs with abnormally increased RS, 17 (77%) had abnormally elevated CBV. In contrast, none of the patients with abnormally decreased RS or unchanged RS had abnormally elevated CBV. No significant differences in CMRO₂ were apparent between controls and MCA ROIs with abnormally increased RS. Regarding other comparisons, the CMRO₂ differed significantly in descending order, as follows: controls or abnormally increased RS, abnormally decreased RS, and unchanged RS. Of the 28 MCA ROIs with unchanged RS, 8 (29%) showed abnormally reduced CMRO₂. In contrast, of the 47 MCA ROIs with abnormally decreased RS, only 7 (14%) had abnormally reduced CMRO₂, and no MCA ROIs with abnormally increased RS had abnormally reduced CMRO₂. Furthermore, in the 47 MCA ROIs with abnormally decreased RS, whereas CBV did not differ significantly in the 11 MCA ROIs with unchanged RS (3.35 ± 0.84 mL/100 g) and the 36 MCA ROIs with abnormally decreased RS (3.40 ± 0.3862 mL/100 g), CMRO₂ was significantly lower in the former MCA ROIs (2.65 ± 0.52 mL/100 g/min) than in the latter (3.06 ± 0.38 mL/100 g/min; \( P = .0182 \)).

Representative chronological images of susceptibility difference obtained from QSM after ACZ administration and PET images performed before MR imaging in a patient with abnormally increased CBV in the symptomatic cerebral hemisphere are shown in Fig 4.

**DISCUSSION**

The present study demonstrated that dynamic changes in susceptibility after ACZ administration on 7T QSM are associated with pre-existing CBV and CMRO₂ on PET in patients with ICA or MCA steno-occlusive disease.

In this study, ACZ was administered intravenously for 1 minute. This method has been applied for measuring CVR on brain perfusion SPECT in adult patients with ICA or MCA steno-occlusive disease. Considering the adverse effects of ACZ administration performed in the same fashion, 63% of patients reportedly developed only minor symptoms such as headache, nausea, dizziness, tinnitus, numbness of the extremities, and general malaise, and these symptoms resolved within 72 hours after ACZ administration. These patients experienced neither severe general symptoms nor neurologic deficits. Similarly, no patients in the present study experienced such severe adverse effects due to ACZ administration.

A study using ¹⁵O PET demonstrated that while CBF increased at 10 and 20 minutes after ACZ administration in healthy humans, CMRO₂ was unchanged compared with preadministration. Changes in susceptibility after ACZ administration on QSM theoretically reflect both changes in CBF and oxygen metabolism by ACZ. Although whether CMRO₂ after ACZ administration is constant in patients with cerebrovascular diseases remains unknown, dynamic changes in susceptibility after ACZ administration on QSM may more strongly reflect changes in CBF than changes in oxygen metabolism: Increases and decreases in the susceptibility may predominantly indicate decreases and increases in CBF, respectively. One study of CVR to ACZ using QSM on 3T MR imaging showed that OEF decreased from baseline to 16 minutes after ACZ administration in the cerebral hemisphere with anterior circulation steno-occlusive disease, as well as in the cerebral hemisphere without such disease, and the decrease was more substantial in the latter. These changes in susceptibility might be caused by changes of CBF.

In the present study, dynamic changes of susceptibility after ACZ administration were classified into 3 patterns: abnormally increasing, peaking at 5 or 10 minutes after ACZ administration and later returning toward baseline; abnormally decreasing within 20 minutes after ACZ administration and maintaining this decrease; and remaining unchanged after ACZ administration. A dynamic change comprising an abnormal increase in susceptibility early after ACZ administration and subsequent return was comparable with a finding obtained using continuous measurements of BOLD signal after ACZ administration. In MCA ROIs with such dynamic changes, pre-existing CBV was abnormally elevated and pre-existing CMRO₂ was normally maintained. Mean transit time (MMT) is a function of CBV/CBF, and an increase in CBV usually implies a long MMT in MCA or ICA steno-occlusive disease. A long MMT can result in low drug delivery, leading to little or no CVR immediately after ACZ administration in the region affected by such a condition, whereas CVR in surrounding or other brain tissues with a normal MMT reacts rapidly because of normal drug delivery. In addition, the small arterioles or intraparenchymal vessels are dilated in regions with an increase in CBV. These pre-existing dilated
vessels cannot respond immediately after ACZ administration, whereas normal vessels in surrounding or other brain tissue rapidly dilate. These differences in CVR to ACZ administration lead to a sudden reduction in cerebral perfusion pressure of normal vessels relative to pre-existing dilated vessels. This dynamically developed pressure gradient may result in the early steal of blood flow from pre-existing dilated vessels to normal vessels. Subsequently, pre-existing dilated vessels may gradually respond to ACZ, and a vasodilation effect of ACZ in normal vessels may gradually reduce with time. Such mechanisms could explain the abnormal increase in susceptibility early after ACZ administration (early steal phenomenon) and subsequent return.

A dynamic change of the abnormal decrease in susceptibility within 20 minutes after ACZ administration and maintenance of this decrease might imply an increase in CBF—that is, normal CVR to ACZ. Actually, in MCA ROIs showing such a dynamic change, pre-existing CBV was normal. Pre-existing CMRO2 was reduced compared with that in controls but remained within normal limits in 86% of MCA ROIs showing that susceptibility change. On the other hand, pre-existing CMRO2 was reduced more in MCA ROIs with an abnormal decrease in susceptibility ≥10 minutes after ACZ administration than in those with an abnormal decrease in susceptibility 5 minutes after ACZ administration. Because CBV did not differ significantly between these 2 groups, the difference in dynamic change after ACZ administration might be attributable to pre-existing CMRO2. ACZ acts via carbonic anhydrase, and a reduction in its activity slows CVR to ACZ. The activity of carbonic anhydrase may depend on cerebral metabolism. This hypothesis may also explain the maintenance of unchanged susceptibility after ACZ administration. In MCA ROIs with such a dynamic change in susceptibility, pre-existing CBV was normal, but pre-existing CMRO2 was quite markedly reduced; this finding suggests that the activity of carbonic anhydrase was severely reduced and cerebral vessels were unresponsive to ACZ.

Using continuous measurement of CVR to ACZ with BOLD imaging, Wu et al demonstrated that BOLD signal began increasing gradually following ACZ administration, approaching a plateau at ~8.5 minutes after administration in the cerebral hemisphere without ICA or MCA steno-occlusive disease. This dynamic change in susceptibility differed compared with that of MCA ROIs with normal CBV and normal CMRO2 in the present study: an abnormal decrease in susceptibility 5 minutes after ACZ administration and maintenance of this decrease. The difference in peak susceptibility may be due to the method of ACZ administration. ACZ (1000 mg) was administered intravenously for 3–5 minutes in the study by Wu et al and for 1 minute in the present study.

A preoperative decrease in CVR to ACZ or a preoperative increase in CBV in the cerebral hemisphere ipsilateral to carotid endarterectomy or carotid artery stent placement for cervical ICA stenosis or arterial bypass surgery for Moyamoya disease has been identified as a powerful and independent predictor of postoperative cerebral hyperperfusion. Postoperative cerebral hyperperfusion induces unilateral headache, face and eye pain, seizure, and focal symptoms that occur secondary to cerebral edema or intracerebral hemorrhage. This phenomenon, even when asymptomatic, induces slight-but-diffuse brain damage, resulting in postoperative cognitive decline. Identification of reduced CVR to ACZ or elevated CBV therefore enables improved risk stratification for such patients. However, the predictive accuracy of CVR to ACZ on brain perfusion SPECT is not always high, with a positive-predictive value of <50%. Our findings that dynamic changes in susceptibility, particularly susceptibility 5 minutes after ACZ administration, detected increased CBV with high positive (77%) and negative (100%) predictive values suggests that these dynamic changes may predict the risk of cerebrovascular adverse events more accurately than conventional measurements of CVR to ACZ. In addition, the examination time of the present method using 7T QSM was <12 minutes if the susceptibility change was measured until a midscan time of 5 minutes after ACZ administration.

QSM source data were obtained with a section thickness of 2 mm. Such a thick section may reduce the accuracy of susceptibility on QSM due to the partial volume effect for small veins. Another limitation, the scan time for 7T QSM is 3 minutes 25 seconds longer than that for BOLD or QSM imaging with 3T MR imaging. Shortening the scan time may help clarify dynamic changes in susceptibility after ACZ administration. Furthermore, considerable variation in cerebral hemodynamics might be anticipated within an MCA ROI, especially in the borderzone, and use of the MCA ROI might reduce such hemodynamic variation. Ranges of CBV in the cerebral gray matter are quite large. This high heterogeneity might also be reduced using MCA ROI analysis.

CONCLUSIONS

The present study demonstrated that the dynamic changes in susceptibility after ACZ administration on 7T QSM are associated with pre-existing CBV and CMRO2 on PET in patients with ICA or MCA steno-occlusive disease.

REFERENCES

MRI study in patients with high-grade unilateral internal carotid artery stenosis. Stroke 2000;31:1561–65 CrossRef Medline
Spatial Resolution and the Magnitude of Infarct Volume Measurement Error in DWI in Acute Ischemic Stroke

J.M. Ospel, A. Jaffray, V. Schulze-Zachau, S. Kozerke, and C. Federau

ABSTRACT

BACKGROUND AND PURPOSE: Infarct volume in acute ischemic stroke is an important prognostic marker and determines endovascular treatment decisions. This study evaluates the magnitude and potential clinical impact of the error related to partial volume effects in infarct volume measurement on diffusion-weighted MR imaging in acute stroke and explores how increasing spatial resolution could reduce this error.

MATERIALS AND METHODS: Diffusion-weighted imaging of 393 patients with acute stroke, of whom 56 had anterior circulation large-vessel occlusion, was coregistered to standard space. Lesion boundaries were manually segmented. A 3D lesion-volume model was resampled for voxel sizes from 4 × 4 × 8 to 1 × 1 × 2 mm, and the surface-volume, corresponding to the partial volume error, was calculated. The number of cases with anterior circulation large-vessel occlusion, in which the endovascular therapy core threshold of 70 mL was contained within the margin of error, was calculated as a function of imaging resolution.

RESULTS: The mean infarct core volume was 27.2 ± 49.9 mL. The mean surface volume was 14.7 ± 20.8 mL for 2 × 2 × 4 mm resolution and 7.4 ± 10.7 mL for 1 × 1 × 2 mm resolution. With a resolution of 2 × 2 × 4 mm, 70 mL was contained within the margin of error in 7/56 cases (12.5%) with large-vessel occlusion, while with a 1 × 1 × 2 mm voxel size, the margin of error was 3/56 (5%). The lesion-volume range of potentially misclassified lesions dropped from 46.5–94.1 mL for a 2 × 2 × 4 mm resolution to 64.4–80.1 mL for a 1 × 1 × 2 mm resolution.

CONCLUSIONS: Partial volume effect is an important source of error in infarct volume measurement in acute stroke. Increasing spatial resolution substantially decreases the mean error. Standard use of high-resolution DWI should be considered to increase the reliability of infarct volume measurements.

ABBREVIATION: LVO = large-vessel occlusion

Infarct core volume is a known predictor of outcome in patients with acute stroke; smaller infarct cores are associated with better outcomes and higher rates of functional independence. In patients with a small infarct core, thrombectomy leads to a more favorable result.1-4 Diffusion-weighted imaging is considered the criterion standard for infarct core measurement. A large infarct core, usually above 70 mL, was considered a contraindication for endovascular therapy in many clinical trials and is commonly used as a threshold in clinical practice.5-7

The accuracy of DWI-based infarct core measurement is limited by several factors. For example, although DWI hyperintense lesions are generally considered to represent irreversibly damaged tissue, reversible DWI lesions have been reported.8,9 Another important potential source of error is partial volume effects at the border of the lesion due to the finite spatial resolution reconstruction.

The surface volume (ie, the volume of the surface voxels) represents the volume in which misclassification can occur due to the partial volume effect and is an established measure of spatial resolution–induced measurement error.10 This error is equal to the number of voxels composing the surface, multiplied by the individual voxel volume used in image reconstruction (Fig 1). Because of the complex shapes of most large stroke lesions (eg, those with a gyriform pattern), volume measurement error due
to partial volume effects could potentially lead to an erroneous choice of therapy.

The purpose of our study was to assess the magnitude and potential therapeutic impact of errors related to partial volume effects in infarct core volume measurement on DWI in acute stroke. Furthermore, we aimed to evaluate to what extent increasing the spatial resolution could reduce this error.

MATERIALS AND METHODS

Patients

Approval from the local ethics committee was obtained for this anonymized, retrospective evaluation. MR imaging examinations of patients who presented in the emergency department of the University Hospital of Basel with symptoms of acute ischemic stroke between August 2010 and November 2017 were reviewed and considered for this study. A subgroup analysis was performed in the group with anterior circulation large-vessel occlusion (LVO) who be considered endovascular treatment–eligible according to their occlusion site based on current American Heart Association/American Stroke Association guidelines.11

Technical Parameters

MR imaging examinations were performed on a Magnetom Skyra 3T, a Magnetom Skyra fit 3T, or a Magnetom Avanto fit 1.5T scanner (Siemens) and included diffusion-weighted (acquired at \( b = 1000 \text{ s/mm}^2 \)), fluid-attenuated inversion recovery, and susceptibility-weighted images. The Table provides a summary of the acquisition parameters.

Image Postprocessing

All diffusion-weighted images were coregistered to the standard Montreal Neurological Institute template (http://www.bic.mni.mcgill.ca/ServicesAtlases/ICBM152NLin2009) space rotated in the anterior/posterior commissure direction, using an affine transformation and a mutual information loss, and resampled to a standard resolution of 128 \( \times \) 128 \( \times \) 40 voxels with the Advanced Normalization Tools library (http://stnava.github.io/ANTS/). Images were then saved back in the DICOM format using a homemade C++ code.

Infarct Core Segmentation

Infarct cores, if present, were manually segmented by 2 radiologists (J.M.O., V.S.-Z.) on the coregistered \( b = 1000 \text{ s/mm}^2 \) DWI using an open-source DICOM viewing software (horosproject.org). Tissue was identified as infarct on the basis of hyperintensity compared with the surrounding tissue and the contralateral hemisphere. Imaging artifacts resembling infarct cores were distinguished from real infarct cores on the basis of contextual information from the whole image including symmetry/bilaterality, presence of metals, other causative materials, and typical anatomic location near the skull base or on brain-bone/air interfaces. Parenchymal hematomas in continuity with within infarcts were included in the segmentation. In case of doubt, to distinguish hyperintensity due to acute infarction from leukoaraiosis and chronic white matter changes, we reviewed the other available MR images, including the \( b = 0 \text{ s/mm}^2 \) images and fluid-attenuated inversion recovery images.

![Image](image_url)

**FIG 1.** Inaccuracy of volumetric measurements at the lesion surface. The grid represents voxel boundaries; the solid line represents an acute ischemic lesion. Voxels outlined in gray are the surface voxels, ie, those voxels with lesion margins. The sum of the surface voxel volume is a measure of the potential measurement error due to partial volume effects, which results from discretization at the lesion surface (the net volume uncertainty per surface voxel has been defined as half of the volume of the voxel in both the positive and negative directions). Standard resolution leads to large surface volume and thus has the potential for large measurement errors (A). Increasing the spatial resolution substantially reduces the surface volume and thereby decreases the potential measurement error (B).

**Diffusion-weighted imaging acquisition parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section thickness (mm)</td>
<td>3.03</td>
<td>0.22</td>
<td>3–5</td>
</tr>
<tr>
<td>TR (ms)</td>
<td>6989.76</td>
<td>1518.53</td>
<td>3550–9600</td>
</tr>
<tr>
<td>TE (ms)</td>
<td>99.59</td>
<td>3.32</td>
<td>59–104</td>
</tr>
<tr>
<td>Interslice gap (mm)</td>
<td>0.6</td>
<td>–</td>
<td>0.6–0.6</td>
</tr>
<tr>
<td>Sequence duration (sec)</td>
<td>583</td>
<td>168</td>
<td>312–984</td>
</tr>
<tr>
<td>Field strength</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5T: ( n = 251 ) (63.9%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3T: ( n = 142 ) (36.1%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matrix size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>128: ( n = 282 ) (71.8%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>162: ( n = 106 ) (27.0%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>192: ( n = 5 ) (1.3%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In-plane resolution</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.00 ( \times ) 2.00 mm: ( n = 318 ) (81.0%)</td>
<td>1.36 ( \times ) 1.36 mm: ( n = 46 ) (11.7%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.20 ( \times ) 1.20 mm: ( n = 4 ) (1.0%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.13 ( \times ) 1.13 mm: ( n = 1 ) (0.3%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.68 ( \times ) 0.68 mm: ( n = 24 ) (6.1%)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note:—— indicates that there was no SD as the interslice gap was 0.6 mm in all cases.
Lesion Tessellation
Section-by-section parameterization of segmented ischemic cores using spline interpolants was performed to reconstruct lesion boundaries with arbitrarily high in-plane resolution, and tessellated surfaces were reconstructed from the interpolated lesion boundary points, thereby providing lesion boundaries of varying in-plane resolutions. Reconstruction of the high-resolution 3D lesions with voxels of fixed size within the tessellated surface boundaries was then performed for a continuous range of voxel sizes from 4 × 4 × 8 to 1 × 1 × 2 mm using a Matlab routine (MathWorks). Section-by-section parameterization and lesion tessellation are illustrated in detail in the On-line Appendix.

Data Analysis and Statistics
Differences in the segmented stroke volumes between the 2 radiologists were visualized with a Bland-Altman plot. Infarct core volumes and infarct core surface volumes were calculated for both readers and were averaged for further analysis. Means and SDs were calculated for averaged stroke volume and surface volume, and the number of patients in whom the margin of error (ie, mean infarct volume ± 50% of the mean surface volume) contained 70 mL, which is the commonly used decision threshold for endovascular treatment, was calculated for each voxel size. A Spearman correlation was performed to assess the relationship of spatial resolution and surface volume. Correlation coefficients and corresponding P values were calculated for each patient and averaged.

RESULTS
Data Base Characteristics
MR imaging examinations of 393 patients with DWI-hyperintense lesions consistent with acute infarction were included in this study (220 male patients, 56.0%). The mean age was 71.6 ± 14.2 years. Fifty-six patients had an anterior circulation large-vessel occlusion (26 male patients, 46.4%; mean age, 71.7 ± 14.9 years; median NIHSS score, 8; interquartile range, 12–17.5), with a median time from symptom onset to MR imaging of 26 hours (interquartile range, 21.8–30.8 hours). These included 34 M1 segment middle cerebral artery occlusions, 21 terminal internal carotid artery occlusions, and 1 A1 segment anterior cerebral artery occlusion.

Human Segmentation
Comparison of segmentation volumes performed by both radiologists showed good agreement: In 380/393 (96.03%), the difference was within the limits of agreement (Fig 2). Mean infarct volume was 27.2 ± 49.9 mL, and 69.4 ± 81.1 mL in the subgroup of patients with anterior circulation LVO. Thirty-two of 393 (8.1%) infarct volumes were above 70 mL. In 8/393 (2.0%) infarct volumes, the readers disagreed on whether the infarct volume was greater or less than 70 mL. On-line Fig 3 shows how human segmentations were performed on an exemplary case.

Volume Error Related to Partial Volume Effects
The mean surface volume (ie, volume error related to partial volume effects) was 14.7 ± 20.8 mL for a reconstruction voxel size of 2 × 2 × 4 mm and 7.4 ± 10.7 mL for a size of 1 × 1 × 2 mm. In the patients with anterior circulation LVO, the mean surface volume was 64.8 ± 31.1 mL for a reconstruction voxel size of 2 × 2 × 4 mm and 17.7 ± 16.3 mL for a size of 1 × 1 × 2 mm. Figure 3 shows the simulation results for additional spatial resolution steps. There was a strong and highly significant correlation between spatial resolution and surface volume (P = 0.903, P < .001).

Volume Range Containing the Clinical-Decision Volume Threshold of 70 mL as Function of Reconstruction Resolution
With a reconstruction voxel size of 2 × 2 × 4 mm, the clinical-decision volume threshold of 70 mL was contained in a lesion volume ranging from 46.5 to 94.1 mL—that is, lesions as small as 46.5 mL could have been erroneously classified as >70 mL due to partial volume effects, and lesions of 94.1 mL volume could have been erroneously classified as <70 mL (Fig 4B). Reducing the reconstruction voxel size to 1 × 1 × 2 mm resulted to the clinical-decision volume threshold being contained in the range of 64.4–80.1 mL (Fig 4C).

Number of Patients with Anterior Circulation LVO within the Range of Measurement Error
In the patients with anterior circulation LVO, the clinical-decision volume threshold of 70 mL was contained within the range of measurement error in 7/56 cases (12.5%) with a resolution of 4 × 2 × 2 mm and in 3/56 cases (5.3%) with a resolution of 1 × 1 × 2 mm (Fig 5).

DISCUSSION
Accurate determination of infarct volume in acute ischemic stroke is crucial because infarct core volume was used as an exclusion criterion in several randomized controlled trials and has become a key decision driver in endovascular therapy decision-
making. Our study demonstrates the importance of spatial resolution when performing DWI-based infarct measurements in the context of acute stroke imaging: With a reconstruction resolution of \( \frac{2}{C^2} \times \frac{2}{C^2} \times 4 \) mm, the mean error in our patient cohort was 14.7 mL, and a resolution of \( \frac{1}{C^2} \times \frac{1}{C^2} \times 2 \) mm led to a noteworthy decrease of the error to 7.4 mL. The same held true for the number of cases in which 70 mL was contained within the margin of error in volume measurements (from 12.5% of cases with LVO using a \( \frac{2}{C^2} \times \frac{2}{C^2} \times 4 \) mm resolution to 5% of cases using a \( \frac{1}{C^2} \times \frac{1}{C^2} \times 2 \) mm resolution), while the range of lesion volumes measured in these cases decreased from 46.5–94.1 to 64.4–80.1 mL.

Our findings suggest that DWI in the context of acute stroke should be acquired with the maximal resolution possible. Of note, the maximal achievable resolution is variable across scanners and depends mostly on the imaging gradient strength. While the increase in resolution implies a decrease in the signal-to-noise ratio, this should be acceptable because of the high contrast of the DWI findings positive for lesions compared with the normal brain parenchyma. In case the signal-to-noise ratio needs to be improved, scan time could be increased. However, stroke outcomes are highly time-dependent, and treatment delays should, therefore, be avoided. If a high-resolution DWI sequence were to be used in the acute setting, its acquisition time has to be relatively short. High-resolution DWI sequences with acquisition times of <5 minutes are available and have been shown to increase lesion conspicuity and the overall number of visible lesions in acute ischemic stroke.13,14
Several further sources of errors that can occur in DWI-based infarct core measurement were not considered in this work, such as inhomogeneities in the $B_0$ and diffusion-weighted gradient and imaging gradient fields, errors in section selection, and motion artifacts. Because the effect of these errors is multiplicative, our results most probably substantially underestimated the real volume measurement error. Reducing measurement error due to partial volume averaging is even more important in patients with larger, complex-shaped infarcts (ie, gyriform, white matter–sparing infarct patterns) because the lesion surface and hence the potential error due to partial volume averaging are particularly large in such patients.

**Limitations**

Our study has several limitations: First, this was a single-center study. The images used were mostly acquired after thrombolytic or endovascular therapy, but it is reasonable to assume that the sizes and shapes of the lesions were similar to lesions in the therapeutic window. Infarct segmentation was performed manually, possibly introducing some inaccuracy, though interreader agreement of the infarct segmentations was good. We performed a “virtual increase” using lesion tessellation to simulate high-resolution DWI, and not “real” high-resolution DWI, which has been shown to be feasible at and beyond the smallest voxel size simulated in this study but yields its own set of challenges that have not been addressed in this study. These include a decrease in signal-to-noise ratio with voxel size and increased motion artifacts due to longer required scan times. Furthermore, endovascular therapy trials that used DWI for infarct core measurement and established the 70-mL threshold might have had the very same measurement inaccuracy described in this article but were nevertheless able to show the clinical significance of this threshold. DWI volume is used in most multicentric studies as a surrogate marker for stroke severity, for example in the Diffusion Weighted Imaging Evaluation for Understanding Stroke Evolution Study 2 (DEFUSE 2) studies, but in clinical practice, pretreatment MR imaging is performed only in a limited number of centers. Despite these limitations, our study provides preliminary exploratory evidence for the magnitude of infarct core measurement error on DWI that can arise due to partial volume averaging and offers a starting point for future investigations on how to reduce these inaccuracies.

**CONCLUSIONS**

When one performs DWI-based volumetric stroke infarct measurements with routinely used spatial resolutions, measurement errors due to partial volume averaging are substantial and can potentially affect clinical decision-making. Thus, infarct volumes derived from DWI should be interpreted with caution. The magnitude of error can be minimized when spatial resolution is increased. Because fast and robust high-resolution DWI sequences are available, they should be used in acute stroke imaging whenever possible to accurately measure infarct cores and come to a solid, well-informed endovascular therapy decision.


**REFERENCES**

Structural Connectivity and Cortical Thickness Alterations in Transient Global Amnesia


ABSTRACT

BACKGROUND AND PURPOSE: Transient global amnesia (TGA) is a sudden onset of anterograde and retrograde amnesia. We aimed to assess differences in terms of cortical thickness and structural brain connectome between patients with TGA (at acute and delayed postrecovery stages) and matched controls.

MATERIALS AND METHODS: We report on 18 consecutive patients with TGA who underwent 3T MR imaging, including DTI and MPRAGE sequences, at the acute (mean delay postonset: 44 hours) and delayed post-recovery (mean delay: 35 days) stages. Structural connectome was assessed in patients with TGA and in 18 age- and sex-matched controls by using probabilistic fiber-tracking and segmentation of 164 cortical/subcortical structures (“nodes”). Connectivity graphs were computed and global network metrics were calculated. Network-based statistical analysis (NBS) was applied to compare patients with TGA at each stage with controls. We also compared cortical thickness between patients with TGA and healthy controls.

RESULTS: Global network metrics were not altered in patients with TGA. NBS-analysis showed structural connectome alterations in patients with TGA compared with controls, in core regions involving the limbic network, with 113 nodes and 114 connections (33 left intrahemispheric, 31 right intrahemispheric, and 50 interhemispheric connections) showing significantly decreased structural connectivity ($P_{\text{NBS corrected}} < .05$, $t$-values ranging from 3.03 to 8.73). Lower cortical thickness compared with controls was associated with these structural alterations in patients with TGA, involving the orbitofrontal, cingulate, and inferior temporal cortices. All the abnormalities were visible at both acute and delayed postrecovery stages.

CONCLUSIONS: Our preliminary study suggests there are structural abnormalities of the limbic network in patients with TGA compared with controls, including decreased structural connectivity and cortical thickness.

ABBREVIATIONS: TGA = transient global amnesia; TOF-MRA = time-of-flight MRA; VBM = voxel-based morphometry; AUC = area under the curve; FDR = false discovery rate; NBS = network-based statistical analysis.
hypoperfusion of the parietal, posterior temporal, occipital, and cerebellar areas. More recently, 2 resting-state functional MR imaging studies revealed a decreased functional connectivity in TGA. Peer et al\textsuperscript{13} showed functional alterations involving the episodic memory network including the frontal regions, cingulate cortex, and basal ganglia. Zidda et al\textsuperscript{14} demonstrated alterations of the executive and salience networks including frontal and prefrontal regions, cingulate gyrus, parietal, insular and hippocampus.

However, the connectivity alterations at the acute stage of TGA, and their evolution at delayed postrecovery stage, have not been investigated yet by brain morphometry or whole-brain structural connectome. Considering the clinical phenomenology of TGA and previous neuroimaging studies, we hypothesized that the structural connectome could also be altered in patients with TGA, particularly within the limbic network that includes the hippocampus, temporal pole, parahippocampal gyrus, amygdala, anterior and posterior cingulate gyrus, and medial and lateral orbitofrontal cortex. To test this hypothesis, we aimed to assess differences in terms of cortical thickness and structural brain connectome between patients with TGA (at acute and delayed postrecovery stages) and matched controls.

**MATERIALS AND METHODS**

**Participants**

The Institutional Review Board approved this prospective study. Informed written consent was obtained according to ethical standard procedures. During a study inclusion of 2 years, the inclusion criteria were the following:

1) Right-handed patient.

2) Diagnosis of TGA according to Hodges and Warlow criteria\textsuperscript{1} including: a) presence of an anterograde amnesia that is witnessed by an observer, b) no clouding of consciousness or loss of personal identity, c) cognitive impairment limited to amnesia, d) no focal neurologic or epileptic signs, e) no recent history of head trauma or seizures, and f) resolution of symptoms within 24 hours.

3) No personal history of prior episodes of TGA, drug abuse, or neurologic or psychiatric disorders.

4) Neurologic examination considered as normal at acute and delayed postrecovery stages.

5) 3T MRI available at both acute (less than 48 hours after onset) and delayed (between 30 to 40 days after onset) stages.

Exclusion criteria included: age under 18 years, contraindication for 3T MR imaging, brain lesion or motion artifacts detected on MR images. We excluded patients and controls in case of 1) white matter hyperintensities with a score of 2 or 3 according to the Fazekas scale, 2) acute or old brain infarct(s), 3) more than 3 microbleeds, or 4) brain tumor, because these brain lesions may have potentially altered the tractography of the fibers. Hippocampal hyperintensity was not exclusionary for patients with TGA.

The recruitment of healthy controls was based on clinical examination (including the Mini-Mental State Examination), socio-demographic factors, and MR imaging data. The healthy volunteers were selected from a cohort of healthy subjects once the cohort of patients with TGA was fully established. Indeed, each control was selected to correspond to a patient with TGA in terms of sex and age. The exclusion criteria applied to patients with TGA and controls. Both patients with TGA and healthy volunteers had no personal history of prior episodes of TGA, drug abuse, neurologic or psychiatric disorders, and had normal structural MR imaging.

**MR Imaging Acquisition**

Patients with TGA and healthy controls were scanned at the same site by using 3T MR imaging (Discovery MR750; GE Healthcare) with the same MR sequences: 1) 3D T1-weighted MPRAGE: TR/TE, 2530/1.74 ms; flip angle, 70°; matrix, 256 x 256; isotropic 1-mm voxels; acquisition time, 7 minutes; 2) single-shot echo-planar diffusion tensor imaging: TR/TE, 8080/83 ms; flip angle, 90°; matrix, 128 x 128; isotropic 2-mm voxels; b = 10 s/mm\textsuperscript{2} images (“b0 images”) and 64 directional gradients at b = 1000 seconds/mm\textsuperscript{2}; acquisition time, 10 minutes. DTI trace images and apparent diffusion coefficient maps were systematically generated. In all patients and healthy controls, the MR imaging protocol also included axial T2-weighted, 3D FLAIR, time-of-flight MRA (TOF-MRA), and susceptibility-weighted MR images.

**Analysis of Conventional MR Images**

In each subject, 2 senior neuroradiologists (J.H. and X.L.) analyzed in consensus all the MR images available to ensure the lack of brain lesions, and in search of potential hippocampal lesions on DTI trace images and ADC maps.

**Image Pre-processing and Whole-Brain Tractography**

For each participant, MPRAGE images were segmented in 74 cortical regions and 8 subcortical regions per hemisphere by using FreeSurfer software (v5.3.0, http://surfer.nmr.mgh.harvard.edu/). Moreover, the cortical thickness map in native space was generated by FreeSurfer software. DTI data were preprocessed for eddy current distortions and motion artifacts by using FSL software (FMRIB Software Library; http://www.fmrib.ox.ac.uk/fsl/), and the MPRAGE image was non-linearly registered to the mean b0 image by using ANTS software (http://stnava.github.io/ANTS/). Whole-brain fiber-tracking was performed by using the MRtrix3 software package (http://www.mrtrix.org/). The algorithm generated 1,000,000 fibers of minimum length 20 mm. All steps are detailed in On-line Appendix and were manually checked for errors by visual inspection. In this study, image preprocessing and analysis were performed without knowledge of the presence of hippocampal hyperintensities on DTI trace images.

**Network Construction**

The nodes of graph G represented the 164 brain ROIs and edges, the brain subcortical and cortical structure connections. For any pair of nodes I and J, an edge was added to G if at least 1 fiber tract connected these 2 ROIs. Moreover, the edge was weighted by the sum of fiber tracts connecting these 2 ROIs divided by the volume of the 2 ROIs. Thus, for each subject, a 164 x 164 symmetric weighted graph was constructed.
Statistical Analyses

Demographic Data. Between-group difference in age and sex between controls and patients was assessed by student and $\chi^2$ tests, respectively.

Network Topology Analysis. The weighted graph $G$ was thresholded at different levels of sparsities that ranged from $T = 10\%$ to $35\%$ by using increments of $1\%$, in keeping the highest weighted edges. Then, further analysis was based on the binarized graph $G_T$ for each subject.

At each threshold, global network metrics were calculated by using GRETRA toolbox (https://www.nitrc.org/projects/gretna/): clustering coefficient ($C_p$), shortest path length ($L_p$), global efficiency ($E_g$), and local efficiency ($E_{loc}$). The metrics were described by Rubinov et al. $^{15}$ The area under the curve (AUC) for each network metric was calculated to provide a summarized scalar for topologic organization of brain networks independent of a single threshold selection. $^{16}$ Non-parametric permutation tests (20,000 permutations) were used to test between-group differences in AUC values of network metrics (ie, clustering coefficient, mean path length, local efficiency, and global efficiency). Patients with TGA between acute and delayed stages ($n = 18$) and with age-matched controls ($n = 18$) were compared by using paired and 2-sample $t$-tests, respectively. The results are presented at a statistical threshold of $P < .05$ (false discovery rate [FDR]-corrected).

Structural Connectivity Analysis. To further identify altered structural connections in patients with TGA between acute and delayed stages compared with controls, network-based statistical analysis (NBS) was used on the unthresholded corrected graph, which is a validated nonparametric statistical approach for controlling family-wise error in connectome analyses. $^{17}$ NBS was applied to compare patients and controls by paired and 2-sample $t$-tests as follows: $^7$ mean connectivity strength changes were calculated with permutation tests, $^2$ network components of interconnected edges that survived a $p$-value of .005 uncorrected were retained, $^3$ and the size of the largest cluster was calculated. To generate an empirical null distribution for evaluating the statistical significance of the cluster sizes, groups were randomly shuffled (20,000 permutations), then the largest cluster size null distribution was obtained by repeating steps 1, 2, and 3. The significance level of altered components in the NBS analysis was set to .05 (NBS-corrected).

Cortical Thickness and Subcortical Volumetric Analyses

Because our purpose was to assess the between-group differences in cortical thickness and subcortical volume within brain regions associated with structural connectivity alterations, an ROI analysis was also performed. The mean cortical thickness of cortical ROI and volumetry of subcortical ROI were measured by FreeSurfer software and compared between patients with TGA at acute and delayed stages and controls to assess if connectivity alterations were associated with cortical thickness or subcortical volumetry differences between groups. Nonparametric permutation tests (20,000 permutations) were used to test between-group differences. The results are presented at a statistical threshold of $P < .05$ (FDR-corrected).

RESULTS

Participants

During the 2-year inclusion period, 24 consecutive patients were diagnosed with TGA in our institution. Of these 24 patients, 4 presented motion artifacts on MR imaging and MR imaging was not available at the delayed postrecovery stage in 2. Finally, 18 patients with TGA (14 womens, mean age 64.5 years, range: 40–80 years, all right-handed) were included in the study. All patients ($n = 18$) underwent 3T MR imaging (Discovery MR750) at both acute (mean scan time from onset: 44 hours, range: 24–51 hours) and delayed postrecovery (mean scan time from onset: 35 days, range: 31–39 days) TGA stages. Characteristics of patients with TGA are shown in the Table.

In addition, 18 age and sex-matched, right-handed, healthy volunteers served as controls. The controls perfectly matched with patients in terms of sex (18 subjects including 14 [78%] females) and were close in terms of age (the mean age of patients was 64 ± 11 years; it was 61 ± 13 years in controls). The comparison between patients with TGA and controls did not show any significant statistical difference for age ($P = .46$) or sex ($P = .1$).

Hippocampal Hyperintensities on DTI Trace Images

Punctate hyperintensities were visible on DTI trace images in 12 of the 18 patients with TGA. Results are detailed in the Table. There were no abnormalities detected in controls. Postprocessing data showed no correlation between hippocampal hyperintensities and structural connectivity ($P = .38$, Mann-Whitney test) or volume of hippocampus ($P = .61$, Mann-Whitney test).

Network Topology

No differences of the global network metrics were observed between patients with TGA at acute and delayed stages and healthy controls (Fig 1).

Structural Connectivity

NBS Comparing patients with TGA at Acute Phase and Healthy Controls. In comparison with controls, a cluster of 113 nodes and 114 connections (33 left intrahemispheric, 31 right intrahemispheric, and 50 interhemispheric connections) had significantly decreased structural connectivity (ie, decreased number of fiber tracts) ($P < .05$ NBS corrected, $t$ values ranging from 3.03 to 8.73) in patients with TGA at acute phase (Fig 2A and On-line Table).

In this widespread network, structural changes in patients with TGA were more prominent in core regions (ie, showing the highest number of altered connections) including: left hippocampus, left frontomarginal gyrus and sulcus, left middle anterior part of the cingulate gyrus and sulcus, left subcallosal gyrus, left planum polare of the superior temporal gyrus, left superior segment of the circular sulcus of the insula, left superior and transverse occipital sulci, left and right opercular part of the inferior frontal gyri, right inferior temporal gyrus, right anterior segment of the circular sulcus of the insula, right superior frontal sulcus, right parieto-occipital sulcus, and right pericallosal sulcus (Fig 2B).
NBS Comparing Patients with TGA at Acute and Delayed Stages. No statistical difference in structural connectivity was observed between acute and delayed stages in patients with TGA.

Cortical Thickness and Subcortical Volumetry
In comparison with controls, lower cortical thickness was observed in patients with TGA at the acute stage in the following cortical regions: left and right frontomarginal gyrus and sulcus, left orbital gyr, left transverse frontopolar gyrus and sulci, left middle-anterior part of the cingulate gyrus and sulcus, left lateral orbital sulcus, left posterior-ventral cingulate gyrus, left subcallosal gyrus, left lateral occipito-temporal gyrus, right long insular gyrus and central sulcus of the insula, right triangular part of the inferior frontal gyrus, and right opercular part of the inferior frontal gyrus (Online Figs 1 and 2). No between-group difference was observed in subcortical regions.

Longitudinal Follow-Up at the Delayed Post-recovery TGA Stage
All the structural abnormalities observed in patients with TGA (ie, decreased structural connectivity and decreased cortical thickness and subcortical volumetry) remained completely stable at the delayed stage, in each patient with TGA.

DISCUSSION
This is the first study evaluating NBS-corrected structural connectome and cortical thickness in TGA at both acute and postrecovery stages. We found lower structural connectivity and cortical thickness in some core regions of the limbic network in patients with TGA. This preliminary study highlights the potential value of advanced MR imaging to improve our understanding of TGA. Our findings and those of the other available studies may be of use to determine the most relevant neuropsychological tests in patients with TGA, potentially reducing the time and cost of future research protocols.
Of note the topologic organization of the structural connectome was not altered in patients with TGA. Rather than a global disorganization, our study suggests altered connectivity of a specific network in patients with TGA. Such findings are in agreement with neuropsychological studies demonstrating a selective disorder of episodic memory in TGA without impairment of other components of memory. Indeed, as suggested by previous SPECT, PET, tract-based spatial statistics, and fMRI studies, the present study leads to the concept of TGA as a network disease.

We found structural alterations in several brain regions that are all known to be involved in different stages or components of memory processing, including: hippocampus, temporal lobe, orbitofrontal cortex, and parieto-occipital sulcus or cingulum. Indeed, most of these regions are directly involved in memory processing, in strategic or executive-based processing of mnemonic information meant to ensure its successful encoding or retrieval. The orbitofrontal and dorsal anterior cingulate cortices are involved in value-based modulations and temporal context-retrieval. The ventral frontal regions tap into incidental and intentional memory encoding. The retrosplenial cortex is also assumed to perform the translation between egocentric and allocentric representations formed into the hippocampus, mostly during navigation.

Interestingly, some brain regions with decreased connectivity are also involved in meta-cognitive processing. The dorsal anterior cingulate regions, together with the subcallosal cortex and the insula are all involved in mnemonic monitoring processes such as self-assessment of encoding or retrieval and confidence in self-performance. This network appears critical to ensure successful memory given its connections to frontal and medial temporal regions, parahippocampal cingulate bundle and fornix. Overall, this network allows for strong interactions between memory processing and meta-cognitive processes so that associative regions could interact with interoceptive cortex and salience regions to promote accurate judgment of self-memory performance as well as self-projection in different spatial and temporal dimensions.

To compare the cortical thickness between patients and controls we performed an ROI analysis, which is potentially more sensitive than the voxel-based approach previously used in patients with TGA. Decreased cortical thickness involved the orbitofrontal, cingulate, and inferior temporal cortices of patients with TGA. Interestingly, several regions of the limbic network showed alterations of both structural connectivity and cortical thickness (left subcallosal gyrus, left lateral orbital sulcus, left middle anterior cingulate cortex, and right opercular inferior frontal gyrus), further strengthening our results.

Our study has some limitations. There was no patient scanned during the hyperacute phase. Because TGA is both rare and short-lived, neuroimaging studies performed during an attack of amnesia are also rare. In this pilot study, the reference standard for the diagnosis of TGA was based on the Hodges and Warlow criteria that are widely used in the clinical practice and completely recovered after the event. Further studies are required to correlate the structural abnormalities with neuropsychological assessments of cognitive and emotional functions.

CONCLUSIONS

Compared with controls, both structural connectivity and cortical thickness were significantly decreased in the limbic network in patients with TGA. This preliminary study highlights the potential value of advanced MR imaging to improve our understanding of TGA and may be of use to plan further prospective studies.

REFERENCES


FIG 2. Altered structural connections in patients with TGA. A. Structural connectivity differences between controls and patients with TGA at acute phase. The gray lines mean significant decrease of connectivity in patients compared with controls. The red dots represent the location of ROIs and their size is proportional to the number of significant connections. A network of 113 ROIs and 114 connections had significantly decreased structural connectivity in patients with TGA (P < .05 NBS corrected, t-values ranging from 3.03–8.73). B. In this widespread network, the cortical and subcortical regions (nodes) with a number of significantly altered connections between patients and controls greater than 4 were mainly located in inferior and mesial temporal, cingulate, and frontal cortices.
15. Rubinov M, Sporns O. 
1. Is transient global amnesia a network disease? Eur Neurol 2018;80:345–54 CrossRef

The Impact of Intracortical Lesions on Volumes of Subcortical Structures in Multiple Sclerosis

I. Kalinin, G. Makshakov, and E. Evdoshenko

ABSTRACT

BACKGROUND AND PURPOSE: Recent studies showed thalamic atrophy in the early stages of MS. We investigated the impact of intracortical lesions on the volumes of subcortical structures (especially the thalamus) compared with other lesions in MS.

MATERIALS AND METHODS: Seventy-one patients with MS were included. The volumes of intracortical lesions and white matter lesions were identified on double inversion recovery and FLAIR, respectively, by using 3D Slicer. Volumes of white matter T1 hypointensities and subcortical gray matter, thalamus, caudate, putamen, and pallidum volumes were calculated using FreeSurfer. Age, MS duration, and the Expanded Disability Status Scale score were assessed.

RESULTS: Patients with intracortical lesions were older ($P = .003$), had longer disease duration ($P < .001$), and higher Expanded Disability Status Scale scores ($P = .02$). The presence of intracortical lesions was associated with a significant decrease of subcortical gray matter volume ($P = .02$). In our multiple regression model, intracortical lesion volume was the only predictor of thalamic volume ($R^2 = 0.4$, $b^2 = -0.28$, $P = .03$) independent of white matter lesion volume and T1 hypointensity volume. White matter lesion volume showed an impact on subcortical gray matter volume in patients with relapsing-remitting MS ($P = .04$) and those with disease duration of <5 years ($P = .04$) and on thalamic volume in patients with Expanded Disability Status Scale scores of $<4.0$ ($P = .01$). By contrast, intracortical lesion volume showed an impact on subcortical gray matter and thalamic volumes in the secondary-progressive MS subgroup ($P = .02$ and $P < .001$) in patients with a long-standing disease course ($P < .001$ and $P = .001$) and more profound disability ($P < .001$ and $P < .001$).

CONCLUSIONS: Thalamic atrophy was explained better by intracortical lesions than by white matter lesion and T1 hypointensity volumes, especially in patients with more profound disability.

ABBREVIATIONS: DIR = double inversion recovery; EDSS = Expanded Disability Status Scale; ICL = intracortical lesion; ICV = intracranial volume; PPMS = primary-progressive MS; RRMS = relapsing-remitting MS; SPMS = secondary-progressive MS; T1H = T1 hypointensities; WMH = white matter hypointensities; WML = white matter lesions; SGMS = subcortical grey matter structures; $b^2$ = standardized regression coefficients

Multiple sclerosis (MS) is a chronic debilitating disease of the CNS. Atrophy of subcortical gray matter structures (SGMS) is associated with cognitive dysfunction, fatigue, and overall disability in MS.1 Among all SGMS, most attention is riveted on the thalamus, a gateway for cortical areas and a relay between the cortex and basal ganglia.2 Volumetric MR imaging studies showed thalamic atrophy at early stages of the disease, including clinically and radiologically isolated syndromes.3,4 Thalamic atrophy was associated with fatigue,5,6 cognitive impairments,7,8 decrease of walking speed,9 and higher disability.10,11 Measuring thalamic volume may be a good surrogate biomarker for clinical trials and, probably, routine practice.4,11,12

Thalamic atrophy in MS can be related to either primary damage in the thalamus or distant damage of afferent or efferent thalamic fibers.1 Primary damage includes inflammation and neurodegeneration followed by demyelination, neuronal loss, neuronal shrinkage, and axonal damage.7,13 Distant damage occurs due to mechanisms of Wallerian or transneuronal degeneration as a result of gray matter (GM)14 and white matter (WM)15 pathology.

In early MS, most lesions are located in the WM with less prominent involvement of the GM.16 Cortical demyelination spreads as the disease progresses, and it is most prominent in the secondary-progressive phase.17 However, thalamic atrophy is prominent in early stages and continues throughout the disease course.11 The
The differential impact of localization of demyelinating lesions on the thalamus and SGMS atrophy is unclear. In a neuropathological study of Vercellino et al., demyelination of subcortical GM correlates well with cortical, but not WM, demyelination. By contrast, in the MR imaging study by Henry et al., in patients with clinically isolated syndrome, the evidence of a direct connection between thalamic atrophy and white matter lesions (WML) was shown. In the MR imaging study of Ruggieri et al., atrophy of SGMS, eg, the thalamus, correlated with the number and volume of cortical lesions. The negative correlation of T1 black hole load and SGMS atrophy in patients with relapsing-remitting MS (RRMS) was shown.

To date, not enough data demonstrate the differentiated impact of cortical and WM pathology on subcortical gray matter atrophy. In this study, we aimed to investigate the impact of cortical lesions on volumes of the thalamus and other subcortical structures compared with other lesion locations (eg, WML and T1 white matter hypointensities [WMH] in MS.

**MATERIALS AND METHODS**

**Study Population**

This study had a cross-sectional design with retrospective analysis of data. Seventy-one patients with MS were included. The study was approved by the Local Ethics Committee of the SBIH City Clinical Hospital No. 31, and all subjects provided their written informed consent.

The inclusion criteria were as follows: patients of both sexes with multiple sclerosis, defined with the McDonald 2017 criteria; 18–65 years of age at MR imaging; and providing informed consent. The exclusion criteria were as follows: patients with relapses or steroid use within 30 days before MR imaging; a medical history associated with brain pathology (excluding MS); and pregnancy.

All data were collected for the time of MR imaging study and included sex, age, disease duration (time from the first relapse till the date of MR imaging), and Expanded Disability Status Scale (EDSS) score.

**MR Imaging Acquisition Protocol**

MR imaging was performed on General Electric (GE) Signa (GE Healthcare) 3T machine with isometric voxels (1 × 1 × 1 mm); model discovery, 750w; receive channels, 24. MR imaging sequence protocol was as follows—T1 CUBE (GE Healthcare): imaging plane, sagittal; imaging resolution, 1 × 1 × 1 mm; TR, 7.68 ms; TE, 3.12 ms; TI, 400 ms; flip angle, 11°; echo-train length, 1; bandwidth, 244 Hz/pixel; FLAIR CUBE: imaging plane, sagittal; imaging resolution, 1 × 1 × 1 mm; TR, 4750 ms; TE, 100 ms; flip angle, 90°; echo-train length, 10; bandwidth, 122 Hz/pixel; and double inversion recovery (DIR) CUBE: imaging plane, sagittal; imaging resolution, 1 × 1 × 1 mm; TR, 3500 ms; TE, 140 ms; flip angle, 110°; echo-train length, 10; bandwidth, 140 Hz/pixel.

For this project, we defined only intracortical lesions (iCLs) that were confined within the cortex, with sparing of the white matter and detected using the DIR sequence. WML were detected on FLAIR. Lesion volumes were manually counted using 3D Slicer (http://www.slicer.org) by 2 independent neuroscientists with experience in lesion segmentation (G.M., I.K.) based on predefined criteria. All inconsistencies were solved by agreement.

Volumes of white matter hypointensities, detected on T1 (T1H); subcortical gray matter volume; and thalamus, caudate, putamen, and pallidum volumes were measured automatically using FreeSurfer (https://surfer.nmr.mgh.harvard.edu/). To normalize lesion volumes for brain size, we used ROBEX (https://www.nitrc.org/projects/robox) to count the intracranial volume (ICV) and then transformed all volumes in milliliters to percentage of ICV using the following formula:

\[
\text{Volume} = \frac{\text{ICV} × 100\%}{\text{Volume}^{\text{ICV}}}
\]

**Statistical Analysis**

All datasets were checked for normality with the Kolmogorov-Smirnov test.

On the first step, we divided all patients in 2 groups according to their intracortical lesion status, positive and negative for iCLs to compare demographic, clinical, and radiologic features of both groups depending on the iCL status. A comparison between iCL-negative and iCL-positive groups was performed with nonparametric (Mann-Whitney U test) tests. To compare the categorical data (the sex and MS phenotype), we used the Pearson χ² test. We considered P < .05 as significant. Data are presented as mean (SD) or as median (25th, 75th quartile).

On the second step, we attempted to assess the impact of iCL on volumes of subcortical structures using multivariate analysis of covariance, corrected for age and sex. Then we assessed the impact of iCL on volumes of subcortical structures compared with white matter lesions and T1 WMH using multiple regression analysis. Furthermore, using regression analysis, we estimated the impact of iCL in different disease categories, such as the following: 1) patients with relapsing-remitting MS (RRMS), secondary-progressive MS (SPMS), and primary-progressive MS (PPMS); 2) disease duration before and after 5 years; and 3) EDSS score <4.0 and ≥4.0. Volumes of iCL, WML, WMH, age, and disease duration were included as covariates.

**RESULTS**

Forty-three (61%) women and 28 (39%) men were included in the study. The median age was 37 years (interquartile range (difference between 75th and 25th percentiles), 29–49 years). According to the disease phenotype, there were 54/71 (76%) patients with RRMS, 12/71 (17%) with SPMS, and 5/71 (7%) with PPMS. The median EDSS score and disease duration were 3 months (interquartile range, 2–4.5 months) and 53 months (interquartile range, 15–122) months, respectively.

iCLs were detected in 55/71 (77%) patients. An example of a purely iCL is shown in Fig 1. On the basis of iCL status, patients were divided into groups positive and negative for iCL. Clinical and imaging data were compared between these 2 groups. No difference in the disease phenotype (RRMS, SPMS, PPMS) was detected between groups (P = .11). Patients with iCLs were older (P = .003) and had longer disease duration (P < .001) and the highest EDSS scores (P = .02). Results are shown in the Table 1.
The presence of iCLs was associated with a significant decrease of subcortical gray matter volume \((P = .02)\) and atrophy of the caudate, putamen, and pallidum \((P = .03, P = .02, P = .005\), respectively). Results are presented in Table 2 and Fig 2; examples of the volume measurements of the thalamus in patients positive and negative for iCL are shown in the Fig 3.

Generalized linear model analysis of covariance revealed a significant impact of the volume of iCLs (% ICV) on the volumes of subcortical gray matter \((F = 20.8, P < .001)\), thalamus \((F = 16.15, P < .001)\), caudate \((F = 7.4, P = .008)\), putamen \((F = 7.15, P = .009)\), and pallidum \((F = 8.88, P = .004)\). Age significantly predicted the total volumes of subcortical gray matter \((F = 7.76, P = .007)\) and the putamen \((F = 24.7, P < .001)\). Sex was not a significant predictor in this model.

Our multiple regression model explained 43% of variance \((R^2 = 0.43)\), and iCL volume was the only predictor of the thalamic volume \((b^* = -0.28, P = .03)\) independent of the volume of WML and T1H. iCL volume did not demonstrate an impact on volumes of other structures. T1H volume predicted subcortical gray matter volume \((R^2 = 0.52, b^* = -0.41, P = .02)\) independent of the volumes of iCLs and WML.

### Subgroup Analysis

**Disease Phenotype.** In patients with RRMS, only the WML volume had a significant impact on subcortical gray matter volume \((R^2 = 0.49, b^* = -0.49, P = .04)\) and putamen volume \((R^2 = 0.27, b^* = -0.57, P = .04)\).

By contrast, in the SPMS subgroup, iCL lesion volume showed a significant impact on subcortical gray matter volume \((R^2 = 0.81, b^* = -0.44, P = .02)\) and thalamic volume \((R^2 = 0.89, b^* = -0.78, P < .001)\). Volumes of T1H explained only subcortical gray matter volume \((R^2 = 0.81, b^* = -1.11, P = .04)\). Volumes of WML showed no significant associations.

We did not perform multiple regression for the PPMS subgroup due to the low number of patients (4 subjects).

**Disease Duration.** In the iCL-positive group, 20 patients had disease duration of <5 years and 35 patients had disease duration of >5 years.

In patients with MS with disease duration of <5 years, only the volume of WML had a significant impact on subcortical gray matter volume \((R^2 = 0.34, b^* = -0.70, P = .04)\), putamen volume \((R^2 = 0.36, b^* = -0.79, P = .02)\), and pallidum volume \((R^2 = 0.32, b^* = -0.77, P = .03)\).

In patients with long-standing disease course (>5 years), we showed significant impact of iCL volume on subcortical gray matter volume \((R^2 = 0.71, b^* = -0.41, P < .001)\), thalamus \((R^2 = 0.55, b^* = -0.50, P = .001)\), and caudate volumes \((R^2 = 0.49, b^* = -0.59, P = .01)\). The volume of T1H was associated with subcortical gray matter volume \((R^2 = 0.71, b^* = -0.62, P = .001)\). Volumes of WML showed no significant associations.

---

**Table 1: Demographic characteristics according to iCL status**

<table>
<thead>
<tr>
<th>iCL-Negative Subgroup ((n = 16))</th>
<th>iCL-Positive Subgroup ((n = 55))</th>
<th>(P) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. (% women) 10/16 (63) 33/55 (63)</td>
<td>8.6*</td>
<td></td>
</tr>
<tr>
<td>Median age at MR imaging (yr) (25th, 75th percentile) 27 (25–36.5) 40 (32–53)</td>
<td>0.003b</td>
<td></td>
</tr>
<tr>
<td>Median disease duration (mo) (25th, 75th percentile) 15 (9–34.5) 77 (26–77)</td>
<td>&lt;.000b</td>
<td></td>
</tr>
<tr>
<td>Median EDSS score (25th, 75th percentile) 2 (1.25–3) 3.5 (2–4.5)</td>
<td>.02b</td>
<td></td>
</tr>
<tr>
<td>MS phenotype (No.) (%) Relapsing-remitting 15 (94) 39 (71)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary-progressive 0 12 (22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary-progressive 1 (6) 4 (7)</td>
<td>.11a</td>
<td></td>
</tr>
</tbody>
</table>

*To compare, the Pearson \(\chi^2\) test was used.

bSignificant difference.

---

**Table 2: Brain morphometry analysis according to the iCL status**

<table>
<thead>
<tr>
<th>Structures, Median, and (25th, 75th Percentile)</th>
<th>iCL-Negative Subgroup ((n = 16))</th>
<th>iCL-Positive Subgroup ((n = 55))</th>
<th>(P) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median of ICV (mL) (25th, 75th percentile) 1374.49 (1282.47–1482.68) 1343.47 (1260.50–1342.01)</td>
<td>.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subcortical gray matter volume (median) % ICV (25th, 75th percentile) 4.36 (3.93–4.42) 4.01 (3.75–4.23)</td>
<td>.02a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thalamus volume (median) % ICV (25th, 75th percentile) 1.03 (0.99–1.097) 0.99 (0.93–1.09)</td>
<td>.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caudate volume (median) % ICV (25th, 75th percentile) 0.56 (0.48–0.598) 0.49 (0.46–0.55)</td>
<td>.03a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Putamen volume (median) % ICV (25th, 75th percentile) 0.81 (0.75–0.88) 0.74 (0.68–0.82)</td>
<td>.02a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pallidum volume (median) % ICV (25th, 75th percentile) 0.23 (0.20–0.25) 0.20 (0.17–0.22)</td>
<td>.005a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volume of WML (median) % ICV (25th, 75th percentile) 0.39 (0.16–0.57) 0.65 (0.31–1.34)</td>
<td>.01a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>iCL lesions volume (median) % ICV (25th, 75th percentile) – 0.008 (0.004–0.02)</td>
<td>–</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant difference.

---

806 Kalinin May 2020 www.ajnr.org

**FIG 1.** An example of a purely intracortical lesion (arrowhead) \(A\) on the DIR MR image and its manual segmentation using 3D Slicer \(B\).
EDSS Score. In the iCL-positive group, 28 patients had EDSS scores of <4.0, and 27 patients had EDSS scores ≥4.0.

In the subgroup of patients with MS with EDSS scores of <4.0, the only significant impact was shown for the volume of WML on thalamic volume ($R^2 = 0.34, b^* = -0.60, P = .01$).

In patients with EDSS score ≥4.0, iCL lesion volume showed a significant impact on subcortical gray matter volume ($R^2 = 0.77, b^* = -0.56, P < .001$), thalamic volume ($R^2 = 0.55, b^* = -0.60, P < .001$), and caudate volume ($R^2 = 0.57, b^* = -0.44, P = .007$).

The volume of T1H was associated with subcortical gray matter ($R^2 = 0.77, b^* = -0.79, P = .002$) and thalamic volumes ($R^2 = 0.55, b^* = -0.74, P = .03$). Volumes of WML showed no significant associations.

All data can be found in the Online Table.

DISCUSSION

In this study, we assessed the impact of iCL on volumes of subcortical structures compared with WML and T1H in MS.

Our findings demonstrate that the volume of iCLs may impact the volumes of subcortical gray matter structures, especially on the thalamus in MS. iCL volume was the only predictor of thalamic volume independent of volumes of WML and T1H.

The more interesting results were found in subgroup analysis. Volumes of WML impacted subcortical gray matter volume in patients with RRMS, patients with disease duration of <5 years, and the volume of the thalamus in patients with EDSS scores of <4.0. These findings are in line with those of the study of Tao et al, which has shown a correlation between the volume of WML and thalamic atrophy in patients with RRMS. Acute inflammation, related to the formation of WML, is the major driving force during the RRMS phase, with some diminishing inflammatory activity in the progressive phase. Hence, our findings are highly relevant to this concept. iCLs have a different pathogenesis and may prevail in progressive forms of MS, becoming more important to atrophy of subcortical structures.

iCL volume impacted subcortical gray matter and thalamic volume in patients with long-standing disease course and more profound disability. Although we found that iCL volume impacts subcortical gray matter and thalamic volumes in the SPMS subgroup, we had no iCL-negative patients with SPMS for comparison. In fact, the presence of cortical demyelination has been shown in pathology studies, however, MR imaging methods like the DIR sequence may not be sensitive enough to detect iCLs. Cortical pathology in our iCL-negative subgroup may not have been prominent enough to be detected with the DIR sequence, so in fact, this study may represent only a part of disease mechanisms due to technical limitations.

Because in the early stages of disease, white matter pathology predominates over cortical pathology, we suppose, that in this case, thalamic atrophy can be caused by distant damage of afferent or efferent thalamic fibers. This supposition is in line with previous MR imaging studies that supposed that atrophy of the
subcortical gray matter structures preceded cortical atrophy in early MS.19,26 Rocca et al26 observed that white matter lesions contribute to thalamic tissue loss. In another study, Henry et al19 demonstrated that only WML in the region of thalamocortical fibers significantly correlated with thalamic atrophy in patients with clinically isolated syndrome.

Cortical demyelination becomes much more prominent in the chronic disease stages.17 iCLs can result in thalamic atrophy due to retrograde degeneration of thalamocortical radiations.14 This finding is in accordance with previously published neuropathologic and MR imaging that showed a correlation between cortical demyelination and cortical lesions with thalamic atrophy.14,18

Our study had some limitations. We had a rather small sample size (71 patients in total and 55 with iCLs). The design of the study was not prospective, so the dynamics of changes were not evaluated. A low number of patients with PPMS (4 subjects) made it impossible to perform multiple regression for this subgroup. In addition, we had no healthy controls or a control group of patients with other neurodegenerative disorders with pathology of the basal ganglia.

CONCLUSIONS

Thalamic atrophy was explained better by intracortical lesions, than by white matter lesions and T1 white matter hypointensities, especially in patients with disease duration of >5 years and EDSS scores of ≥4.0.


REFERENCES

ABSTRACT

BACKGROUND AND PURPOSE: Though CT is a highly calibrated imaging modality, head CT is typically interpreted qualitatively. Our aim was to initiate the establishment of a reference quantitative database for clinical head CT.

MATERIALS AND METHODS: An automated segmentation algorithm was developed and applied to 354 clinical head CT scans with radiographically normal findings (ages, 18–101 years; 203 women) to measure brain volume, brain parenchymal fraction, brain radiodensity, and brain parenchymal radiomass. Brain parenchymal fraction was modeled using quantile regression analysis.

RESULTS: Brain parenchymal fraction is highly correlated with age ($R^2 = 0.908$ for men and 0.950 for women), with 11% overall brain volume loss in the adult life span (1%/year from 20 to 50 years and 2%/year after 50 years of age). Third-order polynomial quantile regression curves for brain parenchymal fraction were rationalized and statistically validated. Total brain parenchymal radiodensity shows a decline as a function of age (14.9% for men, 14.7% for women; slopes not significantly different, $P = .760$). Age-related loss of brain radiomass (the product of volume and radiodensity) is approximately 20% for both sexes, significantly greater than the loss of brain volume ($P < .001$).

CONCLUSIONS: An automated segmentation algorithm has been developed and applied to clinical head CT images to initiate the development of a reference database for quantitative brain CT imaging. Such a database can be subject to quantile regression analysis to stratify patient brain CT scans by metrics such as brain parenchymal fraction, radiodensity, and radiomass, to aid in the identification of statistical outliers and lend quantitative assessment to image interpretation.

ABBREVIATIONS: BPF = brain parenchymal fraction; TIV = total intracranial volume; SD = standard deviation

Abnormalities in brain volumetrics have been associated with congenital and acquired diseases. Most in vivo studies have been performed with MR imaging of healthy volunteers and measure global and regional volume loss. Lack of an accepted normative database, together with evidence that measurements are influenced by differences in postprocessing methods, has limited quantitative reporting. Among adults, MR imaging has been used to identify abnormalities of global brain volume in multiple sclerosis, amyotrophic lateral sclerosis, and age-related dementia, with brain parenchymal fraction (BPF) permitting normalization for subject variability. MR imaging also suggests that brain volume is reduced by antipsychotic medications, steroids, alcohol use, and radiation and chemotherapy among other things. Changes in brain radiodensity or radiomass (the product of volume and radiodensity), as a function of disease states are relatively unexplored topics. Identifying and quantifying tissue loss through volumetric measures, measures of radiodensity or radiomass, may aid in the diagnosis or monitoring of brain pathology. To account for intrasubject variability as well as variability as a function of age and sex, correlating metrics with pathology requires a reference database. The initiation of such a reference database is the goal of the current study.

CT imaging appears highly suitable for in vivo study of the brain because it is routinely acquired in the clinical setting and is less subject to motion artifacts than MR imaging. Radiodensity characteristics of the brain and skull enable automated volumetric and radiodensity assessment. Furthermore, the relatively low...
clinical threshold for performing CT provides large numbers of studies with radiographically normal findings, enabling the generation of a large reference database for statistical analysis.

In the current study, we report brain volumes from 354 subjects, both before and after normalization to the intracranial volume, expressed as a function of age and sex. Total radiodensity and radiomass estimations are also calculated. Statistical methods are used with quantile regression applied to brain parenchymal fraction measures. We propose that a clinical database can be used to quantitatively assess new cases in the context of a clinical peer group.

MATERIALS AND METHODS

Study Design
This study was limited to a retrospective analysis of head CTs performed on patients who were identified from the clinical PACS. The study was approved by this institutional review board (Geisinger Medical Center), and a waiver of consent was granted.

Study Cohort
All studies were from a single CT scanner during a 2-year time interval (January 1, 2015, to December 31, 2016). Selected cases were scanned for nonspecific symptoms (headache, syncope, vertigo), were without known systemic disease, and were discharged without incident. All control cases were interpreted as having normal findings (without acute or chronic abnormal findings) by 2 board-certified neuroradiologists.

Imaging Data
The CT scanner (LightSpeed VCT; GE Healthcare) primarily serves the emergency department of a level 1 trauma center. The axial acquisition noncontrast head CT protocol consists of 135 kV(peak) and modulated milliamperere, minimum 50 and maximum 290 mA; rotation time, 0.75 seconds, acquired from the foramen magnum through the vertex with a standard 512 x 512 matrix; and 24-cm FOV at 5.0-mm section thickness. The scanner undergoes a daily quality assurance procedure, which assesses the radiodensity of water. This value must be within allowable limits, generally 0–5 HU. Drift or trending is rarely observed. In addition, scanners undergo an annual inspection by a medical physicist using the American College of Radiology phantom. Acceptable ranges of Hounsfield units for clinical scanners are broad (−7 to +7 HU for water, 110–135 HU for acrylic). This testing is extended to all kV (peak) values used by the scanner. Additionally, service engineers routinely test the calibration at the scanner. DICOM images were converted to the Neuroimaging Informatics Technology Initiative data format using MRICconvert-2.0.7 (https://www.softpedia.com/get/Science-CAD/MRIConvert.shtml). Images were first thresholded from −15 to 50 HU to grossly remove background and skull. Brain extraction was then applied using the FSL Brain Extraction Tool (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET) with a fractional intensity threshold of 0.01. All cases were carefully reviewed for the integrity of brain extraction. For segmentation, a 3-tissue-compartment segmentation using the FMRIB Automated Segmentation Tool (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/fast) was used, with the resulting white matter and gray matter compartments combined into a single brain compartment (Fig 1). The brain parenchymal fraction (BPF) was calculated as the ratio of brain volume to total intracranial volume, consisting of brain plus CSF space. Brain radiomass was calculated as the product of mean brain radiodensity and brain volume.

RESULTS

Brain Parenchymal Volume
A scatterplot of brain parenchymal volume from adult male and female subjects as a function of age is shown in Fig 2A. The mean brain volume for men is 1209 ± 133.6 cm³, and for women, 1056 ± 107.4 cm³. Across all ages, whole-brain volumes were approximately 10% smaller for female subjects. The slopes of the linear regression trendlines for brain volume (men, $R = 0.413, R^2 = 0.17$; women, $R = 0.425, R^2 = 0.18$) are not significantly different between the sexes ($P = .5451$ (on-line Table 1)). The loss of brain volume calculated from the linear trendline from 20 to 100 years is 14.0% for men and 13.9% for women, without a significant difference.

Brain Parenchymal Fraction
BPF from adult male and female subjects as a function of age is shown in Fig 2B. The brain parenchymal fraction ranges from approximately 0.9 to 0.75, showing a curvilinear decline with age. The mean BPF is 0.843 for women and 0.847 for men and the difference was

not significant between the sexes (Student $t$ test, $P = .48$). Linear regression with age shows $R^2 = 0.51$ for men and 0.62 for women. BPF shows a decreased coefficient of variation relative to brain volume (3.085% vs 11.06% for men, 4.632% vs 10.2% for women; male data is shown in On-line Table 1).

The bar charts (Fig 2C and 2D) show decade groupings of the age cohorts. The BPF shows increased correlation with age, with decreased SD relative to brain volume measure. Statistical correlations of brain parenchymal volume and BPF are shown in On-line Table 2.

**Brain Parenchymal Density**
The brain parenchymal density (Fig 3A) is taken as the mean Hounsfield unit number for total brain parenchyma. Considerable
variation is seen in the mean parenchymal density for both sexes. The mean Hounsfield unit density for men is 30.29 ± 3.20, and for women, 29.58 ± 2.70. Male mean brain parenchymal density is slightly-but-significantly greater than for women (P = .0263). The mean brain parenchymal density declines significantly with age from 31.34 to 28.91 for men (P < .0061, −7.74%) and from 30.64 to 28.54 for women (P = .0026, −6.8%). The slopes of the linear regression lines are not significantly different (P = .76). Analysis by decade grouping by age shows that the third decade (20–29 years) is significantly different from the older groups (male data shown in On-line Table 3).

**Brain Parenchymal Radiomass**

Brain volume multiplied by brain radiodensity yields total brain radiomass, measured as cubic centimeter × Hounsfield unit (Fig 3B). The mean brain radiomass was 36,671.6 ± 5954 for men and 29,291 ± 8816 for women, with female brain radiomass being consistently approximately 20% less than for men. For women, the brain mass was 34,876 at 18 years and 27,673 at 100 years, a 20.7% decline. For men, the brain mass was 40,679 at 18 years and 31,709 at 100 years, a 22% decline, with differences in regression slopes not significant (P = .325). Significance between decade cohorts is achieved between the third decade and the older decade cohorts (male data shown in On-line Table 3).

**Quantile Regression**

BPF was chosen to illustrate quantile regression. The overall polynomial regression of third degree of age was fitted on the brain parenchymal fraction, after adjusting for sex. Overall polynomial regression indicated differences between men and women (mean difference = −0.005; 95% confidence interval, 0.01 to −1e-04; P = .046). Therefore, we fitted a quantile regression model of third degree of age on brain parenchymal fraction for men (Fig 4A) and
women (Fig 4B), respectively, with a polynomial curve of third degree of age plotted at 0.05-tile and 0.95-tile, and 0.10-tile and 0.90-tile. The coefficients of the polynomial regression models are shown in the Table.

**DISCUSSION**

Our study is a quantitative analysis of clinical images. Most brain studies of this type are performed with MR imaging, on healthy volunteers under ideal conditions. The goal of the current study was the development of a clinical tool; our study is derived from clinical images so that patients can be compared with their clinical peer group using the same imaging technique under identical conditions. While current image interpretation is largely qualitative, an automated quantitative analysis together with statistical methods such as quantile regression enables a quantitative assessment of brain parameters.

The data show large variance in brain volumes, consistent with previous reports that total brain size can vary almost 2-fold among individuals of the same age. Adult brain volumes are consistent with published literature, with our study focusing on the parenchymal volume. The female brain parenchymal volume is 1056 cm³, approximately 12.6% smaller than for men (mean parenchymal volume, 1209 cm³) not adjusted for body size, with a near-linear 14.9% decrease in male brain volume and 12.6% smaller than the male brain mass, whereas the brain volume is only 10% smaller. Both sets of data illustrate that brain mass declines significantly more than is reflected by a measure of brain volume alone.

Quantile regression is used as a quantitative method for comparing a given case with the total pool of reference cases. BPF was chosen for illustration. Third-order polynomial was best fit for the data. Small-but-measurable differences were seen between men and women, with slightly greater volume loss in men at older ages (Table). Similar regressions could be performed for each type of quantitative metric. New cases could be described in terms of the quantile position relative to the reference database.

**Study Limitations**

A potential criticism of this database is that included studies may not be “normal” because each study is obtained for a clinical reason. The goal of this study was not to identify true normal but to characterize the clinical population and develop a clinically useful database with which future studies might be compared. A large database together with statistical methods will approximate normalcy (i.e., the law of large numbers applies) or at least provide a clinically useful reference. CT entails radiation exposure and recruiting large numbers of healthy volunteers from the clinical archive facilitates the development of a large database of both sexes across the life span.

All data for this study originated from a single CT scanner. Although contemporary scanners are highly calibrated, differences between scanners or scan protocols could introduce an additional variable. Measured volumes and BPFs would be expected to show little variance across scanners, though measured Hounsfield units may vary. Data from different scanners may be merged using statistical methods such as z scoring; more rigorous calibration standards may be necessary for broad implementation of a standard database.
CONCLUSIONS
This study demonstrates that a large pool of clinical CT data can be subject to automated analysis to yield brain metrics supported by the existing literature. The total brain metrics of BFP, brain parenchymal density, and brain parenchymal mass derived from CT images are novel reports and show high correlation with loss of brain matter as a function of the aging process. Clinical head CT data can be subject to analytic methods to quantitatively assess new studies in the context of a clinical peer group.

Disclosures: Keith Cauley—RELATED: Grant: Geisinger Clinical Research Grant; Comments: internal research grant, no money paid to me; UNRELATED: Employment: Geisinger Medical Center.

REFERENCES
ABSTRACT

BACKGROUND AND PURPOSE: Despite the improved prognostic relevance of the 2016 WHO molecular-based classification of lower-grade gliomas, variability in clinical outcome persists within existing molecular subtypes. Our aim was to determine prognostically significant metrics on preoperative MR imaging for lower-grade gliomas within currently defined molecular categories.

MATERIALS AND METHODS: We undertook a retrospective analysis of 306 patients with lower-grade gliomas accrued from an institutional data base and The Cancer Genome Atlas. Two neuroradiologists in consensus analyzed preoperative MRIs of each lower-grade glioma to determine the following: tumor size, tumor location, number of involved lobes, corpus callosum involvement, hydrocephalus, midline shift, eloquent cortex involvement, ependymal extension, margins, contrast enhancement, and necrosis. Adjusted hazard ratios determined the association between MR imaging metrics and overall survival per molecular subtype, after adjustment for patient age, patient sex, World Health Organization grade, and surgical resection status.

RESULTS: For isocitrate dehydrogenase (IDH) wild-type lower-grade gliomas, tumor size (hazard ratio, 3.82; 95% CI, 1.94–7.75; P < .001), number of involved lobes (hazard ratio, 1.70; 95% CI, 1.28–2.27; P < .001), hydrocephalus (hazard ratio, 4.43; 95% CI, 1.12–17.54; P = .034), midline shift (hazard ratio, 1.16; 95% CI, 1.03–1.30; P = .013), margins (P = .031), and contrast enhancement (hazard ratio, 0.34; 95% CI, 0.13–0.90; P = .030) were associated with overall survival. For IDH-mutant 1p/19q-codeleted lower-grade gliomas, tumor size (hazard ratio, 2.85; 95% CI, 1.06–7.70; P = .039) and ependymal extension (hazard ratio, 6.34; 95% CI, 1.07–37.59; P = .042) were associated with overall survival.

CONCLUSIONS: MR imaging metrics offers prognostic information for patients with lower-grade gliomas within molecularly defined classes, with the greatest prognostic value for IDH wild-type lower-grade gliomas.

ABBREVIATIONS: LGG = lower-grade glioma; HR = hazard ratio; IDH = isocitrate dehydrogenase; IDHmut-Codel = IDH mutation and a whole-arm deletion of chromosome arms 1p and 19q; IDHmut-Noncodel = IDH-mutant lacking 1p/19q codeletion; IDHwt = IDH wild-type; IQR = interquartile range; OS overall survival; TCIA = The Cancer Imaging Archive; TERT = telomerase reverse transcriptase; WHO = World Health Organization

Recidence of the biologic and prognostic significance of molecular-based characterization of diffuse lower-grade gliomas (LGGs) heralded major revisions to their classification by the World Health Organization (WHO) in 2016.1–3 LGGs encompass WHO grade II and III astrocytic and oligodendrogial tumors, most of which have a prognostically favorable mutation in the isocitrate dehydrogenase (IDH) gene.2,4 Oligodendrogliomas are defined by the presence of both an IDH mutation and a whole-arm deletion of chromosome arms 1p and 19q (IDHmut-Codel), which confer added prognostic and therapeutic favorability.5,6 IDH-mutant astrocytomas (IDHmut-Noncodel) lack 1p/19q codeletion and typically have tumor protein p53 and alpha-thalassemia/mental retardation syndrome X-linked (ATRX) chromatin remodeler, [ATRX]) gene mutations.7,8 IDH wild-type LGGs (IDHwt) are generally associated with far worse clinical outcomes and largely comprise neoplasms that are genotypically similar to primary glioblastoma.1,9

Despite improved clinical applicability of the revised classification system, there remains substantial heterogeneity in clinical outcomes within existing subtypes of LGGs.10–12 While numerous
studies have reported neuroimaging features that predict currently defined glioma molecular subtypes.15–20 Neuroimaging features might additionally contain information that allows prognostic stratification of gliomas within currently defined molecular categories. The purpose of our investigation was to determine whether neuroimaging features on preoperative anatomic MR imaging have prognostic significance for LGGs within currently defined molecular categories.

MATERIALS AND METHODS
This was a retrospective study with institutional review board approval (University of Virginia Health System, Charlottesville, Virginia) as well as Health Insurance Portability and Accountability Act compliance.

Patient Selection
Cases were accrued from a diffuse LGG data base maintained at our institution (n = 255) as well as from The Cancer Imaging Archive (TCIA) (n = 198).21 Inclusion criteria were the following: 1) known IDH mutation status and 1p/19q-codeletion status; 2) known surgical resection status (gross total resection, subtotal resection, or biopsy); and 3) available preoperative MR imaging including (at a minimum) precontrast T1WI, contrast-enhanced T1WI, and either T2WI or FLAIR.

From the institutional dataset, 178 cases met the inclusion criteria; 43 cases were excluded for lack of IDH and/or 1p/19q-codeletion status, and 34 cases were excluded for lack of requisite preoperative MR imaging. From the TCIA cohort, 128 cases met the inclusion criteria; 1 case was excluded for lack of IDH and 1p/19q-codeletion status and 69 cases lacked requisite preoperative MR imaging. In total, our study cohort included 306 cases. Overall survival times were available for all included patients, defined as the time between the date of pathologic diagnosis to the date of death or last contact.

Neuroimaging and Analysis
Two neuroradiologists with 6 and 14 years of experience, blinded to molecular status, WHO grade, patient demographics, and clinical outcome, analyzed the preoperative MR imaging to determine the following metrics:—1) tumor size: maximum long-axis diameter (centimeters); 2) location: any glioma signal abnormality/enhancement involving the basal ganglia, thalamus, or brain stem (yes = central; no = peripheral); 3) number (n) of involved lobes: each of the following counted as 1 lobe (per hemisphere)—frontal lobe, parietal lobe, temporal lobe, insula, occipital lobe, brain stem/cerebellum; 4) corpus callosum involvement: any glioma signal abnormality/enhancement involving the corpus callosum (yes or no); 5) hydrocephalus: (yes or no); 6) midline shift: greatest degree of contralateral brain displacement (centimeters); 7) eloquent cortex involvement: any glioma signal abnormality/enhancement involving the eloquent cortex, defined per Chang et al.22 (yes or no); 8) ependymal extension: any glioma signal abnormality/enhancement involving the ventricular ependyma (yes or no); 9) margins: <33%, 33%–66%, or ≥66% sharp/circumscribed glioma margins; 10) contrast enhancement: any glioma contrast enhancement (yes or no); and 11) necrosis: any region of glioma necrosis characterized by peripheral contrast enhancement and central nonenhancement (yes or no).

Neuropathology
Pathologic data for patients in the institutional dataset were retrieved from the electronic medical record. Molecular markers were tested in the Clinical Laboratory Improvement Amendments–certified molecular pathology laboratory at our institution. A clinically validated IDH1 antibody was used to determine IDH1 mutation status.23,24 Clinically validated pyrosequencing assays (PyroMark Q24 system; https://www.qiagen.com/us/shop/automated-solutions/pyromark-q24/) were used in immunohistochemistry-negative cases in accordance with the manufacturers to determine IDH1/IDH2 mutation status.

<p>| Table 1: Patient characteristics for each molecular subtype |
|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>IDHmut-Codeletion</th>
<th>IDHmut-Noncodel</th>
<th>IDHwt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total patients*</td>
<td>104 (34.0%)</td>
<td>129 (42.2%)</td>
</tr>
<tr>
<td>Female sex</td>
<td>58 (55.8%)</td>
<td>61 (47.3%)</td>
</tr>
<tr>
<td>Age ≥50</td>
<td>48 (77.4%)</td>
<td>35 (53.8%)</td>
</tr>
<tr>
<td>WHO grade II</td>
<td>72 (69.2%)</td>
<td>80 (62.0%)</td>
</tr>
<tr>
<td>WHO grade III</td>
<td>32 (30.8%)</td>
<td>46 (35.7%)</td>
</tr>
<tr>
<td>*Data are listed as count (percentage of total cohort).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>*Data are listed as median (range). Other data are listed as count (percentage of cases within given molecular subtype).</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Empiric distribution summaries for the preoperative MR imaging metrics according to LGG molecular subtype

<table>
<thead>
<tr>
<th>MR Imaging Metric</th>
<th>IDHmut-Codeletion</th>
<th>IDHmut-Noncodel</th>
<th>IDHwt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor size (cm)</td>
<td>6.6</td>
<td>6.9</td>
<td>6.2</td>
</tr>
<tr>
<td>IQR (1st to 3rd quartile)</td>
<td>(5.2–8.6)</td>
<td>(5.3–9.3)</td>
<td>(4.6–8.4)</td>
</tr>
<tr>
<td>Range (minimum–maximum)</td>
<td>(1.9–14.2)</td>
<td>(1.7–15.3)</td>
<td>(1.1–15.0)</td>
</tr>
<tr>
<td>Location</td>
<td>26 (25.0%)</td>
<td>43 (33.3%)</td>
<td>29 (39.7%)</td>
</tr>
<tr>
<td>Central</td>
<td>78 (75.0%)</td>
<td>86 (66.7%)</td>
<td>44 (60.3%)</td>
</tr>
<tr>
<td>Peripheral</td>
<td>0.00</td>
<td>0.30</td>
<td>0.00</td>
</tr>
<tr>
<td>No. of involved lobes</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>IQR (1st to 3rd quartile)</td>
<td>(1.0–3.0)</td>
<td>(1.0–3.0)</td>
<td>(1.0–3.0)</td>
</tr>
<tr>
<td>Range (minimum–maximum)</td>
<td>(1.0–6.0)</td>
<td>(1.0–9.0)</td>
<td>(1.0–8.0)</td>
</tr>
<tr>
<td>Corpus callosum extension</td>
<td>44 (42.3%)</td>
<td>46 (35.7%)</td>
<td>18 (24.2%)</td>
</tr>
<tr>
<td>Hydrocephalus</td>
<td>12 (11.5%)</td>
<td>27 (20.9%)</td>
<td>7 (9.6%)</td>
</tr>
<tr>
<td>Midline shift (cm)</td>
<td>0.00</td>
<td>0.30</td>
<td>0.00</td>
</tr>
<tr>
<td>Ependymal extension</td>
<td>33 (31.7%)</td>
<td>45 (34.9%)</td>
<td>24 (32.9%)</td>
</tr>
<tr>
<td>Margin</td>
<td>65 (62.5%)</td>
<td>80 (62.0%)</td>
<td>58 (79.5%)</td>
</tr>
<tr>
<td>&lt;33% sharp/circumscribed</td>
<td>32 (30.8%)</td>
<td>31 (24.0%)</td>
<td>28 (38.4%)</td>
</tr>
<tr>
<td>33%–66% sharp/circumscribed</td>
<td>55 (52.9%)</td>
<td>38 (29.5%)</td>
<td>31 (42.5%)</td>
</tr>
<tr>
<td>≥66% sharp/circumscribed</td>
<td>17 (16.3%)</td>
<td>60 (46.5%)</td>
<td>14 (19.2%)</td>
</tr>
<tr>
<td>Contrast enhancement</td>
<td>21 (31.8%)</td>
<td>51 (39.5%)</td>
<td>30 (41.1%)</td>
</tr>
<tr>
<td>Necrosis</td>
<td>17 (16.3%)</td>
<td>14 (10.9%)</td>
<td>18 (24.7%)</td>
</tr>
</tbody>
</table>
The 1p/19q-codeletion status was determined by fluorescence in situ hybridization; dual-color human probes localizing the 1p, 1q, 19p, and 19q loci were used (Vysis LSI 1p36/1q25 and LSI 19q13/19p13 FISH Probe Kit; Abbott Laboratories, Abbott Park, IL). Histopathologic and molecular data for the TCIA cohort were retrieved from supplemental material in Ceccarelli et al.25

**Statistical Analysis**

Data Summarization. Categoric variables are summarized by frequencies and percentages, and continuous scaled data are summarized by the median, the interquartile range (IQR), and the range of the empiric distribution.

Survival Analyses. Cox proportional hazards regression was used to examine the prognostic utility of the preoperative MR imaging metrics to predict survival within the currently defined molecular categories of LGGs. The survival analyses were conducted per molecular category (ie, IDHmut-Codel, IDHmut-NonCodel, and IDHwt) and per MR imaging metrics. For each MR imaging metric, a multivariate Cox proportional hazards regression analysis was conducted in which the MR imaging metric of interest served as the primary predictor variable and patient age, patient sex, WHO grade, and surgical resection status served as concomitant adjustment variables. The follow-up times of survivors were treated as right-censored survival times in the Cox model. With

**FIG 1.** Kaplan-Meier survival curves based on the LGG molecular subtype. Vertical tick marks identify right-censored survival times. The survival curves differed among all 3 LGG molecular subtypes (IDH-mutant 1p/19q-codeleted molecular subtype versus IDH-mutant 1p/19q-noncodeleted molecular subtype: \( P = .021 \); IDH-mutant 1p/19q-codeleted molecular subtype versus IDH wild-type molecular subtype: \( P < .001 \); and IDH-mutant 1p/19q-noncodeleted molecular subtype versus IDH-wild-type molecular subtype: \( P < .001 \)). Survival curves for patients who composed a subset of the current patient cohort are shown in Patel et al,46 in 2019.

**Table 3: Associations between preoperative MR imaging metrics and overall survival per LGG molecular subtype, after adjustment for patient age, sex, tumor grade, and surgical resection status**

<table>
<thead>
<tr>
<th>Relative Comparison (Non-Reference:Reference)(^a)</th>
<th>IDHmut-Codel</th>
<th>IDHmut-Noncodel</th>
<th>IDHwt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor size (3rd to 1st quartile)</td>
<td>2.85 (1.06–7.70) (( P = .039 ))</td>
<td>2.90 (1.54–5.48) (( P &lt; .001 ))</td>
<td>3.82 (1.94–7.75) (( P &lt; .001 ))</td>
</tr>
<tr>
<td>Location (central:peripheral)</td>
<td>0.59 (0.16–2.18) (( P = .430 ))</td>
<td>0.50 (0.22–1.17) (( P = .101 ))</td>
<td>0.78 (0.36–1.70) (( P = .534 ))</td>
</tr>
<tr>
<td>No. of involved lobes (X+1X)</td>
<td>1.37 (0.84–2.23) (( P = .211 ))</td>
<td>1.25 (0.99–1.58) (( P = .065 ))</td>
<td>1.70 (1.28–2.27) (( P = .001 ))</td>
</tr>
<tr>
<td>Corpus callosum extension (yes:no)</td>
<td>2.60 (1.69–9.84) (( P = .160 ))</td>
<td>1.42 (0.59–3.41) (( P = .433 ))</td>
<td>1.89 (0.78–4.58) (( P = .160 ))</td>
</tr>
<tr>
<td>Hydrocephalus (yes:no)</td>
<td>2.57 (0.42–15.81) (( P = .308 ))</td>
<td>0.63 (0.19–2.11) (( P = .457 ))</td>
<td>4.43 (1.12–17.54) (( P = .034 ))</td>
</tr>
<tr>
<td>Midline shift (3rd to 1st quartile)</td>
<td>1.32 (0.77–2.26) (( P = .320 ))</td>
<td>1.14 (0.54–2.41) (( P = .730 ))</td>
<td>1.16 (0.10–1.30) (( P = .013 ))</td>
</tr>
<tr>
<td>Eloquent cortex (yes:no)</td>
<td>1.00 (0.29–3.43) (( P = .740 ))</td>
<td>1.91 (0.77–4.78) (( P = .165 ))</td>
<td>2.01 (0.90–4.49) (( P = .087 ))</td>
</tr>
<tr>
<td>Ependymal extension (yes:no)</td>
<td>6.34 (1.07–37.59) (( P = .042 ))</td>
<td>1.72 (0.64–4.64) (( P = .289 ))</td>
<td>1.51 (0.52–4.35) (( P = .447 ))</td>
</tr>
<tr>
<td>Margins (global test ( P ) value)(^b)</td>
<td>(( P = .775 ))</td>
<td>(( P = .190 ))</td>
<td>(( P = .031 ))</td>
</tr>
<tr>
<td>33%–66%&lt;33%(^d)</td>
<td>1.57 (0.42–5.97) (( P = .505 ))</td>
<td>0.51 (0.19–13.33) (( P = .366 ))</td>
<td>0.39 (0.17–0.90) (( P = .027 ))</td>
</tr>
<tr>
<td>&gt;66%&lt;33%(^d)</td>
<td>0.98 (0.09–10.85) (( P = .985 ))</td>
<td>0.39 (0.13–11.9) (( P = .098 ))</td>
<td>0.24 (0.06–1.05) (( P = .057 ))</td>
</tr>
<tr>
<td>&gt;66%33%–66%(^d)</td>
<td>0.62 (0.06–6.40) (( P = .688 ))</td>
<td>0.77 (0.24–2.38) (( P = .646 ))</td>
<td>0.63 (0.14–2.87) (( P = .554 ))</td>
</tr>
<tr>
<td>Contrast enhancement (yes:no)</td>
<td>3.11 (0.71–13.68) (( P = .332 ))</td>
<td>1.67 (0.62–4.53) (( P = .313 ))</td>
<td>0.34 (0.13–0.90) (( P = .030 ))</td>
</tr>
<tr>
<td>Necrosis (yes:no)</td>
<td>1.84 (0.32–10.78) (( P = .498 ))</td>
<td>0.28 (0.03–2.42) (( P = .247 ))</td>
<td>1.93 (0.82–4.58) (( P = .134 ))</td>
</tr>
</tbody>
</table>

\(^a\)Data are listed as adjusted hazard ratios (95% confidence interval) and corresponding \( P \) values.

\(^b\)Relative comparison (nonreference:reference) identifies the nonreference predictor variable value/level and the reference predictor variable value/level at which the adjusted instantaneous risk of death (ie, hazard ratio) was evaluated. For example, 3rd to 1st quartile represents a comparison of the adjusted instantaneous risk of death (ie, hazard) between 2 patients, 1 patient whose predictor variable value is at the 3rd quartile of the predictor variable empiric distribution (nonreference) and 1 patient whose predictor variable value is at the 1st quartile of the predictor variable empiric distribution (reference). A relative comparison denoted as (\( X + 1 \)) represents a comparison of the adjusted instantaneous risk of death (ie, hazard) between 2 patients, 1 patient whose predictor variable value is \( X + 1 \) units (nonreference) and 1 patient whose predictor variable value is \( X \) units (reference). A relative comparison, denoted as (yes:no), represents a comparison of the adjusted instantaneous risk of death (ie, hazard) between 2 patients, 1 patient who has the factor of interest (nonreference = yes) and 1 patient who does not have the factor of interest (reference = no). Note, if the adjusted hazard ratio is >1 (<1), it indicates that the estimate for the instantaneous risk of death is greater (less) for the patient who has the nonreference predictor variable/level.

\(^c\)Global test \( P \) value is the \( P \) value for testing the null hypothesis that the instantaneous risk of death is same for all “margin” categories.

\(^d\)Percentage of LGGs displaying sharp/circumscribed margin.
FIG 2. Representative cases of IDHwt LGG. A 51-year-old man with an IDH wild-type diffuse astrocytoma. FLAIR (A) shows a 6-cm mass in the right temporal lobe with ill-defined margins. Contrast-enhanced TIWI (B) shows no contrast enhancement of the mass. This patient is deceased, with a survival time of 262 days. Further molecular testing in this case was positive for IDHwt LGG, with a BRAF V600E mutation. A 26-year-old woman with an IDH wild-type diffuse astrocytoma. FLAIR (C) and contrast-enhanced TIWI (D) show a 1.6-cm left frontal lobe mass with fairly well-circumscribed margins and contrast enhancement. This patient was alive at last follow-up, with a survival time of 2757 days. Further molecular testing in this case was positive for the BRAF V600E mutation.

regard to hypothesis testing, the Wald $\chi^2$ test was used to test the null hypothesis that the instantaneous risk of death (ie, hazard) is not associated with the values/categories of the MR imaging metric of interest after adjustment for patient age, patient sex, WHO grade, and surgical resection status. A P ≤ .05 decision rule was used as the null hypothesis rejection criterion for the tests of association.

Tumor Grade Analyses. MR imaging metrics versus WHO grade association were examined via multivariate logistic regression, per LGG molecular subtype. The dependent variable of the multivariate logistic regression model was an indicator variable (Y) that distinguished WHO grade III tumors (Y = 1) from WHO grade II tumors (Y = 0), and the MR imaging metrics served as the predictor variables. Regarding hypothesis testing, the type III Wald $\chi^2$ test was used to test the null hypothesis that the adjusted odds for a tumor being WHO grade III are not associated with the MR imaging metric. A P ≤ .05 decision rule was used as the null hypothesis rejection criterion.

RESULTS

The study population included 306 patients, with 154 women (50.3%) and 152 men (49.7%). Patient characteristics per LGG molecular subtype are shown in Table 1, and preoperative MR imaging metric empirical distribution summaries per LGG molecular subtype are shown in Table 2.

Survival curves for overall survival (OS) are shown in Fig 1 for the LGG molecular subtypes. OS differed among all the LGG molecular subtypes (P < .02 for all comparisons), with patients with IDHwt LGG having the shortest median OS (49.5 months; 95% CI lower bound, 25.5 months), followed by the patients with IDHmut-Noncodel (134.6 months; 95% CI lower bound, 98.7 months), and patients with IDHmut-Codel (196.6 months, 95% CI lower bound, 141.4 months).

The associations between preoperative MR imaging metrics and OS per LGG molecular subtype are expressed as hazard ratios (HRs) in Table 3. After adjustment for patient age, patient sex, tumor grade, and surgical resection status. For IDHwt LGG, greater tumor size (HR, 1.70; 95% CI, 1.28–2.27; P < .001), greater number of involved lobes (HR, 1.70; 95% CI, 1.28–2.27; P < .001), and ependymal extension (HR, 6.34; 95% CI, 1.07–37.59; P = .042) were associated with shorter OS. A secondary analysis determined unique associations between MR imaging metrics and tumor grade (II or III) per LGG molecular subtype (Table 4). Contrast enhancement was associated with grade III for both IDHwt LGG (OR, 16.75; 95% CI, 3.47–80.86; P < .001) and IDHmut-Noncodel LGG (OR, 6.08; 95% CI, 2.12–17.41; P = .001), but not for IDHmut-Codel LGG. For IDHwt LGG, ependymal extension was associated with grade II (OR, 0.09; 95% CI, 0.01–0.59; P = .012), and for IDHmut-Noncodel LGG, central...
location was associated with grade II (OR, 0.21; 95% CI, 0.05–0.80; P = .035).

**DISCUSSION**

Despite the markedly improved prognostic relevance of the 2016 WHO classification of LGGs, clinical outcome variability persists for LGGs within existing molecular subtypes.26-29 For each LGG molecular subtype, we separately explored the association between preoperative MR imaging metrics and patient OS after adjustment for WHO grade, surgical resection status, patient age, and patient sex. We further explored the relationship between MR imaging metrics and tumor grade separately for each LGG subtype. Among our results, we found that preoperative MR imaging features have more prognostic value for IDHwt LGGs than for IDH-mutant LGGs. Tumor size is significantly associated with OS in all LGG molecular subtypes, and MR imaging associations with tumor grade vary among LGG molecular subtypes.

Our finding that preoperative MR imaging has the greatest prognostic value for the IDHwt subtype is potentially explained by the increasingly well-recognized biologic and clinical heterogeneity of IDHwt LGGs, in particular based on molecular alterations such as telomerase reverse transcriptase (TERT) promoter mutation (Fig 2A, -B), epidermal growth factor receptor (EGFR) gene amplification, and chromosome 7/10 alterations.12,26,27,30 These molecular alterations have not yet been formally incorporated into the WHO classification scheme, and it is conceivable that some of our neuroimaging metrics correlated with these prognostically significant molecular alterations among our IDHwt cohort. We further speculate that our apparently counterintuitive finding of a positive correlation between contrast enhancement and survival time for the IDHwt cohort could be explained by the potential inclusion of B-raf proto-oncogene (BRAF)-mutant IDHwt gliomas in our cohort (Fig 2C, -D). BRAF-mutant diffuse gliomas commonly demonstrate contrast enhancement and are associated with favorable clinical outcomes.31-33

On the other hand, we found relatively few significant associations with overall survival for the IDH-mutant LGG subtypes. These results might reflect that the IDH-mutant LGG subtypes represent more uniform genomic entities compared with the IDHwt subtype. Nonetheless, for both IDHwt-Noncodel and IDHmut-Codel subtypes, larger tumor size was associated with worse OS. Moreover, we found that ependymal extension is associated with worse OS for the IDHmut-Codel subtype. It is known that subventricular zone extension by glioblastoma portends a worse prognosis, possibly due to recruitment of migratory progenitor cells in this location.34 It is unclear whether a similar process contributes to a more aggressive clinical course among the IDHmut-Noncodel LGGs, and further investigation into the impact of ependymal extension in the IDHmut-Codel subtype could build on our results.

To our knowledge, the literature investigating prognostically relevant imaging metrics in molecularly defined LGG subtypes is sparse. Perhaps the most applicable study is by Wu et al,35 in 2019, who found that tumor size and elevated relative cerebral blood volume on preoperative imaging correlate with a more aggressive subtype of IDHmut-Noncodel gliomas. While our results support these findings in regard to tumor size, we observed a similar relationship for IDHmut-Codel and IDHwt subgroups as well. Small sample size and lack of adjustment for WHO grade by Wu et al are limitations that may have contributed to differences observed between our studies. Suchorska et al,36 in 2019, found that contrast enhancement is a viable prognostic metric for IDH-mutant subgroups, a finding that was not reproduced in our study. Finally, the apparent diffusion...
coefficient has been reported as a potential prognostic marker for IDHwt LGGs.37

Multiple prior studies have investigated imaging findings that correlate with WHO grade.38-42 However, to the best of our knowledge, we are the first to report the relationship between neuroimaging metrics and WHO grade separately for each LGG molecular subtype. Our results indicate that contrast enhancement predicts WHO grade III for the IDHmut-Noncodel and IDHwt LGGs, but not for IDHmut- Codel LGGs. A previous study of oligodendrogliomas by White et al.43 in 2005, also reported no association between contrast enhancement and WHO grade; however, 1p/19q testing was not undertaken in their study, and their cohort was small (n = 24). Our results strengthen this conclusion and may serve to modify the well-established dictum that contrast enhancement correlates with tumor grade in adult diffuse gliomas.44,45

While efforts to identify preoperative neuroimaging predictors of LGG molecular status are valuable, the clinical utility is frequently limited because nearly all such cases undergo biopsy or resection for a definitive pathologic diagnosis. Our investigation aimed to uncover simple neuroimaging metrics that could relay prognostic information for LGGs beyond what can be inferred from their molecular and histologic characterization. An added strength of our study design was the adjustment of our analysis for both tumor grade and surgical resection status. The extent of surgical resection in particular is well-recognized for its impact on the overall survival of a patient with LGG.46,47 Nonetheless, our study has limitations. This is a retrospective study and thus inherently limited by design in its ability to infer relationships beyond association. Further prospective multi-institutional investigations would be necessary to confirm our results. Another potential limitation is the binary method of analysis for several metrics (eg, yes or no contrast enhancement). While this method was chosen to reduce ambiguity in the data acquisition and maintain a simple set of imaging metrics, a more graded approach to measurement may allow a more nuanced and potentially revealing analysis. Similarly, we measured “tumor size” as a single long-axis diameter, and volumetric measurements based on 3D MR imaging acquisitions would be preferable. Furthermore, we investigated a limited set of MR imaging pulse sequences (pre-contrast T1WI, contrast-enhanced T1WI, and either T2WI or FLAIR), and additional study into the prognostic value of more advanced techniques (eg, perfusion-weighted imaging, diffusion tensor imaging, susceptibility-weighted imaging) might be fruitful. We did not adjust our analysis for nonsurgical treatment (eg, chemotherapy and radiation therapy), which presumably varied across our cohort. Finally, future investigations evaluating links between MR imaging metrics and molecular alterations pertinent to IDHwt LGGs, such as BRAF or TERT mutations, would be highly valuable.

CONCLUSIONS

Preoperative MR imaging metrics offer prognostic information for patients with LGG within molecularly defined classes. Multiple neuroimaging features had unique prognostic significance for the IDHwt subtype, including hydrocephalus, midline shift, margin features, and contrast enhancement, likely reflecting the known biological and clinical heterogeneity of IDHwt LGGs. Contrast enhancement was associated with WHO grade III among IDHwt and IDHmut-Noncodel LGGs, but not IDHmut- Codel LGGs.


REFERENCES


ABSTRACT

BACKGROUND AND PURPOSE: Previous studies in acute ischemic stroke have demonstrated the importance of minimizing delays to endovascular treatment and keeping thrombectomy procedural times at <30–60 minutes. The purpose of this study was to investigate the impact of thrombectomy procedural times on clinical outcomes.

MATERIALS AND METHODS: We retrospectively compared 319 patients having undergone thrombectomy according to procedural time (<30 minutes, 30–60 minutes, and >60 minutes) and time from stroke onset to endovascular therapy (≤6 or >6 hours). Clinical characteristics of patients with postprocedural intracranial hemorrhage were also assessed. Logistic regression was used to determine independent predictors of poor outcome at 90 days (mRS ≥3).

RESULTS: Greater age (OR, 1.03; 95% CI, 1.01–1.06; P = .016), higher admission NIHSS score (OR, 1.10; 95% CI, 1.04–1.16; P = .001), history of diabetes mellitus (OR, 1.96; 95% CI, 1.05–3.65; P = .034), and postprocedural intracranial hemorrhage were independently associated with greater odds of poor outcome. Modified TICI scale scores of 2c (OR, 0.11; 95% CI, 0.04–0.28; P < .001) and 3 (OR, 0.15; 95% CI, 0.06–0.38; P < .001) were associated with reduced odds of poor outcome. Although not statistically significant on univariate analysis, onset to endovascular therapy of >6 hours was independently associated with increased odds of poor outcome (OR, 2.20; 95% CI, 1.11–4.36; P = .024) in the final multivariate model (area under the curve = 0.820). Procedural time was not independently associated with clinical outcome in the final multivariate model (P > .05).

CONCLUSIONS: Thrombectomy procedural times beyond 60 minutes are associated with lower revascularization rates and worse 90-day outcomes. Procedural time itself was not an independent predictor of outcome. While stroke thrombectomy procedures should be performed rapidly, our study emphasizes the significance of achieving revascularization despite the requisite procedural time. However, the potential for revascularization must be weighed against the risks associated with multiple thrombectomy attempts.

ABBREVIATIONS: ICH = intracranial hemorrhage; mTICI = modified TICI; OTE = onset to endovascular therapy; sICH = symptomatic intracranial hemorrhage

Ischemic stroke is a major source of morbidity and the fifth leading cause of death in the United States. Mechanical thrombectomy has been well-established as the standard of care for acute ischemic stroke treatment in selected patients presenting with a large-vessel occlusion. Advances in prehospital stroke care and triage, along with technical advances in mechanical thrombectomy, have resulted in significant improvement in timing, revascularization rates, and clinical outcomes following stroke thrombectomy.

In recent years, there has been investigation into the impact of procedural timing on clinical outcomes. The time from groin puncture to revascularization of <60 minutes, termed the “golden hour,” has been associated with improved outcomes following stroke thrombectomy. Additionally, a shorter time from stroke onset to endovascular therapy (OTE) has been linked to more favorable outcomes, including significant early recovery and increased odds of discharge to home. Notably, the association between OTE and improved outcome has been shown to be most significant within the first 2 hours following stroke onset, termed the “2 golden hours.”

In the current study, we investigated the significance of thrombectomy procedure duration and the time from stroke onset to endovascular treatment on clinical outcomes after
mechanical thrombectomy. Additionally, we investigated the impact of revascularization and postprocedural intracranial hemorrhage (ICH), among other variables, on clinical outcomes in patients treated with mechanical thrombectomy.

**MATERIALS AND METHODS**

This study was conducted with local institutional review board approval (University of Maryland, Baltimore). We performed a retrospective review of a prospectively maintained data base of all patients undergoing mechanical thrombectomy for the treatment of acute ischemic stroke at a single comprehensive stroke center from April 2012 through February 2019. Patients eligible for inclusion in the study included patients 18 years of age or older who underwent mechanical thrombectomy within 24 hours of stroke onset for treatment of an acute ischemic stroke involving the anterior cerebral circulation.

Data collected included patient demographics, clinical presentation, procedural details, angiographic/radiographic findings, and clinical outcome at 90-day follow-up. Anterior circulation occlusions were defined as occlusions of the ICA, M1 segment of the MCA, or M2 segment of the MCA. Angiographic revascularization grade classified according to the modified TICI score (mTICI) was determined by the neurointerventional surgeon (D.G., T.R. M., G.J., E.L.) and 1 neurointerventional fellow. Differences were adjudicated in a consensus reading when necessary. Successful revascularization was defined as mTICI 2b, 2c, or 3.

A thrombectomy pass was defined as any stent retrieval thrombectomy attempt or an attempt using aspiration without a stent retriever. Procedural time was grouped according to groin puncture to procedure end time of <30 minutes, 30–60 minutes, or >60 minutes. Procedural time was also dichotomized to less than or greater than 30 minutes for additional analyses. OTE was defined as the time from stroke onset or last known well to groin puncture. Patients were dichotomized according to OTE ≤6 hours or >6 hours. Postprocedural ICH was identified on postthrombectomy CT or MR imaging of the brain. Symptomatic intracranial hemorrhage (sICH) was defined as a deterioration in neurologic examination findings attributed to the hemorrhage, as defined by the European Cooperative Acute Stroke Study III definition.8,9 The primary outcome of interest was the mRS score at 90 days based on assessment by a neurologist or specialty-trained neurology nurse who did not perform the thrombectomy procedure. An mRS score of 0–2 was considered a good clinical outcome, and 3–6, a poor clinical outcome.

Procedures were performed using institutional protocols and standard of care. Patient selection for thrombectomy was based on findings from noncontrast CT, CT angiography, and/or MR imaging/MR angiography. Exclusion criteria for the study were the following: 1) mild stroke symptoms, defined as admission NIHSS score of less than four; 2) the presence of a completed large territorial infarction observed on noncontrast CT (ASPECTS <4) or MR imaging (>90-mL infarct volume); 3) functional dependence at baseline, defined as a prestroke mRS less than three; 4) intracranial hemorrhage noted within the targeted stroke territory; or 5) >24 hours from time of known stroke onset. Any discrepancy in the ASPECTS or eligibility for thrombectomy was adjudicated by multidisciplinary consensus between the attending neuroradiological surgeon and attending stroke neurologist. Procedures were performed in a biplane neuroangiography suite (Artis zee; Siemens) by 1 of 4 attending neuroradiological surgeons. Procedures were completed with the patient under moderate/conscious sedation or general anesthesia at the discretion of the attending neuroradiological surgeon.

**Statistical Analysis**

The distribution, frequency, and mean values of all clinical and demographic variables were assessed. We compared patients grouped according to procedural time of <30 minutes, 30–60 minutes, or >60 minutes and OTE of ≤6 hours or >6 hours using the χ² test for categorical variables and the Student t test/ANOVA for continuous variables. Bivariate analysis was used to compare clinical characteristics among patients who had postprocedural symptomatic hemorrhage and all other patients. Univariate and multivariate logistic regression was used to determine independent predictors of poor outcome at 90 days defined by mRS scores of 3–6.

Logistic regression was performed on all 319 patients. Variables on univariate analysis with a P value ≤ .10 were included in the final multivariable logistic regression. Select variables with P values >.10 on univariate analysis were included in the final logistic regression model due to their clinical relevance. The following variables were included in the final multivariate model: age, NIHSS score, tPA, medical history (hypertension, diabetes mellitus, hyperlipidemia, atrial fibrillation, smoking), OTE of ≤6 hours, procedural time (<30 minutes, 30–60 minutes, >60 minutes), mTICI score, moderate sedation versus general anesthesia, number of thrombectomy passes, and postprocedural ICH. The number of thrombectomy passes was grouped according to 1 pass, 2 passes, or ≥3 passes for all analyses. mTICI scores were categorized as 2a, 2b, 2c, or 3 for logistic regression analysis. We assessed potential interactions among all independent variables before inclusion in the final model. A receiver operating characteristic analysis was performed to measure the predictive power of the final regression model. All analyses were conducted using STATA/SE (Version16 for Windows; StataCorp). A P value < .05 was considered statistically significant for all tests.

**RESULTS**

**Total Cohort Demographics, Clinical/Procedural Characteristics, and Outcomes**

Three hundred thirty-four eligible patients underwent mechanical thrombectomy for ischemic stroke during the study period. Patients were excluded due to ischemic stroke involving the posterior circulation (n = 12), missing clinical information (n = 1), and loss to follow-up (n = 2), leaving a total of 319 study subjects. The average patient age was 65.10 ± 0.78 years. The mean admission NIHSS score was 17.71 ± 0.31. Sixty-four percent of patients had an occlusion of the M1 segment, while 22% and 14% had occlusions of the ICA and M2 segment, respectively.
Revascularization (mTICI 2b, 2c, 3) was achieved in 79% of cases, and full revascularization (mTICI 2c or 3), in 56% of cases. Mean groin puncture to procedure end time was 58.7 (2.0) minutes. Forty percent of patients (n = 129) from the 319 study subjects achieved a good outcome at 90 days. Patient demographics, clinical characteristics, procedural data, and outcomes are outlined in the On-line Table.

**Procedural Time**
A comparison of patients grouped according to procedural time of <30 minutes (n = 74), 30–60 minutes (n = 112), and >60 minutes (n = 133) is outlined in the On-line Table. There was no statistically significant difference in patient age, admission NIHSS score, IV-tPA usage, occlusion site/side, medical history, and time from stroke onset-to-endovascular therapy (P > .05). Shorter procedural times were associated with greater revascularization rates and fewer thrombectomy passes (P < .001) (On-line Table). Patients with procedural times of <30 minutes and 30–60 minutes had greater rates of good outcome at 90 days (46% and 48%, respectively) relative to those with a procedural time of >60 minutes (31%) (P = .012) (Fig 1A).

**Stroke Onset to Endovascular Therapy**
The mean time from stroke onset or last known well time to endovascular therapy was 244.29 ± 67.24 minutes in the OTE ≤6 hours group and 646.93 ± 312.86 minutes in the OTE >6 hours group (P < .001). The admission NIHSS score was greater for patients in the OTE ≤6 hours group (18.46 ± 5.35 versus 15.85 ± 5.75; P < .001), and there was a trend toward a greater percentage of M1 occlusions relative to ICA or M2 occlusions among patients in the OTE ≤6 hours group (P = .060). Patients with OTE of ≤6 hours were more likely to receive tPA before thrombectomy relative to patients with OTE >6 hours (57% versus 25%; P < .001). The mean time from groin puncture to procedure end was greater among patients in the OTE ≤6 hours group (61.47 ± 37.5 minutes versus 51.74 ± 30.37 minutes; P = .029). There was no significant difference in revascularization rates or the number of thrombectomy passes between the ≤6 hours versus >6 hours OTE groups. A comparison of patient demographics and clinical characteristics according to OTE ≤6 hours or >6 hours is shown in the On-line Table. Forty-three percent of patients in the OTE ≤6 hours group achieved a good outcome at 90 days relative to 35% in the OTE >6 hours group (P = .225) (Fig 1B).

**Postprocedural Hemorrhage**
Fifty-seven patients (18%) had a postprocedural ICH following thrombectomy, 16 of whom (5%) developed an sICH. There was no difference in admission NIHSS, tPA usage, occlusion site/side, medical history, and time from stroke onset-to-endovascular therapy (P > .05). The development of sICH was associated with a greater number of thrombectomy passes (P = .019). All patients who developed sICH (n = 16) experienced poor outcomes at 90 days (Table 1).

**Independent Predictors of Poor Outcome**
On univariate analysis, several patient demographic and clinical characteristics were associated with poor clinical outcome at
90 days (Table 2). On multivariable logistic regression, greater age (OR, 1.03; 95% CI, 1.01–1.06; \( P = .016 \)), higher admission NIHSS score (OR, 1.10; 95% CI, 1.04–1.16; \( P = .001 \)), ICH (OR, 3.48; 95% CI, 1.55–7.82; \( P = .003 \)), and a history of diabetes mellitus (OR, 1.96; 95% CI, 1.05–3.65; \( P = .034 \)) were independently associated with greater odds of poor outcome. Modified TICI scale scores of 2c (OR, 0.11; 95% CI, 0.04–0.28; \( P < .001 \)) and 3 (OR, 0.15; 95% CI, 0.06–0.38; \( P < .001 \)) were associated with reduced odds of poor outcome. Although not significant on univariate analysis, OTE \( > 6 \) hours was independently associated with increased odds of poor outcome (OR, 2.20; 95% CI, 1.11–4.36; \( P = .024 \)) after adjusting for confounding variables included in the final multivariate model (area under the curve = 0.820). Although procedural times of \( < 30 \) minutes and \( 30–60 \) minutes had significantly better rates of good outcome at 90 days relative to those with procedural times of \( > 60 \) minutes, procedural time itself and the number of thrombectomy passes did not have an independent association with clinical outcome on the final multivariate logistic regression (Table 2). An additional bivariate analysis comparing patients grouped according to procedural times less than or \( \geq 30 \) minutes showed no significant difference in outcome (mRS 0–2: 46% versus 39%, respectively; \( P = .271 \)).

**DISCUSSION**

The golden hour in stroke thrombectomy, defined by Spiotta et al\(^{5}\) as a procedural time of \( \leq 60 \) minutes, has been associated with improved outcomes in comparison with longer procedural times. Subsequent recent reports by Alawieh et al\(^{10}\) and Huang et al\(^{11}\) have reinforced the importance of shorter procedural times, noting worse outcomes and greater rates of sICH among patients with extended procedural times, specifically those extending beyond 30 minutes. Similar to these recent reports,\(^{10,11}\) greater rates of good functional outcomes among patients who underwent procedures lasting \( < 30 \) minutes or \( 30–60 \) minutes relative to procedures of \( > 60 \) minutes (\( P = .012 \)) were found. Our data also demonstrate a trend toward an increased risk of sICH in those with procedural times of \( > 30 \) minutes. Final TICI scores were significantly different across the procedural time groups. When we accounted for this notable discrepancy in the final multivariate regression analysis, procedural time and the number of thrombectomy passes were no longer independent predictors of outcome.

Procedural time, when dichotomized to less than or greater than \( 30 \) minutes, was, again, not independently associated with outcomes on multivariate analysis.\(^{10,11}\) Our results differ in this regard when compared with the reports by Alawieh et al\(^{10}\) and Huang et al.\(^{11}\)

Our results are similar, however, to recent findings of Tonetti et al\(^{12}\) and Jindal et al.\(^{13}\) The data from these studies suggest that the treatment effect still favors recanalization for patients, despite requiring a higher number of thrombectomy passes. Both of these groups concluded that successful recanalization should be pursued despite the number of attempts necessary. Our findings deviate from those reported in the original description of the first pass effect by Zaidat et al.\(^{14}\) However, the data from this report did not account for lower rates of revascularization in higher pass cohorts. We continue to believe that it is critical to expedite thrombectomy procedures as much as possible because lengthy procedures delay revascularization and confer some additional risk to the patient with each subsequent thrombectomy pass.

More recently, the 2 golden hours, coined by Peretz et al,\(^{6}\) defined the crucial 2-hour period between the time of stroke onset to endovascular therapy in which patients experience the greatest benefit from mechanical thrombectomy. These results are consistent with those of prior reports that described the importance of early stroke intervention.\(^{6,15,16}\) Our data similarly demonstrate that OTE \( > 6 \) hours is associated with increased odds of poor outcome (OR, 2.20; 95% CI, 1.11–4.36; \( P = .024 \)) on

---

### Table 1: Clinical and procedural characteristics associated with postthrombectomy symptomatic hemorrhage*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Postprocedural Symptomatic Hemorrhage (n = 16)</th>
<th>Asymptomatic/No Hemorrhage (n = 303)</th>
<th>( P ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIHSS</td>
<td>17.88 (6.48)</td>
<td>17.70 (5.55)</td>
<td>.905</td>
</tr>
<tr>
<td>tPA</td>
<td>9 (56%)</td>
<td>145 (48%)</td>
<td>.521</td>
</tr>
<tr>
<td>Occlusion site</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICA</td>
<td>3 (19%)</td>
<td>67 (22%)</td>
<td>.927</td>
</tr>
<tr>
<td>M1</td>
<td>11 (69%)</td>
<td>194 (64%)</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>2 (13%)</td>
<td>42 (14%)</td>
<td></td>
</tr>
<tr>
<td>mTICI scale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mTICI 0</td>
<td>0 (0%)</td>
<td>25 (8%)</td>
<td>.425</td>
</tr>
<tr>
<td>mTICI 1</td>
<td>1 (6%)</td>
<td>11 (4%)</td>
<td></td>
</tr>
<tr>
<td>mTICI 2a</td>
<td>3 (19%)</td>
<td>25 (8%)</td>
<td></td>
</tr>
<tr>
<td>mTICI 2b</td>
<td>4 (25%)</td>
<td>71 (23%)</td>
<td></td>
</tr>
<tr>
<td>mTICI 2c</td>
<td>1 (6%)</td>
<td>56 (18%)</td>
<td></td>
</tr>
<tr>
<td>mTICI 3</td>
<td>7 (44%)</td>
<td>115 (38%)</td>
<td></td>
</tr>
<tr>
<td>OTE (min)</td>
<td>308.06 (87.68)</td>
<td>361.85 (258.76)</td>
<td>.408</td>
</tr>
<tr>
<td>Groin puncture to procedure end (min)</td>
<td>64.45 (35.01)</td>
<td>58.39 (35.97)</td>
<td>.512</td>
</tr>
<tr>
<td>Procedural sedation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate sedation</td>
<td>9 (56%)</td>
<td>166 (55%)</td>
<td>.909</td>
</tr>
<tr>
<td>General anesthesia</td>
<td>7 (44%)</td>
<td>137 (45%)</td>
<td></td>
</tr>
<tr>
<td>No. thrombectomy passes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Pass</td>
<td>4 (25%)</td>
<td>109 (36%)</td>
<td>( P = .019 )</td>
</tr>
<tr>
<td>2 Passes</td>
<td>2 (12%)</td>
<td>104 (34%)</td>
<td></td>
</tr>
<tr>
<td>( \geq 3 ) Passes</td>
<td>10 (63%)</td>
<td>90 (30%)</td>
<td></td>
</tr>
<tr>
<td>Procedure time</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( &lt; 30 ) min</td>
<td>1 (6%)</td>
<td>73 (24%)</td>
<td>( P = .099 )</td>
</tr>
<tr>
<td>( \geq 30 ) min</td>
<td>15 (94%)</td>
<td>230 (76%)</td>
<td></td>
</tr>
<tr>
<td>mRS at 90 days</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mRS 0–2 (good)</td>
<td>0 (0%)</td>
<td>129 (43%)</td>
<td>.001</td>
</tr>
<tr>
<td>mRS 3–6 (poor)</td>
<td>16 (100%)</td>
<td>174 (57%)</td>
<td></td>
</tr>
<tr>
<td>Death</td>
<td>10 (63%)</td>
<td>59 (19%)</td>
<td>( &lt; .001 )</td>
</tr>
</tbody>
</table>

*Continuous variables are expressed as mean (± standard error or SD). Categoric binary variables are expressed as No. (%) \( P < .05 \) is considered statistically significant.
multivariable logistic regression analysis. An OTE of >6 hours was associated with lower rates of tPA usage, lower NIHSS scores, and a trend toward more M2-level occlusions. These factors are known to impact outcomes, and these were accounted for in our multivariable analysis. Patients in the OTE >6 hours group also had shorter procedural times than those in the OTE ≤6 hours group (51.74 [30.37] minutes versus 61.47 ([37.58] minutes; P = .029); this finding may reflect an operator tendency to spend less time on patients in the “late window” and/or might be related to the higher incidence of M2 occlusions seen in the OTE >6 hours group. We did not evaluate the impact of the 2 golden hours of OTE on patient outcomes, because only 15 of our patients were treated within 2 hours of onset. Most of our patients with stroke treated with thrombectomy were transferred to our facility from outside hospitals during the study period.6

Greater age and NIHSS scores, which are well-established predictors of outcome following stroke, were independent predictors of poor outcome in our final multivariate analysis (P < .05).2,17 Postprocedural ICH was associated with poor outcomes on multivariate analysis, and all patients with sICH experienced a poor outcome at 90 days. This outcome is similar to that in prior reports.10 Also similar to prior reports, a history of diabetes mellitus was found to be predictive of poor outcome (OR, 1.96; 95% CI, 1.05–3.65; P = .034).13,18

The present study has limitations. This is a single-center retrospective study, and the triage and thrombectomy techniques we used evolved during the course of the study. A single-center series, however, does ensure some homogeneity with regard to periprocedural and procedural care. Angiographic TICI scoring was initially performed by the operating physician, but subsequent verification of TICI scoring was performed by only 1 attending physician and 1 neurointerventional fellow. Although we created a thorough multivariate model to predict poor outcome with a high degree of predictive accuracy (area under the curve = 0.820), it is possible that we failed to include unmeasured markers of disease severity or other relevant confounding variables. While we attempted to account for differences between the 2 OTE groups in the multivariate analysis, these differences represent a limitation in comparing these groups. Moreover, although we included 319 patients, a larger, higher-powered study is likely needed to help us more precisely understand the role of procedural time itself on outcomes after stroke thrombectomy, because longer procedures and a greater number of thrombectomy attempts confer added risk to the patient.

CONCLUSIONS

When accounting for lower rates of revascularization in longer procedures, procedural time was not in itself highly predictive of patient outcomes. Thus, our study emphasizes the significance of achieving revascularization despite the requisite procedural time. However, pursuing longer procedures must be weighed against achieving revascularization despite the requisite procedural time. Therefore, future studies should focus on optimizing procedural times and improving outcomes in these challenging cases.
the added risk of intracranial hemorrhage associated with multiple thrombectomy attempts.

Disclosures: Gregory Cannarsa—UNRELATED: Employment: University of Maryland Medical Center. Timothy R. Miller—UNRELATED: Grants/Grants Pending: active grants: August, 01, 2013, to present; Co-Inv, Principal Investigator: Gaurav Jindal, (1%)

Intracranial Aneurysms

grants: August, 01, 2013, to present; Co-Inv, Principal Investigator: Gaurav Jindal, (1%)

“ULTRA: Ultra Coils from Start to Finish for the Endovascular Repair of Small Intracranial Aneurysms”: multicenter trial sponsor: Stryker; award amount: $774,774; assist in patient recruitment, data collection, September 1, 2014, to present; Co-Inv, Site Principal Investigator: Dheeraj Gandhi, (1%)

“BARREL: Prospective, Multi-Center, “Single-Arm Study of the Reverse Medical Barrell™ Vascular Reconstruction Device (VRD) for Adjunctive Treatment to Embolic Coils for Wide-Neck, Intracranial, Bifurcating/Branching Aneurysms of Middle Cerebral and Basal Arteries”; multicenter trial, sponsor: Medtronic; award amount: $26,232; assist in patient recruitment, data collection, September 15, 2015, to present; Co-Inv, Site Principal Investigator: Gaurav Jindal (1%)

“A Prospective, Multi-center, Single Arm Study to Evaluate the Safety and Effectiveness of the CODMAN ENTERPRISE® Vascular Reconstruction Device and Delivery System when Used in Conjunction with Endovascular Coil Embolization in the Treatment of Wide-necked Saccular Intracranial Aneurysms”; multicenter trial, sponsor: Ceragenus, DePuy Synthe; award amount: $999,995; assist in patient recruitment, data collection, January 30, 2018, to present; Co-Inv, Site Principal Investigator: Howard Eisenberg, (1%)

“A Pilot Clinical Trial of the Management of the Medically-Refractory Dyskinesia Symptoms or Motor Fluctuation of Advanced Idiopathic Parkinson’s Disease With Unilateral Lesioning of the Globus Pallidum Using the ExcAblate Neuro System”; multicenter trial, sponsor: InSightec; award amount: $830,971; assist in treatment planning and execution, February 21, 2018, to present; Co-Inv, Site Principal Investigator: Gaurav Jindal (1%)

“Ruptured Aneurysms Treated with Hydrogel Coils–RAGE”; multicenter trial, sponsor: MicroVention; award amount: $102,150; assist in patient recruitment, data collection, July 11, 2018, to present; Co-Inv, Site Principal Investigator: Graeme Woodworth, (1%)

“A Study to Evaluate the Safety and Feasibility of Exalate Model 4000 Type-2 to Temporarily Mediate Blood-Brain Barrier Disruption (BBBD) in Patients with Suspected Globolastoma in the Setting of Planned Surgical Interventions”; single-center trial, sponsor: InSightec; award amount: $239,027; assist in treatment planning and execution, August 16, 2018, to present; Co-Inv, Site Principal Investigator: Dheeraj Gandhi, (1%)

“PHIL: DAVF. Study of PHIL Embolic System in the Treatment of Intracranial Dural Arteriovenous Fistulas”; multicenter trial; sponsor: MicroVention; award amount: $310,000; assist in patient recruitment, data collection, completed grants, January 13, 2016, to June 1, 2019; Site Principal Investigator, (3%) “Apollo Onyx Delivery Micro Catheter Post Market Safety Study”; multicenter trial, sponsor: Medtronic; award amount: $33,500; assist in patient recruitment, data gathering, data analysis, November 2, 2011, to June 2, 2017; Co-Inv, Site Principal Investigator: Gaurav Jindal, (1%)

“Hydrogel Endovascular Aneurysm Treatment Trial (HEAT)”; multicenter trial, sponsor: Stryker; award amount: $34,889; assist in patient recruitment, data collection, November 16, 2013, to November 25, 2014; Co-Inv, Site Principal Investigator: Gaurav Jindal, (1%)

“Supra-aortic Vessels”; multicenter trial; sponsor: Bayer; award amount: $79,295; assist in patient recruitment, data collection, November 25, 2013, to December 11, 2017; Site Principal Investigator, (5%) “Trevor Retriever Registry: Post Market Surveillance Study”; multicenter trial, sponsor: Stryker; award amount: $337,430; assist in patient recruitment, data gathering, data analysis, November 10, 2014, to August 20, 2015; Co-Inv, Site Principal Investigator: Dheeraj Gandhi, (1%)

“Pivotal Study of MicroVention, Inc. Neurovascular Self-expanding Retrieval Stent System LVIS”; multicenter trial; sponsor: MicroVention, award amount: $453,350; assist in patient recruitment, data collection, December 8, 2015, to April 3, 2017; Co-Inv, Site Principal Investigator: Gaurav Jindal, (1%)


REFERENCES


ABSTRACT

BACKGROUND AND PURPOSE: Recanalization after coil embolization is widely studied. However, there are limited data on how recanalized aneurysms rupture. Herein, we describe our experience with the rupture of recanalized aneurysms and discuss the type of recanalized aneurysms at greatest rupture risk.

MATERIALS AND METHODS: A total of 426 unruptured aneurysms and 169 ruptured aneurysms underwent coil embolization in our institution between January 2009 and December 2017. Recanalization occurred in 38 (8.9%) of 426 unruptured aneurysms (unruptured group) and 37 (21.9%) of 169 ruptured aneurysms (ruptured group). The Modified Raymond-Roy classification on DSA was used to categorize the recanalization type. Follow-up DSA was scheduled until 6 months after treatment, and follow-up MRA was scheduled yearly. If recanalization was suspected on MRA, DSA was performed.

RESULTS: In the unruptured group, the median follow-up term was 74.0 months. Retreatment for recanalization was performed in 18 aneurysms. Four of 20 untreated recanalized aneurysms (0.94% of total coiled aneurysms) ruptured. In untreated recanalized aneurysms, class IIIb aneurysms ruptured significantly more frequently than class II and IIIa ($P = 0.025$). In the ruptured group, the median follow-up term was 28.0 months. Retreatment for recanalization was performed in 16 aneurysms. Four of 21 untreated recanalized aneurysms (2.37% of total coiled aneurysms) ruptured. Class IIIb aneurysms ruptured significantly more frequently than class II and IIIa ($P = 0.02$).

CONCLUSIONS: The types of recanalization after coil embolization may be predictors of rupture. Coiled aneurysms with class IIIb recanalization should undergo early retreatment because of an increased rupture risk.

ABBREVIATIONS: AcomA = anterior communicating artery; PcomA = posterior communicating artery

Endovascular coiling of cerebral aneurysms is widely performed, with continued improvement in related techniques and devices. However, an important problem with endovascular coiling is recanalization after coil embolization. Coiled aneurysms that show major recanalization require additional coiling to prevent rupture. The incidence of recanalization after coiling ranges from 6.1% to 33.6%, and the reported risk factors for recanalization include aneurysm morphologic features, coil compaction and/or migration, and various endovascular embolization techniques.

Although recanalization after coil embolization has been previously evaluated, there are insufficient data on the way in which recanalized aneurysms rupture. Thus, in the present study, we describe our experience with the rupture of aneurysms that have recanalized after coil embolization in Kobe City Medical Center General Hospital and discuss the type of recanalized aneurysms at greatest risk of rupture and the appropriate timing of follow-up angiography and retreatment.

MATERIALS AND METHODS

Patient and Aneurysm Characteristics

Between January 2009 and December 2017, coil embolization was performed in a total of 426 unruptured aneurysms and 169 ruptured aneurysms in our institution, excluding fusiform aneurysms, dissecting aneurysms, infectious aneurysms, and aneurysms combined with arteriovenous malformations or Moyamoya disease. Aneurysms that underwent parent artery occlusion and that were treated with an off-label stent used in a clinical trial and a stent in the acute phase of rupture, recanalized aneurysms previously
treated with coiling, and aneurysms that reruptured in the acute phase (<1 month after the initial rupture) were excluded from the present study. Recanalization after coil embolization occurred in 38 of 426 unruptured aneurysms (unruptured group) and 37 of 169 ruptured aneurysms (ruptured group). Table 1 summarizes the characteristics of the patients and recanalized aneurysms, endovascular procedures, and follow-up term in these 2 groups compared with all patients and aneurysms.

The present investigation was approved by our local ethics committee (institutional review board of Kobe City Medical Center General Hospital). The study was conducted in accordance with the 1964 Declaration of Helsinki (as revised in Fortaleza, Brazil, October 2013).

**Perioperative Antiplatelet Management**

Patients undergoing coil embolization for unruptured aneurysms were generally administered dual-antiplatelet therapy with 100 mg of aspirin and 75 mg of clopidogrel daily for at least 7 days before the procedure. Either aspirin or clopidogrel was decreased gradually from 6 months after the procedure and discontinued after at least 1 year. Single-antiplatelet therapy was generally discontinued after 1 or 2 years. When stent-assisted coil embolization was performed, antiplatelet activity was assessed using aspirin and P2Y12 assays (VerifyNow; Accumetrics) on the day before the procedure. When the aspirin reaction unit measurements were >550, the aspirin dose was increased to 200 mg daily. When the P2Y12 reaction unit measurements were >220, a loading dose of 20 mg of prasugrel was given, and the dose was then decreased to 3.75 mg from the following day. Prasugrel was decreased gradually from 6 months after the procedure. Although prasugrel and either aspirin or clopidogrel were discontinued after 1 year, single-antiplatelet therapy was continued as a life-long treatment. Patients undergoing coil embolization for ruptured aneurysms were generally administered single-antiplatelet therapy with aspirin or cilostazol. Single-antiplatelet therapy was discontinued depending on the DSA findings.

**Endovascular Procedures**

Preoperative DSA findings were used to determine the best approach and technique for the endovascular procedure (ie, simple technique, double-catheter technique, balloon-assisted coil embolization, or stent-assisted coil embolization). The procedure was performed by a neurovascular team, including a board-certified neuroendovascular surgeon. The procedure for unruptured aneurysms was generally performed with the patient under local anesthesia, because neurologic evaluation can be performed during the procedure. The procedure for ruptured aneurysms was performed with the patient under general anesthesia. All patients received heparin after an arterial puncture to maintain an activated clotting time of approximately 300 seconds during the procedure for unruptured aneurysms and approximately 250 seconds for ruptured aneurysms. The volume embolization ratio was measured on DSA after the procedure, and the angiographic findings were evaluated.

---

**Table 1: Patient and aneurysm characteristics, volume embolization ratio, use of stents, and follow-up term**

<table>
<thead>
<tr>
<th>Patient characteristic</th>
<th>Unruptured Group</th>
<th>Recanalized Aneurysms</th>
<th>Ruptured Group</th>
<th>Recanalized Aneurysms</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Aneurysms (n = 426)</td>
<td>All Aneurysms (n = 38)</td>
<td>All Aneurysms (n = 169)</td>
<td>All Aneurysms (n = 37)</td>
<td></td>
</tr>
<tr>
<td>Patient characteristic</td>
<td>Ruptured Group</td>
<td>Recanalized Aneurysms</td>
<td>Ruptured Group</td>
<td>Recanalized Aneurysms</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>63.0 (52.0–70.0)</td>
<td>65.0 (49.8–71.3)</td>
<td>66.0 (52.0–77.5)</td>
<td>59.0 (44.5–76.5)</td>
</tr>
<tr>
<td>Male sex</td>
<td>105 (24.6%)</td>
<td>6 (15.8%)</td>
<td>56 (33.1%)</td>
<td>10 (27.0%)</td>
</tr>
<tr>
<td>Aneurysm location</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICA-cavernous</td>
<td>12 (2.8%)</td>
<td>2 (5.3%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>ICA-paraclinoid</td>
<td>188 (44.1%)</td>
<td>9 (23.7%)</td>
<td>5 (3.0%)</td>
<td>2 (5.4%)</td>
</tr>
<tr>
<td>ICA-PcomA</td>
<td>49 (11.5%)</td>
<td>9 (23.7%)</td>
<td>55 (32.5%)</td>
<td>16 (43.2%)</td>
</tr>
<tr>
<td>ICA-AchoA</td>
<td>10 (2.3%)</td>
<td>0 (0.0%)</td>
<td>13 (7.7%)</td>
<td>4 (10.8%)</td>
</tr>
<tr>
<td>ICA-bifurcation</td>
<td>11 (2.6%)</td>
<td>2 (5.3%)</td>
<td>3 (1.8%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>MCA</td>
<td>6 (1.4%)</td>
<td>0 (0.0%)</td>
<td>4 (2.4%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>AcomA</td>
<td>66 (15.5%)</td>
<td>1 (2.6%)</td>
<td>57 (33.7%)</td>
<td>11 (29.7%)</td>
</tr>
<tr>
<td>VA</td>
<td>15 (3.5%)</td>
<td>4 (10.5%)</td>
<td>8 (4.7%)</td>
<td>1 (2.7%)</td>
</tr>
<tr>
<td>BA-SCA</td>
<td>16 (3.8%)</td>
<td>2 (5.3%)</td>
<td>1 (2.2%)</td>
<td>2 (2.7%)</td>
</tr>
<tr>
<td>BA-bifurcation</td>
<td>41 (9.6%)</td>
<td>9 (23.7%)</td>
<td>12 (7.1%)</td>
<td>2 (5.4%)</td>
</tr>
<tr>
<td>Others</td>
<td>12 (2.8%)</td>
<td>0 (0.0%)</td>
<td>10 (5.9%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Aneurysm size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dome</td>
<td>6.2 (4.9–8.4)</td>
<td>10.1 (7.0–14.1)</td>
<td>6.5 (5.0–9.3)</td>
<td>7.7 (5.4–13.0)</td>
</tr>
<tr>
<td>Neck</td>
<td>3.8 (2.9–5.1)</td>
<td>5.9 (4.7–7.7)</td>
<td>3.2 (2.4–4.2)</td>
<td>4.0 (2.7–5.9)</td>
</tr>
<tr>
<td>Endovascular procedures</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VER</td>
<td>26.9 (21.9–32.8)</td>
<td>23.0 (18.0–29.2)</td>
<td>25.8 (19.8–33.2)</td>
<td>23.7 (18.9–29.5)</td>
</tr>
<tr>
<td>Without stent</td>
<td>240 (56.3%)</td>
<td>19 (50.0%)</td>
<td>169 (100.0%)</td>
<td>37 (100.0%)</td>
</tr>
<tr>
<td>With stent</td>
<td>186 (43.7%)</td>
<td>19 (50.0%)</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
</tr>
<tr>
<td>Follow-up Duration (mo)</td>
<td>58.5 (24.0–86.0)</td>
<td>74.0 (30.5–96.3)</td>
<td>18.0 (3.0–48.0)</td>
<td>28.0 (10.0–56.0)</td>
</tr>
</tbody>
</table>

*Note: AchoA indicates anterior choroidal artery; VA, vertebral artery; BA, basilar artery; SCA, superior cerebellar artery; VER, volume embolization ratio.*

---

Evaluation of Coiled Aneurysms

The Raymond-Roy occlusion classification is the standard for evaluating aneurysm occlusion, while the Modified Raymond-Roy classification was recently reported to be a predictor of progressive occlusion or recurrence (Fig 1). In the Modified Raymond-Roy classification, complete obliteration is defined as class I, a residual neck is defined as class II, and a residual aneurysm is defined as class III. Class III is further separated into class IIIa (contrast within the coil interstices) and IIIb (contrast along the aneurysm wall). The present study used the Modified Raymond-Roy classification to evaluate aneurysm occlusion and recanalization. Any further filling of the aneurysm sac across time on follow-up DSA was termed "recanalization." The appearance of new blebs was distinguished from an increase in the diameter of the coiled aneurysm. Except for the appearance of new blebs, an increase in the diameter of the coiled aneurysm was termed "aneurysm growth" (Fig 1).

Two observers, who were board-certified neuroendovascular surgeons, evaluated aneurysm occlusion and recanalization using the Modified Raymond-Roy classification, and 1 investigator evaluated these parameters twice to verify inter- and intrarater reliability.

Follow-Up

In the unruptured group, follow-up DSA was scheduled at 6 months after treatment, and follow-up MRA (using ≥1.5T magnet strength) was scheduled yearly. In the ruptured group, follow-up DSA was generally scheduled for 1 week after treatment to evaluate vasospasm and for 1 month after treatment to detect early recanalization. From 6 months after treatment, follow-up evaluations were performed as described for the unruptured group. In principle, the diagnosis of recanalization was based only on DSA findings. If recanalization was suspected on follow-up MRA, DSA was performed. When recanalization was identified on follow-up DSA, the patient was scheduled to undergo retreatment or an earlier follow-up DSA.

Statistical Analysis

Statistical analyses were performed using statistical software (JMP software, Version 14; SAS Institute). Sex, aneurysm location, use of a stent, type of recanalization, aneurysm growth, and the appearance of new blebs were evaluated using the $\chi^2$ test or the Fisher exact test. Age, aneurysm dome diameter, aneurysm neck size, and volume embolization ratio were analyzed using the Mann-Whitney U test. $P<.05$ was considered statistically significant.

RESULTS

Outcomes of Aneurysms and Patients

Outcomes of aneurysms and patients are summarized in Table 2. In the unruptured group, the median follow-up term of recanalized aneurysms was 74.0 months. The type of recanalization on follow-up DSA, the timing of retreatment and rupture, and the type of recanalization immediately before retreatment and rupture are shown in Table 3. From a total of 38 recanalized aneurysms, the type of recanalization on final DSA was classified as class II in 3 (7.9%) aneurysms, class IIIa in 12 (31.6%), and class IIIb in 23 (60.5%). On follow-up DSA, aneurysm growth was identified in 12 (31.6%) aneurysms, and the appearance of new blebs, in 2 (5.3%) aneurysms. Retreatment was performed in 18 (47.4%) aneurysms,
while 20 (52.6%) were untreated. The 2 aneurysms with the appearance of new blebs were retreated. Four of the 20 untreated aneurysms ruptured, 0.94% of the total coiled aneurysms ruptured, and the rupture risk was 0.20% per year. Thirty-one (91.2%) of 34 patients without a ruptured recanalized aneurysm had a final mRS of 0, while 2 (50.0%) of 4 patients with a ruptured recanalized aneurysm had a final mRS of 0–2.

In the ruptured group, the median follow-up term of recanalized aneurysms was 28.0 months. The type of recanalization on follow-up DSA, the timing of retreatment and rupture, and the type of recanalization immediately before retreatment and rupture are shown in Table 4. From a total of 37 recanalized aneurysms, the type of recanalization on final DSA was classified as class II in 7 (18.9%) aneurysms, class IIIa in 7 (18.9%), and class IIIb in 23 (62.2%). On follow-up DSA, aneurysm growth was identified in 3 (8.1%) aneurysms, and new blebs appeared in 1 (2.7%) aneurysm. Retreatment was performed in 16 (43.2%) aneurysms, and new blebs appeared in 1 (2.7%) aneurysm.

Factors Associated with Rupture after Coil Embolization

Factors associated with rupture after coil embolization are summarized in Tables 5 and 6 for the unruptured and ruptured groups,
Table 5: Factors associated with rupture after coil embolization in untreated recanalized aneurysms in the unruptured groupa

<table>
<thead>
<tr>
<th>Untreated Recanalized Aneurysms (n = 20)</th>
<th>Unruptured (n = 16)</th>
<th>Ruptured (n = 4)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>67.0 (60.3–72.8)</td>
<td>63.0 (55.8–70.3)</td>
<td>.570</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1 (6.3%)</td>
<td>1 (25.0%)</td>
<td>.264</td>
</tr>
<tr>
<td>Female</td>
<td>15 (93.7%)</td>
<td>3 (75.0%)</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICA</td>
<td>5 (31.2%)</td>
<td>1 (25.0%)</td>
<td>.937</td>
</tr>
<tr>
<td>ICA-PcomA</td>
<td>4 (25.0%)</td>
<td>1 (25.0%)</td>
<td></td>
</tr>
<tr>
<td>AcomA</td>
<td>1 (6.3%)</td>
<td>0 (0.0%)</td>
<td></td>
</tr>
<tr>
<td>Posterior</td>
<td>6 (37.5%)</td>
<td>2 (50.0%)</td>
<td></td>
</tr>
<tr>
<td>Dome</td>
<td>9.2 (7.6–13.4)</td>
<td>9.9 (6.6–11.1)</td>
<td>.813</td>
</tr>
<tr>
<td>Neck</td>
<td>5.5 (4.4–7.0)</td>
<td>4.5 (3.2–8.8)</td>
<td>.741</td>
</tr>
<tr>
<td>VER</td>
<td>22.1 (13.9–26.1)</td>
<td>25.1 (18.5–37.3)</td>
<td>.321</td>
</tr>
<tr>
<td>Use of stent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>9 (56.3%)</td>
<td>2 (50.0%)</td>
<td>.822</td>
</tr>
<tr>
<td>Yes</td>
<td>7 (43.7%)</td>
<td>2 (50.0%)</td>
<td></td>
</tr>
<tr>
<td>Type of recanalization on final DSA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II + IIIa</td>
<td>10 (62.5%)</td>
<td>0 (0.0%)</td>
<td>.025b</td>
</tr>
<tr>
<td>IIb</td>
<td>6 (37.5%)</td>
<td>4 (100.0%)</td>
<td></td>
</tr>
<tr>
<td>Aneurysm growth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>13 (81.3%)</td>
<td>3 (75.0%)</td>
<td>.780</td>
</tr>
<tr>
<td>Yes</td>
<td>3 (18.7%)</td>
<td>1 (25.0%)</td>
<td></td>
</tr>
</tbody>
</table>

Note:—ICA indicates ICA other than ICA-Pcom; VA, vertebral artery; BA, basilar artery; SCA, superior cerebellar artery; VER, volume embolization ratio; Posterior, posterior circulation including the VA, BA-SCA, and BA-bifurcation.

aData are recanalized aneurysms in the unruptured group (n = 38). Twenty of 38 recanalized aneurysms were untreated. Among 20 untreated recanalized aneurysms, 16 aneurysms unruptured and 4 aneurysms ruptured.

bStatistical significance.

Table 6: Factors associated with rupture after coil embolization in untreated recanalized aneurysms in the ruptured groupa

<table>
<thead>
<tr>
<th>Untreated Recanalized Aneurysms (n = 21)</th>
<th>Unruptured (n = 17)</th>
<th>Ruptured (n = 4)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>62.0 (44.5–78.0)</td>
<td>55.5 (47.5–75.5)</td>
<td>.929</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>6 (35.3%)</td>
<td>1 (25.0%)</td>
<td>.694</td>
</tr>
<tr>
<td>Female</td>
<td>11 (64.7%)</td>
<td>3 (75.0%)</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICA</td>
<td>3 (17.7%)</td>
<td>0 (0.0%)</td>
<td>.628</td>
</tr>
<tr>
<td>ICA-PcomA</td>
<td>9 (52.9%)</td>
<td>2 (50.0%)</td>
<td></td>
</tr>
<tr>
<td>AcomA</td>
<td>4 (23.5%)</td>
<td>2 (50.0%)</td>
<td></td>
</tr>
<tr>
<td>Posterior</td>
<td>1 (5.9%)</td>
<td>0 (0.0%)</td>
<td></td>
</tr>
<tr>
<td>Dome</td>
<td>6.2 (5.2–8.7)</td>
<td>9.4 (5.3–14.4)</td>
<td>.347</td>
</tr>
<tr>
<td>Neck</td>
<td>3.2 (2.5–4.6)</td>
<td>5.2 (3.6–9.1)</td>
<td>.066</td>
</tr>
<tr>
<td>VER</td>
<td>24.9 (21.3–29.5)</td>
<td>25.5 (16.7–28.5)</td>
<td>.741</td>
</tr>
<tr>
<td>Use of stent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>17 (100.0%)</td>
<td>4 (100.0%)</td>
<td>–</td>
</tr>
<tr>
<td>Yes</td>
<td>0 (0.0%)</td>
<td>0 (0.0%)</td>
<td></td>
</tr>
<tr>
<td>Type of recanalization on final DSA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II + IIIa</td>
<td>11 (64.7%)</td>
<td>0 (0.0%)</td>
<td>.020b</td>
</tr>
<tr>
<td>IIb</td>
<td>6 (35.3%)</td>
<td>4 (100.0%)</td>
<td></td>
</tr>
<tr>
<td>Aneurysm growth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>16 (94.1%)</td>
<td>4 (100.0%)</td>
<td>.619</td>
</tr>
<tr>
<td>Yes</td>
<td>1 (5.9%)</td>
<td>0 (0.0%)</td>
<td></td>
</tr>
</tbody>
</table>

Note:— indicates not available.

aData are recanalized aneurysms in the ruptured group (n = 37). Twenty-one of 37 recanalized aneurysms were untreated. Among 21 untreated recanalized aneurysms, 17 aneurysms unruptured and 4 aneurysms ruptured.

bStatistical significance.

DISCUSSION

The cases of ruptured recanalized aneurysms are summarized in the Online Table. The DSA findings of ruptured recanalized aneurysms in the untreated and ruptured groups are shown in Figs 2 and 3, respectively. In patient 4 in the untreated group, DSA could not be performed when the recanalized aneurysm ruptured. Thus, the MRA findings are shown instead. In the ruptured group, 4 recanalized aneurysms ruptured. In patient 1 with a large basilar artery–bifurcation aneurysm, the type of recanalization changed from class II to IIIb, and the aneurysm ruptured at 8 months after coil embolization. In patient 2 with an ICA–posterior communicating artery (PcomA) aneurysm, patient 3 with an ICA–paracinald aneurysm, and patient 4 with a large basilar artery–bifurcation aneurysm, the type of recanalization was class IIIb, and the aneurysms ruptured with a new bleb formation at 76, 1, 2, and 24 months after coil embolization, respectively. All 4 patients with ruptured recanalized aneurysms were receiving continued antiplatelet therapy at the time of rupture.

In the ruptured group, 4 recanalized aneurysms ruptured. In patients 1 and 2, both with an anterior communicating artery (AcomA) aneurysm, the type of recanalization changed from class II to IIIb, and both aneurysms ruptured with formation of a new bleb at 1 month after coil embolization. In patients 3 and 4, both with an ICA–PcomA aneurysm, the type of recanalization was class IIIb, and the aneurysms ruptured at 29 and 7 months after coil embolization, respectively. All 4 patients with ruptured recanalized aneurysms were receiving continued antiplatelet therapy at the time of rupture. Recanalized aneurysms in the ruptured group tended to rupture earlier than those in the untreated group.

The Epidemiology, natural history, management options, and familial screening of unruptured aneurysms are well-established. Several studies have also reported treatment outcomes for ruptured
Recanalization is a risk factor for rupture of coiled aneurysms because major recanalization can result in aneurysm rupture. However, not all recanalized aneurysms rupture, and the risk of rupture seems to depend on the type of recanalization. The Modified Raymond-Roy classification is an indicator of how aneurysms are occluded, and it can be used to predict progressive occlusion or recurrence. Because the type of occlusion is a predictor of recanalization, the types of occlusion and recanalization may be predictors of rupture of coiled aneurysms. In the present study, although the aneurysm neck diameter of the ruptured recanalized aneurysms tended to be larger than that of the unruptured recanalized aneurysms in the ruptured group (P = .020), the type of recanalization was the only significant factor in both the unruptured (P = .025) and ruptured (P = .050) groups. There was no rupture of coiled aneurysms with class II and IIIa recanalization (ie, no contrast along the aneurysm wall). However, 4 recanalized aneurysms with class IIIb recanalization (ie, contrast along the aneurysm wall) ruptured in each group. Although there were few ruptured aneurysms and statistical evaluation was insufficient, the present study has important information regarding how the occlusion and recanalization of coiled aneurysms affected the risk of rupture.

A previous study reported that the risk factors for early recanalization of coiled aneurysms within 6 months included posterior circulation aneurysms, SAH presentation, a second coiling procedure for recanalization, and an aneurysm of ≥7 mm. In addition, long-term follow-up is recommended even for aneurysms of 4–7 mm because of the risk of late recanalization. In the present study, recanalized aneurysms in the ruptured group tended to rupture earlier than those in the unruptured group. In our institution, on the basis of our experience, follow-up DSA was scheduled at 6 months after treatment for unruptured aneurysms; for ruptured aneurysms, follow-up DSA was scheduled at 1 week after treatment to evaluate vasospasm and at 1 month after treatment to detect early recanalization. The recanalized aneurysms in patients 1 and 2 in the ruptured group ruptured at 1 month after treatment (just before the scheduled follow-up DSA).

In the present study, most retreatments for recanalization in the ruptured group were performed at 1–6 months after the initial treatment. Follow-up DSA within 1 month can prevent the rupture
of some recanalized aneurysms. For unruptured aneurysms, follow-up DSA at 6 months after coil embolization is considered reasonable. Compared with the ruptured group, the timing of coiled aneurysm rupture in the unruptured group tended to be later in the present study. Most retreatments for recanalization in the unruptured group were performed at 6 months to 1 year after the initial treatment. Although retreatment at this time can prevent the rupture of recanalized aneurysms, there are rare exceptions that rupture shortly after coil embolization: For example, the unruptured aneurysm after coil embolization in patient 3 ruptured 2 months after treatment. In addition, there were cases in both the unruptured and ruptured groups in which the coiled aneurysms ruptured during long-term follow-up. Coiled aneurysms with class IIIb recanalization should undergo early retreatment because of the high risk of rupture. Although coiled aneurysms with class II or IIIa recanalization should also be retreated if possible, these aneurysms can be treated with conservative therapy because of the low risk of rupture. However, class II or IIIa recanalization can progress into class IIIb recanalization. Thus, follow-up examinations are essential.

Periprocedural antiplatelet therapy was associated with a low risk of symptomatic thromboembolic events after coil embolization for aneurysms. In the present study, all patients with ruptured recanalized aneurysms were receiving continued antiplatelet therapy at the time of rupture. In particular, long-term dual-antiplatelet therapy was continued, considering the thromboembolic event in patient 4 in the unruptured group. Long-term single-antiplatelet therapy was also continued in patient 2 in the unruptured group and in patient 3 in the ruptured group. Although it is unclear whether antiplatelet therapy in the present cases could influence the rupture of recanalized aneurysms, indiscriminate continuation of antiplatelet therapy should be avoided.

The risk of rupture for aneurysms with growth is 12 times higher than that for aneurysms without growth. In the present study, although some coiled aneurysms showed aneurysm growth and the appearance of new blebs, there were no significant differences in aneurysm growth in the unruptured (P = .780) and ruptured (P = .619) groups. The appearance of new blebs was also not a significant risk factor for rupture of recanalized aneurysms. However, all aneurysms with the appearance of new blebs on follow-up DSA were retreated, and the appearance of new blebs was detected in 3 of 4 ruptured recanalized aneurysms in the unruptured group and in 2 of 4 ruptured recanalized aneurysms in the ruptured group at the time of rupture. Although it is difficult to detect the appearance of new blebs on follow-up DSA because aneurysms with formation of new blebs can rupture early, the appearance of new blebs can be a risk factor for rupture of recanalized aneurysms and thus must not be overlooked. Additionally, the appearance of new blebs both on follow-up DSA and at the time of rupture was identified only in coiled aneurysms with class IIIb recanalization; this finding may be because blood flow along the aneurysm wall can form new blebs. In recanalized aneurysms, the type of recanalization can be important because class IIIb recanalization can result in the appearance of new blebs, which may be a risk factor for early rupture of recanalized aneurysms.

**Study Limitations**

The present study was a nonrandomized retrospective observation study, though the series comprised consecutive cases. In addition,
interpretation of the radiographic data was performed in a non-blinded manner. Although there were 4 cases of ruptured recanalized aneurysms in both the untreated and ruptured groups, the number of ruptured recanalized aneurysms was small. Thus, statistical multivariate analysis could not be performed, though multiple factors can be related to rupture of recanalized aneurysms. Because the ruptured group included many severe cases that could not be followed up at our institution, there was also a lack of cases with sufficient data. Furthermore, because retreatment was generally performed in cases in which retreatment for recanalized aneurysms was possible and the patient provided consent, the true course of recanalized aneurysms leading to rupture without retreatment remains unclear.

CONCLUSIONS

The present study determined the type of recanalized aneurysm at risk of rupture on the basis of our experience. The types of aneurysm occlusion and recanalization are predictors of the rupture of coiled aneurysms. Coiled aneurysms with class IIIb recanalization should undergo early retreatment because of the high risk of rupture.

ACKNOWLEDGMENTS

We thank Kelly Zammit, BVSc, from Edanz Editing (www.edanzediting.com/ac), for editing a draft of the manuscript.


REFERENCES

Signal of Carotid Intraplaque Hemorrhage on MR T1-Weighted Imaging: Association with Acute Cerebral Infarct

D. Yang, Y. Liu, Y. Han, D. Li, W. Wang, R. Li, C. Yuan, and X. Zhao

BACKGROUND AND PURPOSE: Identifying the mere presence of carotid intraplaque hemorrhage would be insufficient to accurately discriminate the presence of acute cerebral infarct. We aimed to investigate the association between signal intensity ratios of carotid intraplaque hemorrhage on T1-weighted MR imaging and acute cerebral infarct in patients with hemorrhagic carotid plaques using MR vessel wall imaging.

MATERIALS AND METHODS: Symptomatic patients with carotid intraplaque hemorrhage were included. The signal intensity ratios of carotid intraplaque hemorrhage against muscle on T1-weighted, TOF, and MPRAGE images were measured. The acute cerebral infarct was determined on the hemisphere ipsilateral to the carotid intraplaque hemorrhage. The association between signal intensity ratios of carotid intraplaque hemorrhage and acute cerebral infarct was analyzed.

RESULTS: Of 109 included patients (mean, 66.8 ± 9.9 years of age; 96 men), 40 (36.7%) had acute cerebral infarct. Patients with acute cerebral infarct had significantly higher signal intensity ratios of carotid intraplaque hemorrhage on T1-weighted images than those without (Median, 144; 25–75 Percentiles, 114-182 versus Median, 127; 25–75 Percentiles, 106–155, P = .022). Logistic regression analysis revealed that the signal intensity ratio of carotid intraplaque hemorrhage on T1-weighted images was significantly associated with acute cerebral infarct before (OR, 4.08; 95% CI, 1.34–12.40; P = .013) and after (OR, 3.34; 95% CI, 1.08–10.31; P = .036) adjustment for clinical confounding factors. However, this association was not significant when further adjusted for occlusion of the carotid artery (P = .058) and volumes of intraplaque hemorrhage and lipid-rich necrotic core (P = .458).

CONCLUSIONS: The signal intensity ratio of carotid intraplaque hemorrhage on T1-weighted images is associated with acute cerebral infarct in symptomatic patients with carotid hemorrhagic plaques. This association is independent of traditional risk factors but not of the size of plaque composition. The possibility of applying T1 signals of carotid intraplaque hemorrhage to predict subsequent cerebrovascular ischemic events needs to be prospectively verified.

ABBREVIATIONS: ACI = acute cerebral infarct; IPH = intraplaque hemorrhage; LRNC = lipid-rich necrotic core; SIR = signal intensity ratio
Sun et al.\textsuperscript{11} suggested that carotid IPH could persist for years with no clinical symptoms. However, the presence of IPH could accelerate the progression of atherosclerotic plaque and increase the risk of plaque rupture.\textsuperscript{12} Therefore, identifying the mere presence of IPH would be insufficient to precisely predict the presence of cerebrovascular ischemic events.

T1-weighted MR images have commonly been used to detect IPH due to the degradation of hemorrhage into methemoglobin, which leads to T1-shortening\textsuperscript{13} and correspondingly causes high signal intensity on T1-weighted MR images. Previous studies have indicated the importance of the signal intensity ratios (SIRs) of carotid plaques on different MR images, which could be taken into account for factors beyond the presence of IPH.\textsuperscript{14,15} A recent study by Wang et al.\textsuperscript{16} indicated that the SIRs of carotid IPH on different T1-weighted MR images were not distinct.

The purpose of this study was to investigate the association between the SIRs of carotid IPH on different T1-weighted MR images and the presence of ipsilateral ACI lesions in patients with hemorrhagic carotid plaques using multicontrast MR vessel wall imaging.

### MATERIALS AND METHODS

#### Patients

All patients were enrolled from a cross-sectional, multicenter study of Chinese Atherosclerosis Risk Evaluation (CARE-II), which aimed to assess the prevalence and characteristics of carotid high-risk atherosclerotic plaques in Chinese patients with recent ischemic stroke or transient ischemic attack using multicontrast MR vessel wall imaging. The study protocol of the CARE-II study has been reported previously.\textsuperscript{17} The inclusion criteria are as follows: 1) 18–80 years of age, 2) a recent stroke or transient ischemic attack (<2 weeks after the onset of symptoms), 3) atherosclerotic plaques in at least one carotid artery determined by B-mode sonography, and 4) carotid plaques with IPH confirmed by MR vessel wall imaging. Patients with the following conditions were excluded from this study: 1) cardioembolic or hemorrhagic stroke, 2) a history of radiation therapy to the neck, or 3) contraindications to MR imaging. All patients underwent carotid multicontrast MR vessel wall imaging and clinically routine brain MR imaging. The demographics and clinical characteristics of all patients including age, sex, body mass index, history of smoking, hypertension, diabetes, hyperlipidemia, statin use, antihypertension medication use, coronary heart disease, and lipid levels were acquired from the medical records. The study protocol was approved by the institutional review board of Tsinghua University and local institutions, and a written consent form was obtained from each subject.

#### Carotid and Brain MR Imaging

All patients underwent carotid and brain MR imaging on 3T MR scanners (Achieva TX; Philips Healthcare, Best, the Netherlands) with 8-channel phased array carotid coils and head coils. A standardized multisequence protocol was performed for the carotid artery by acquiring TOF, T1-weighted, and MPRAGE sequences. Carotid artery imaging was centered on the carotid artery bifurcation on the symptomatic side. Standard brain MR images were acquired using T1-weighted and DWI sequences. The imaging parameters are detailed in Table 1.

#### MR Image Analysis

All the carotid images were reviewed in consensus by 2 experienced radiologists who had >3 years of experience in neurovascular imaging, blinded to clinical information using customized software (CASCADE; University of Washington, Seattle, Washington).\textsuperscript{18} All carotid images were registered with the landmark of carotid bifurcation. The lumen and outer wall boundaries were manually traced. The lumen, wall, total vessel areas, and the mean wall thickness on each axial image were measured. Normalized wall index, defined as wall area divided by the total vessel area, was calculated. The presence or absence of calcification, lipid-rich necrotic core (LRNC), IPH, and fibrous cap rupture at each axial location was identified according to the published criteria, which were validated by histology.\textsuperscript{19} In particular, the identification of carotid IPH was mainly based on a pattern of hyperintense signal intensity on MPRAGE images (1.5 times the signal intensity compared with adjacent muscle).\textsuperscript{20} The volume of each plaque component, including calcification, LRNC, and IPH, and the corresponding percentage volume (volume of each carotid plaque component divided by plaque volume) were calculated. The volume of carotid plaque was calculated as the sum of the plaque area within slices multiplied by slice thickness. The length of carotid IPH was defined as the...
number of slices with carotid IPH on MPRAGE images multiplied by slice thickness. The ROIs were drawn to enclose the IPH region and the adjacent muscle, respectively, on each axial MPRAGE image with IPH and were automatically mapped to other MR imaging sequences (On-line Figure). The maximum signal intensity of carotid IPH and the mean signal intensity of the adjacent sternocleidomastoid muscle on T1-weighted, TOF, and MPRAGE images were recorded.

The SIR of carotid IPH against muscle on different imaging sequences was calculated by dividing the maximum signal intensity of the IPH by the mean signal intensity of the sternocleidomastoid muscle on each axial image.21 Luminal stenosis was measured by radiologists on TOF MR angiography using the NASCET criteria.22 In the present study, we included some subjects with total luminal occlusion. In analyzing these subjects, we usually reviewed only the slices without luminal occlusion. Brain MR images were evaluated by 2 experienced radiologists who had >3 years of experience in neurovascular imaging with consensus, blinded to clinical information and carotid MR images. The presence or absence of ACI lesions in the internal carotid artery territory was determined on the hemisphere ipsilateral to carotid IPH. ACI was defined as lesions that showed hyperintensity on DWI ($b = 1000$ s/mm$^2$), but iso- or hypointensity on T1-weighted images.

### Statistical Analysis

For patients with bilateral carotid IPHs, the MR imaging features of the carotid artery associated with symptoms were selected for final statistical analysis. The continuous variables were presented as mean ± SD or median and interquartile range, and the binary variables were expressed as count and percentage. The plaque features on carotid MR imaging were compared between patients with and without ACI using the independent $t$ test, Mann-Whitney $U$ test, $\chi^2$ test, or Fisher exact test. Univariate and multivariate logistic regressions were performed to calculate the OR and corresponding 95% CI of the carotid SIR of carotid IPH against muscle features on different sequences in discriminating the presence of ipsilateral ACI lesions before and after adjustment for confounding factors. $P < .05$ was considered statistically significant. All statistical analyses were performed using SPSS 16.0 (IBM, Armonk, New York).

### RESULTS

The flow chart of patient recruitment in this study is shown in Fig 1. In total, 201 patients with carotid IPH in at least one carotid artery were recruited in this study. Of the 201 patients, 92 patients were excluded due to the following reasons: 1) poor registration of multisequences ($n = 30$); 2) insufficient coverage for carotid images ($n = 18$); 3) missing DWI images ($n = 28$); and 4) missing carotid T1-weighted images ($n = 16$). Finally, 109 patients with carotid IPH were qualified for the final analysis.

### Clinical Characteristics

Of the 109 patients, 40 (36.7%) had ACI lesions in the ipsilateral hemispheres of carotid IPH, of which 28 (70%) had a recent stroke and 12 (30%) had a transient ischemic attack. The clinical

---

**FIG 1.** Flow chart of the study sample.

**Table 2: Clinical characteristics of study population**

<table>
<thead>
<tr>
<th></th>
<th>Mean ± SD or (No.) (%)</th>
<th>Patients with ACI ($n = 40$)</th>
<th>Patients without ACI ($n = 69$)</th>
<th>$P$ Value$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (yr)</strong></td>
<td>64.2 ± 10.2</td>
<td>68.3 ± 9.5</td>
<td>.035</td>
<td></td>
</tr>
<tr>
<td><strong>Sex, male</strong></td>
<td>37 (92.5)</td>
<td>59 (85.5)</td>
<td>.366</td>
<td></td>
</tr>
<tr>
<td><strong>Body mass index (kg/m²)</strong></td>
<td>25.1 ± 3.6</td>
<td>24.0 ± 2.7</td>
<td>.117</td>
<td></td>
</tr>
<tr>
<td><strong>History of smoking</strong></td>
<td>30 (75.0)</td>
<td>38 (55.1)</td>
<td>.038</td>
<td></td>
</tr>
<tr>
<td><strong>History of hypertension</strong></td>
<td>33 (82.5)</td>
<td>58 (84.1)</td>
<td>.833</td>
<td></td>
</tr>
<tr>
<td><strong>Systolic blood pressure (mm Hg)</strong></td>
<td>144.2 ± 20.9</td>
<td>143.9 ± 21.3</td>
<td>.840</td>
<td></td>
</tr>
<tr>
<td><strong>Diastolic blood pressure (mm Hg)</strong></td>
<td>85.2 ± 9.2</td>
<td>85.8 ± 14.1</td>
<td>.720</td>
<td></td>
</tr>
<tr>
<td><strong>History of hyperlipidemia</strong></td>
<td>35 (87.5)</td>
<td>49 (71.0)</td>
<td>.108</td>
<td></td>
</tr>
<tr>
<td><strong>LDL (mmol/L)</strong></td>
<td>2.7 ± 1.0</td>
<td>3.0 ± 1.0</td>
<td>.182</td>
<td></td>
</tr>
<tr>
<td><strong>HDL (mmol/L)</strong></td>
<td>1.1 ± 0.2</td>
<td>1.1 ± 0.2</td>
<td>.780</td>
<td></td>
</tr>
<tr>
<td><strong>TC (mmol/L)</strong></td>
<td>4.3 ± 1.2</td>
<td>4.7 ± 1.1</td>
<td>.161</td>
<td></td>
</tr>
<tr>
<td><strong>TG (mmol/L)</strong></td>
<td>1.9 ± 1.2</td>
<td>2.0 ± 1.1</td>
<td>.519</td>
<td></td>
</tr>
<tr>
<td><strong>History of diabetes mellitus</strong></td>
<td>15 (37.5)</td>
<td>25 (36.2)</td>
<td>.895</td>
<td></td>
</tr>
<tr>
<td><strong>History of coronary heart disease</strong></td>
<td>11 (27.5)</td>
<td>14 (20.3)</td>
<td>.388</td>
<td></td>
</tr>
<tr>
<td><strong>Statin use</strong></td>
<td>15 (37.5)</td>
<td>25 (36.2)</td>
<td>.895</td>
<td></td>
</tr>
<tr>
<td><strong>Antihypertension medication use</strong></td>
<td>28 (70.0)</td>
<td>49 (71.0)</td>
<td>.911</td>
<td></td>
</tr>
<tr>
<td><strong>TIA</strong></td>
<td>12 (30.0)</td>
<td>32 (46.4)</td>
<td>.108</td>
<td></td>
</tr>
<tr>
<td><strong>Recent stroke</strong></td>
<td>28 (70.0)</td>
<td>37 (53.6)</td>
<td>.108</td>
<td></td>
</tr>
</tbody>
</table>

$^a$ $P$ values were calculated by the Mann-Whitney $U$ test or Fisher exact test.

---

Note: -- LDL indicates low-density lipoprotein cholesterol; HDL, high-density lipoprotein cholesterol; TC, total cholesterol; TG, triglycerides; TIA, transient ischemia attack.
characteristics between patients with and without ipsilateral ACI lesions are presented in Table 2. Patients with ipsilateral ACI lesions were younger (mean, 64.2 ± 10.2 versus 68.3 ± 9.5 years of age, \( P = .035 \)) and had a higher prevalence of smoking (75.0% versus 55.1%, \( P = .038 \)) compared with those without. No significant differences were found in other clinical characteristics between patients with and without ipsilateral ACI.

**Plaque Morphology and Component Characteristics**

The comparison results on the carotid plaque morphological and compositional characteristics between patients with and without ipsilateral ACI lesions are shown in Table 3. Patients with ipsilateral ACI lesions had significantly greater mean wall area (48.3 ± 14.9 versus 41.5 ± 11.1 mm\(^2\), \( P = .028 \)), mean wall thickness (1.9 ± 0.6 versus 1.6 ± 0.4 mm, \( P = .010 \)), and a higher prevalence of total luminal occlusion of the carotid artery (52.5% versus 15.9%, \( P < .001 \)) than those without. There were no statistically significant differences in other carotid plaque burden measurements, including lumen area, total vessel area, mean normalized wall index and carotid plaque volume between the two groups (all \( P > .05 \)). Of all 109 patients, 93 (85.3%) had carotid calcification and 37 (33.9%) had carotid fibrous cap rupture. There were no significant differences in the prevalence of calcification (\( P = .526 \)) and fibrous cap rupture (\( P = .859 \)) between patients with and without ipsilateral ACI lesions. Patients with ipsilateral ACI lesions had a significantly larger volume of LRNC (312.7 \([174.1–625.3]\) versus 227.7 \([122.1–351.4]\) mm\(^3\), \( P = .038 \)), a larger percentage LRNC volume (29.4% versus 23.6%, \( P = .039 \)), and a larger IPH volume (93.2 \([31.4–201.2]\) versus 56.4 \([21.9–104.9]\) mm\(^3\), \( P = .016 \)) than those without. No significant differences were found in the percentage of IPH volume, volume of calcification, percentage of calcification volume, and carotid plaque length between the two groups (all, \( P > .05 \)).

**IPH Signal Characteristics**

Patients with ipsilateral ACI lesions were found to have significantly higher SIRs of carotid IPH (1.44 \([1.14–1.82]\) versus 1.27 \([1.06–1.55]\), \( P = .022 \)) on T1-weighted images than those without (Fig 2). No significant differences were found in the SIRs of carotid IPH on TOF (1.36 \([1.11–1.72]\) versus 1.19 \([0.92–1.58]\), \( P = .096 \)) and

![FIG 2. Boxplots of the SIR of carotid IPH on different T1-weighted MR imaging sequences in patients with and without ACI. Patients with ipsilateral ACI lesions had significantly higher SIRs of carotid IPH on T1-weighted images than those without.](https://www.ajnr.org)
MPRAGE (2.92 [2.23–4.11] versus 3.04 [2.37–3.99], P = .523) images between patients with and without ipsilateral ACI lesions (Fig 2).

**Association between SIRs of Carotid IPH and the Ipsilateral ACI**

Univariate regression analysis (Table 4) revealed that the SIRs of carotid IPH on T1-weighted images were significantly associated with the presence of ipsilateral ACI lesions (OR, 4.08; 95% CI, 1.34–12.40; P = .013). After adjustment for clinical confounding factors of age, sex, and history of smoking (model 1), this association remained statistically significant (OR, 3.34; 95% CI, 1.08–10.31; P = .036). After further adjustment for total luminal occlusion of the carotid artery (model 2), this association appeared marginally significant (OR, 3.12; 95% CI, 0.96–10.11; P = .058). However, after further adjustment for the volume of carotid IPH and LRNC (model 3), this association was not significant (OR, 1.57; 95% CI, 0.48–5.20; P = .458). Figure 3 shows that a patient with higher SIRs of carotid IPH against muscle on T1-weighted images had ACI lesions in the ipsilateral hemisphere.

FIG 3. Comparison of the SIR of carotid IPH on T1-weighted MR imaging sequences, including T1-weighted, TOF, and MPRAGE, between patients with ACI (4 images above, hyperintense on DWI as the arrow indicates) and without ACI (4 images below, no abnormality on DWI). It shows that the patient with the higher SIR on T1-weighted images had ACI lesions in the ipsilateral hemisphere.

Table 4: Association between signal intensity ratio of carotid IPH and the presence of ipsilateral ACI lesions

<table>
<thead>
<tr>
<th>SIRIPH-to-muscle on MR Image</th>
<th>Presence of ACI</th>
<th>Univariate Regression</th>
<th>Multivariate Model 1a</th>
<th>Multivariate Model 2a</th>
<th>Multivariate Model 3a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>95% CI</td>
<td>P&lt;sup&gt;b&lt;/sup&gt;</td>
<td>OR</td>
<td>95% CI</td>
</tr>
<tr>
<td>T1-weighted</td>
<td>4.08</td>
<td>1.34–12.40</td>
<td>.013</td>
<td>3.34</td>
<td>1.08–10.31</td>
</tr>
<tr>
<td>TOF</td>
<td>1.98</td>
<td>0.80–4.93</td>
<td>.141</td>
<td>2.54</td>
<td>0.94–6.82</td>
</tr>
<tr>
<td>MPRAGE</td>
<td>0.86</td>
<td>0.51–1.45</td>
<td>.561</td>
<td>0.85</td>
<td>0.48–1.48</td>
</tr>
</tbody>
</table>

Note: SIRIPH-to-muscle indicates the maximum signal intensity of carotid IPH/the mean signal intensity of the sternocleidomastoid muscle.

<sup>a</sup> Multivariate model 1 was adjusted for age, sex, and history of smoking; model 2 was further adjusted using model 1 and total luminal occlusion of the carotid artery. Model 3 was further adjusted using model 1 and the volume of carotid IPH and LRNC.

<sup>b</sup> Associations between patients with and without ipsilateral ACI lesions were assessed using the OR values.

DISCUSSION

The present study investigated the association between the SIRs of carotid IPH on different T1-weighted MR imaging sequences and ipsilateral ACI lesions in patients with carotid hemorrhagic plaques using multicontast MR vessel wall imaging. Although we found that patients with ipsilateral ACI lesions had significantly higher SIRs of carotid IPH on T1-weighted images than those without statistically, there was still an overlap in the SIRs of carotid IPH on T1-weighted images between the...
two groups. Only extremely low (range, 1.06–1.14) or high (range, 1.55–1.82) SIR values will allow a fairly accurate prediction of whether an ACI is present. Further studies are warranted combining SIRs of IPH with other metrics for predicting ACI, particularly for those IPHs with SIRs within this overlap. Furthermore, there was a significant association between the SIR of carotid IPH on T1-weighted images and the presence of ipsilateral ACI lesions, which was independent of traditional risk factors. After further adjustment for total luminal occlusion of the carotid artery, this association appeared marginally significant. This finding implies that the association between the SIRs of IPH on T1-weighted images and ACI might be influenced by the luminal occlusion. However, this association was not significant after further adjustment for the volume of carotid IPH and LRNC, suggesting that the differences of carotid SIRs on T1-weighted images between the two groups might be, in part, attributed to the different sizes of carotid IPH and LRNC. Our findings might provide additional imaging information in predicting the presence of cerebrovascular ischemic events beyond the mere presence of IPH.

In the present study, patients with ACI lesions were found to be younger than those without, which was consistent with the study by Turin et al., in which a lifetime risk of stroke was higher in younger age than in older age. Another recent study by Feigin et al. indicated that the global lifetime risk of stroke declined with age in patients older than 70 years, due to age-related competing risks from other diseases. In addition, the present study also demonstrated that patients with ipsilateral ACI lesions were more likely to smoke than those without, which was in line with previous studies. Wolf et al. and Donnan et al. have reported that cigarette smoking is an independent risk factor for cerebral infarct. Nicotine has been considered a key component that contributes to the influence of cigarette smoking on the cerebral infarct. Li et al. demonstrated that mitochondrial oxidative stress induced by nicotine may have important implications for the pathogenesis and pathophysiology of ischemic stroke.

In this study, we assessed the association between the SIRs of carotid IPH on different MR images and ipsilateral ACI lesions in symptomatic patients with carotid hemorrhagic plaques. We found that the SIRs of carotid IPH on T1-weighted images were associated with the presence of ipsilateral ACI lesions in symptomatic patients regardless of clinical confounding factors, including age, sex, and a history of smoking. However, this association was not significant after further adjustment for total luminal occlusion of the carotid artery and the volume of carotid IPH and LRNC. In contrast, previous studies that included patients both with and without carotid IPH also demonstrated the association between the T1-weighted signal of carotid plaque and cerebral ischemic lesions. Takemoto et al. reported that a higher T1-weighted signal intensity of carotid plaque was present more often in symptomatic patients with lesions positive for cerebral ischemia on DWI. In a study of patients with carotid artery stenosis scheduled for carotid endarterectomy or carotid artery stenting, Kurowski et al. documented that plaques with higher signal intensity on T1-weighted images were at higher risk of causing ischemic events. However, the mechanism for the association between the SIRs of carotid IPH on T1-weighted images and the presence of ACI remained unclear.

Previous studies have demonstrated that the signal intensity of IPH was associated with the age of hemorrhage within plaques, and high signal intensity of IPH on T1-weighted images may indicate a recent hemorrhage. In addition, the increasing size of IPH, which means a higher concentration of hemoglobin locally within the plaque, may also play a role in the occurrence of stronger SIRs of IPH on T1-weighted images. A study by Leung and Moody revealed that erythrocyte degradation caused intraplaque deposition of hemoglobin, which had a strong T1-shortening effect. In the present study, no significant association was found between SIRs of carotid IPH on T1-weighted images and the presence of ACI when adjusted for clinical confounding factors and carotid IPH and LRNC volume, suggesting that the difference in carotid SIRs on T1-weighted images between the two patient groups might be, in part, attributed to the different sizes of IPH and LRNC.

In addition, we found that patients with ACI lesions had significantly larger carotid IPH and LRNC, respectively, than those without. Takaya et al. reported that carotid plaques with a larger area of IPH (hazard ratio for a 10-mm² increase, 2.6; 95% CI, 1.4–4.6; P = .006) were significantly associated with subsequent ipsilateral carotid cerebrovascular events. However, the mechanism for this phenomenon remained uncertain. Previous studies indicated that the volume of IPH was significantly independently associated with minor fibrous cap disruption in carotid arteries (OR, 2.152; 95% CI, 1.241–3.730; P = .006). Another study by Zhao et al. indicated that patients with larger LRNC plaques might have larger cerebral infarcts. This phenomenon might be because that the large LRNC volume would accompany a larger fibrous cap area and increase the risk of rupture. Therefore, the size of carotid IPH and LRNC may influence, in part, the association between SIRs of carotid IPH and ACI. In this study, it is possible that carotid SIRs on T1-weighted imaging may combine both IPH and LRNC. Histologically, it is hypothesized that IPH occurs within the large lipid-rich necrotic core. Although the IPH was outlined on MPRAGE images, it is possible to include the LRNC in the ROI of IPH when the IPH is inhomogeneous. Our findings suggest that the SIR of IPH might be a surrogate for the size of the IPH or LRNC in predicting ACI.

An interesting issue we observed is that the SIRs of carotid IPH on MPRAGE images in patients with ipsilateral ACI lesions tended to be similar to those without (Median, 2.92; 25–75 Percentiles, 2.23–4.11 versus Median, 3.04; 25–75 Percentiles, 2.37–3.99, P = .523), a finding that was inconsistent with those in previous studies. A recent study by Wang et al. indicated that carotid IPH signals on MPRAGE images on the symptomatic side were significantly higher (mean, 5.8 ± 2.4 versus 4.7 ± 1.8, P = .004) than those on the asymptomatic side in symptomatic patients with bilateral carotid IPHs. There were a number of methodologic differences, including different sampling techniques and sample sizes, between the study by Wang et al. and the present study. For example, in our study, patients included were divided into two groups due to the presence or absence of ACI lesions on DWI, while in the study by Wang et
al, carotid plaques withIPH were divided into symptomatic and asymptomatic sides according to the recent symptoms, and the ACI lesions in that study on the symptomatic side were not distinct. More studies need to be performed to investigate the differences in this issue.

There are some limitations to this study. First, it only investigated the association between the SIRs of carotid IPH and ischemic stroke but did not analyze the difference in the age of IPH between patients with and without ACI lesions. The age of IPH may influence the association between the SIRs of IPH and ipsilateral ACI lesions. Future studies are warranted. Second, the current in-plane spatial resolution used in this study was 0.55 mm, which is limited to accurately quantifying small IPHs. Third, in this study, approximately one-fourth of the 201 patients were excluded due to “poor registration of multisequences,” suggesting that the method, at least in a certain number of subjects, is just not feasible. Future studies with larger samples and more improved technical methods are needed. Finally, although patients were recruited from a multicenter study, carotid IPH signal measurements on different sequences were performed on a single platform using a unified MR imaging protocol. Future studies using similar but not identical techniques are needed.

CONCLUSIONS

The signal intensity ratio of carotid IPH on T1-weighted images is associated with ACI in symptomatic patients with carotid hemorrhagic plaques. This association is independent of traditional risk factors but not of the size of plaque compositions. The possibility of applying T1 signals of carotid IPH to predict subsequent cerebrovascular ischemic events needs to be prospectively verified.


*Money paid to the institution.

REFERENCES

Distinguishing Recurrent Thyroid Cancer from Residual Nonmalignant Thyroid Tissue Using Multiphasic Multidetector CT

J.M. Debnam, N. Guha-Thakurta, J. Sun, W. Wei, M.E. Zafereo, M.E. Cabanillas, N.M. Buisson, and D. Schellingerhout

ABSTRACT

BACKGROUND AND PURPOSE: During thyroidectomy incomplete resection of the thyroid gland may occur. This complicates the imaging surveillance of these patients as residual thyroid needs to be distinguished from local recurrence. Therefore, the purpose of this study was to determine if multiphasic multi-detector computed tomography (4D-MDCT) can differentiate residual nonmalignant thyroid tissue and recurrent thyroid carcinoma after thyroidectomy.

MATERIALS AND METHODS: In this retrospective study, Hounsfield unit values on multiphasic multidetector CT in precontrast, arterial (25 seconds), venous (55 seconds), and delayed (85 seconds) phases were compared in 29 lesions of recurrent thyroid cancer, 29 with normal thyroid, and 29 with diseased thyroid (thyroiditis/multinodular thyroid). The comparison of Hounsfield unit values among lesion types by phase was performed using ANOVA. The performance of Hounsfield unit values to predict recurrence was evaluated by logistic regression and receiver operating characteristic analysis.

RESULTS: All 3 tissue types had near-parallel enhancement characteristics, with a wash-in–washout pattern. Statistically different Hounsfield unit density was noted between the recurrence (lowest Hounsfield unit), diseased (intermediate Hounsfield unit), and normal (highest Hounsfield unit) thyroid groups throughout all 4 phases (P < .001 for each group and in each phase). Dichotomized recurrence-versus-diseased/normal thyroid tissue with univariate logistic regression analysis demonstrated that the area under the receiver operating characteristic curve for differentiating benign from malignant thyroid for the various phases of enhancement was greatest in the precontrast phase at 0.983 (95% CI, 0.954–1), with a cutoff value of ≤ 62 (sensitivity/specificity, 0.966/0.983) followed by the arterial phase.

CONCLUSIONS: Recurrent thyroid carcinoma can be distinguished from residual nonmalignant thyroid tissue using multiphasic multidetector CT with high accuracy. The maximum information for discrimination is in the precontrast images, then the arterial phase. An optimal clinical protocol could be built from any number of phases but should include a precontrast phase.

ABBREVIATIONS: 4D-MDCT = multiphasic multidetector CT; ROC = receiver operating characteristic

In the United States, between 1975 and 2013, the incidence of thyroid cancer increased by 211% and the average annual increases in incidence and mortality rates were 3.6% and 1.1%, respectively. Papillary thyroid carcinoma, the least aggressive type of thyroid cancer, accounts for most new cases. In patients with thyroid cancer, thyroid surgery is the mainstay of treatment.

After the operation, approximately 20%–30% of patients experience local recurrence of soft-tissue or nodal metastases. Routine follow-up imaging to detect early recurrence includes CT and sonography to assess structural recurrence and thyroid-stimulating hormone and thyroglobulin levels for biochemical assessment.

During thyroidectomy, an area of difficulty for surgeons is the Berry ligament, which tethers the thyroid to the trachea, because the recurrent laryngeal nerve is generally intimately associated with this ligament. In inexperienced hands, incomplete resection of the thyroid at the Berry ligament may occur as a result of the surgeon’s attempt to avoid damaging this nerve. D’Andrea et al studied the rate of residual thyroid tissue after total thyroidectomy and found that 34 of 102 patients (33.3%) had “significant thyroid tissue remnants” in the thyroid bed.
The problem this creates in the imaging follow-up of these patients is that residual tissue needs to be correctly characterized as either benign or malignant. Accurate delineation of thyroid bed tissue as residual thyroid may avoid an unnecessary operation with complications including damage to the recurrent laryngeal nerve as well as the inherent risks of anesthesia. On the other hand, treatment of recurrent disease is necessary to avoid unfavorable outcomes including locoregional complications such as invasion of the trachea and esophagus and to avoid distant metastases.\textsuperscript{10}

Differentiation between these 2 entities is not always straightforward.\textsuperscript{7} An additional diagnostic challenge is superimposed if there is background benign thyroid disease in the residual thyroid tissue, such as thyroiditis or a multinodular goiter.\textsuperscript{11} To the best of our knowledge, no prior studies have addressed this issue using CT.

In this study, we address this gap in imaging knowledge and determine whether multiphasic multidetector CT (4D-MDCT) can be used to characterize residual thyroid tissue after an operation as either benign or malignant. 4D-MDCT involves high-resolution helical CT before and during the administration of a bolus of iodinated contrast at specific points in time to allow the changing distribution of contrast to be assessed. 4D-MDCT is in common clinical use for the localization of parathyroid adenomas, which show characteristic contrast handling attributes that allow accurate localization.\textsuperscript{12-17}

We hypothesized that 4D-MDCT can be used similarly to differentiate recurrent thyroid carcinoma from residual normal and diseased thyroid tissue. We also hypothesized that recurrent tumor and normal and diseased thyroid would have diagnostic patterns of contrast enhancement that allow them to be clearly discriminated.

**MATERIALS AND METHODS**

Our institutional review board approved this study and waived the requirement for informed consent. Data were acquired in compliance with all applicable Health Insurance Portability and Accountability Act regulations.

**Study Population**

We reviewed the medical records of 152 consecutive patients who underwent 4D-MDCT of the thyroid beds for localization of suspected recurrent thyroid carcinoma following total thyroidectomy. In these patients, recurrent lesions were suspected due to an elevated thyroglobulin level, surveillance CT findings, surveillance ultrasound findings, and/or the results of an ultrasound-guided fine-needle aspiration. For inclusion in this group (group 1), patients were required to have pathologically proved disease by either fine-needle aspiration or subsequent surgical resection. The 4D-MDCT studies in group 1 were performed between 48 days and 14 years 9 months (mean, 4 years 3 months) after thyroidectomy.

Seventy-five separate patients underwent 4D-MDCT for localization of a parathyroid adenoma. From these patients, we identified 2 groups of patients: patients with a normal thyroid, defined as a homogeneously enhancing thyroid gland on 4D-MDCT (group 2); and patients with a diseased-but-nonmalignant thyroid, defined as a thyroiditis with increased vascular flow or benign multinodular goiter on sonography or pathology, and a heterogeneously enhancing thyroid gland on 4D-MDCT (group 3).

For inclusion in the study, patients with a normal or diseased thyroid had to have no evidence of a primary cancer either at the time of the 4D-MDCT study or previously.

The medical records were searched until equal numbers of recurrent lesions (group 1) and both normal thyroid (group 2) and diseased thyroid (group 3) cases in each group were identified. The review was completed when there were an equal number of cases in each of the 3 study groups.

**Image Acquisition**

The 4D-MDCT examinations were performed on a multidetector CT LightSpeed scanner (GE Healthcare) with the following parameters: 140 kV, 220–250 mA, and a 1.25-mm section thickness covering the neck from the bottom of the orbits to the arch of the aorta. We acquired 4 phases: an initial noncontrast helix and then 3 additional helices acquired after intravenous contrast agent injection in the arterial phase (25 seconds after the start of the contrast injection), venous phase (55 seconds after injection), and delayed phase (85 seconds after injection).

**ROI Analysis**

CT images were transferred to an ADW-2 workstation (GE Healthcare) with Volume Share 2, Version 4.4, software. A head and neck radiologist (J.M.D.) with >15 years of experience in head and neck radiology who was not blinded to the tissue type results drew ROIs around the recurrent lesions and areas of normal and diseased thyroid using freehand contouring. In the freehand contouring technique (Fig 1), the ROI was contoured to the visual outline of the lesion or the thyroid gland through the different phases of enhancement, with no requirement for the ROI size to remain constant. The outline from the largest area of enhancement representing the entire lesion was copied onto a similar site on the precontrast phase image. An attempt was made to place the contour 1 mm from the edge of the lesion to minimize the effects of partial volume averaging. Any visible surgical clips or calcifications associated with the lesions were excluded.

Mean, SD, and minimum-maximum of the grouped data were calculated to compare recurrence, normal thyroid, and diseased thyroid Hounsfield unit values through the 4 phases of enhancement. First, ANOVA was used to compare Hounsfield unit values among 3 tissue types. Tukey-Kramer adjustment was used to control the overall type I error rate at 5%. A 2-sample \( t \) test was used to compare Hounsfield unit values of dichotomized cases of recurrence versus benign thyroid (normal and diseased combined).

Second, a univariate logistic regression model was used to associate Hounsfield units with tissue status (recurrence versus benign thyroid). As a high correlation among the different phases was determined, a multivariate model was reduced to a univariate model by the backward deletion. Third, the performance of Hounsfield units to discriminate recurrence from benign thyroid tissue using both the results of the univariate and bivariate analyses was evaluated using receiver operating characteristic (ROC) curve analysis. ROC analysis was used to determine the optimal cutoff point by maximizing the Youden index, sensitivity + specificity–1. Next, ROC
curve analysis of Hounsfield units to discriminate recurrence from benign thyroid tissue was performed using univariate and multivariate analyses with 2, 3, or 4 combinations of the various 4 phases (precontrast, arterial, venous, and delayed).

Similar overall root mean square error values were then calculated for each phase. Matrix scatterplots with Pearson correlation were used to compare the Hounsfield unit density through the 4 phases. All tests were 2-sided, and \( P \leq .05 \) was considered statistically significant. Statistical analysis was performed using SAS, Version 9.4 (SAS Institute) and R statistical and computing software, Version 3.5.3 (http://www.r-project.org/).

RESULTS

Study Population

Twenty-nine recurrent thyroid cancer lesions (group 1) were identified in 24 patients (16 women, 8 men; 21–80 years of age; mean, 52.5 ± 15.8 years); 5 patients each had 2 separate lesions. The types of primary thyroid cancer were as follows: papillary thyroid carcinoma (n = 21), follicular variant papillary thyroid carcinoma, (n = 3), medullary thyroid carcinoma (n = 3), and Hurthle cell carcinoma (n = 2). Sixteen of 24 patients (19 lesions) received radioactive iodine following thyroidectomy. The lesions ranged in size from 0.7 to 2.4 cm (mean, 1.2 ± 0.4 cm) and were diagnosed by surgical resection (n = 28) or fine-needle aspiration (n = 1). Twenty-nine patients with normal thyroid (group 2) (25 women,

Table 1: Summary of Hounsfield units by phase and lesion type using 3 categories—normal, diseased, and recurrent

<table>
<thead>
<tr>
<th>Group</th>
<th>No.</th>
<th>Mean</th>
<th>SD</th>
<th>F</th>
<th>( P ) Value*</th>
<th>Tukey HSD Comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Disease</td>
<td>Normal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Recurrence</td>
</tr>
<tr>
<td>Precontrast</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>29</td>
<td>109.03</td>
<td>16.19</td>
<td>133.73</td>
<td>&lt;.001</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Disease</td>
<td>29</td>
<td>88.07</td>
<td>17.54</td>
<td></td>
<td></td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Recurrence</td>
<td>29</td>
<td>37.09</td>
<td>17.91</td>
<td></td>
<td></td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Arterial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>29</td>
<td>233.93</td>
<td>32.56</td>
<td>105.45</td>
<td>&lt;.001</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Disease</td>
<td>29</td>
<td>185.24</td>
<td>36.25</td>
<td></td>
<td></td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Recurrence</td>
<td>29</td>
<td>99.98</td>
<td>37.67</td>
<td></td>
<td></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Venous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>29</td>
<td>191.72</td>
<td>29.5</td>
<td>72.52</td>
<td>&lt;.001</td>
<td>0.0017</td>
</tr>
<tr>
<td>Disease</td>
<td>29</td>
<td>161.03</td>
<td>36.74</td>
<td></td>
<td></td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Recurrence</td>
<td>29</td>
<td>90.61</td>
<td>31.69</td>
<td></td>
<td></td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Delayed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>29</td>
<td>172.31</td>
<td>24.55</td>
<td>67.46</td>
<td>&lt;.001</td>
<td>0.0008</td>
</tr>
<tr>
<td>Disease</td>
<td>29</td>
<td>144.07</td>
<td>31.06</td>
<td></td>
<td></td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Recurrence</td>
<td>29</td>
<td>87.81</td>
<td>28.6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note:—HSD indicates honestly significant difference.
* \( P \) values were from ANOVA.

Table 2: Summary of Hounsfield units by phase and lesion type using 2 categories—recurrence and benign (includes both normal and diseased thyroid)

<table>
<thead>
<tr>
<th>Phase/Group</th>
<th>No.</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>( P ) Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precontrast</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benign thyroid</td>
<td>58</td>
<td>98.55</td>
<td>19.79</td>
<td>54</td>
<td>138</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Recurrence</td>
<td>29</td>
<td>37.09</td>
<td>17.91</td>
<td>10.8</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>Arterial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benign thyroid</td>
<td>58</td>
<td>209.59</td>
<td>42.06</td>
<td>131</td>
<td>331</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Recurrence</td>
<td>29</td>
<td>99.98</td>
<td>37.67</td>
<td>39.7</td>
<td>174</td>
<td></td>
</tr>
<tr>
<td>Venous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benign thyroid</td>
<td>58</td>
<td>176.38</td>
<td>36.47</td>
<td>96</td>
<td>263</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Recurrence</td>
<td>29</td>
<td>90.61</td>
<td>31.69</td>
<td>28.2</td>
<td>156</td>
<td></td>
</tr>
<tr>
<td>Delayed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benign thyroid</td>
<td>58</td>
<td>158.19</td>
<td>31.19</td>
<td>94</td>
<td>238</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Recurrence</td>
<td>29</td>
<td>87.83</td>
<td>28.6</td>
<td>24.2</td>
<td>155</td>
<td></td>
</tr>
</tbody>
</table>

Note:—Min indicates minimum; Max, maximum.
Discrimination among Normal, Diseased, and Recurrent Thyroid

Hounsfield unit densities for recurrent, normal, and diseased thyroids were compared precontrast and at 25 (arterial), 55 (venous), and 85 (delayed) seconds. ANOVA analysis showed that recurrent disease demonstrated significantly lower Hounsfield unit values in all 4 phases compared with diseased thyroid \((P < .001)\). Diseased thyroid also had a significantly lower values than normal thyroid (Table 1). Combined benign thyroid (normal and diseased thyroid) was also significantly different from recurrence \((P < .001)\) in all 4 phases using a 2-sample \(t\) test (Table 2). A line graph with error bars was drawn to show this evolution of Hounsfield unit density with time for the recurrent disease and benign thyroid groups (Fig 2).

**Statistical Analysis**

On univariate logistic regression analysis of benign thyroid versus recurrence, the odds ratios for differentiating the various phases were as follows: precontrast odds ratio, 0.856 (95% CI, 0.767–
with the precontrast phase, none of these combinations were statistically significant ($P > .05$) (Table 5).

The optimal cutoff point to differentiate benign thyroid from recurrence in each of the 4 phases by ROC analysis was as follows: precontrast, $\leq 62$ HU; arterial phase, $\leq 129$ HU; venous phase, $\leq 120.6$ HU; and delayed phase, $\leq 125$ HU (Fig 4). Overall root mean square error values were as follows: precontrast, 34.8; arterial, 29.5; venous, 26.4; and delayed, 26.

To validate these findings, we made an effort to combine the phases into an indicator with better performance, but matrix scatterplots revealed 2 outlier cases that were misclassified regardless of phase, indicating that a combination of phases could not be used to improve performance (though there might be clinical reasons to include $>1$ phase).

**DISCUSSION**

Our results demonstrate that 4D-MDCT can be used to distinguish recurrent thyroid carcinoma from benign thyroid tissue, including both normal and diseased nonmalignant thyroids. In all 4 phases, the Hounsfield unit value was significantly lower for recurrent thyroid carcinoma than for benign tissue, with diseased thyroid intermediate between recurrent and normal thyroid. The mean time-enhancement curves ran parallel and did not cross in any phase. Based on the ROC curve analysis, the best single phase to distinguish recurrent disease from benign thyroid tissue, both normal and diseased, was the precontrast phase with a cutoff value of $\geq 62$ HU, yielding a sensitivity of 0.966 and specificity of 0.983. Arterial, venous, and delayed phases each followed with cutoff values of 129, 120.6, and 125, respectively, with sensitivity/specificities of 0.828/1, 0.897/0.931, and 0.966/0.879, respectively. Each individual phase, in other words, yields quite acceptable performance, but the precontrast phase, according to our data, contains the most diagnostic information for this application. The ROC curve analysis of univariate and multivariate data showed that the area under the curve could be slightly improved from 0.983 (precontrast phase alone) in various phase combinations of 2, 3, and 4 phases, but none of these were statistically significant ($P > .05$) in comparison with the precontrast phase alone.

On precontrast MDCT, both normal and diseased thyroid had a higher Hounsfield unit density compared with recurrence. This is attributed, at least in part, to the iodine content of normal thyroid tissue and the inclusion of this thyroid tissue in the ROI measurements of the diseased thyroid.
Histologically, the thyroid gland is composed of epithelial cells surrounding follicles. Thyroid hormones produced from thyroglobulin and iodide are stored within colloid and account for 25% of the iodide of the body. Because colloid volume is about 3 times greater than the epithelial cell volume, the high electron density of the iodine colloid leads to a high attenuation of the thyroid gland on MDCT. Our results are confirmed in a report by Paty et al, who found that the higher density of the thyroid gland in benign and malignant nodules compared with normal hyperdense thyroid tissue in our study, the innate higher density of benign and malignant thyroid nodules. Their results in the precontrast phase were similar to ours because the precontrast attenuation of normal thyroid tissue in the arterial phase (25 seconds) was approximately 50 HU over baseline and parathyroid adenomas increased approximately 100 HU from baseline. Ours and the study of Hunter et al also found the greatest separation between thyroid tissue and recurrence, and adenomas and thyroid tissue in the arterial phase, respectively. Fitzgerald et al found no significant difference in HU values between the benign (128 HU) and malignant nodules (144 HU) in the arterial phase (P = .7). Differences between their study and ours are likely related to the inclusion of normal thyroid tissue in the diseased thyroid group in our study, which would not be present in the thyroid nodules studied by Fitzgerald et al.

In our study, there was washout of contrast in the delayed and late phases in all 3 studied groups. Similarly, Hunter et al found that nodular and parathyroid adenomas increased approximately 100 HU from baseline. Ours and the study of Hunter et al also found the greatest separation between thyroid tissue and recurrence, and adenomas and thyroid tissue in the arterial phase, respectively. Fitzgerald et al found no significant difference in HU values between the benign (128 HU) and malignant nodules (144 HU) in the arterial phase (P = .7). Differences between their study and ours are likely related to the inclusion of normal thyroid tissue in the diseased thyroid group in our study, which would not be present in the thyroid nodules studied by Fitzgerald et al.

Yoon et al examined 734 patients with incidental thyroid nodules, 9% of which were malignant. They found that nodular or rim calcifications and Hounsfield unit values of >130 HU on postcontrast CT with a 45-second delay were suggestive of malignancy. On venous phase imaging (55 seconds), we found mean Hounsfield unit values of 91 HU for recurrent thyroid carcinoma. We speculate that the differences in results between our study and that of Yoon et al may also be related to the presence of calcifications in the incidentally detected thyroid carcinomas described in their study.

Limitations include the retrospective nature of the study and the relatively small number of patients (29 in each category). Another limitation is that the lesions in our study may have been detected earlier in their development, given that the patients in our study were being screened for recurrence. During a longer time course, recurrence and residual tissue may have a different appearance. However, the current study design is useful in the early follow-up period when such information is most useful clinically and could be used to decide on questions related to the inclusion of thyroid tissue in the diseased thyroid group in our study, which would not be present in the thyroid nodules studied by Fitzgerald et al.

Table 4: Area under ROC curve (univariate analysis)

<table>
<thead>
<tr>
<th>Phase</th>
<th>AUC</th>
<th>95% LCL</th>
<th>95% UCL</th>
<th>Optimal Cutoff (HU)</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precontrast</td>
<td>0.983</td>
<td>0.954</td>
<td>1</td>
<td>62</td>
<td>0.966</td>
<td>0.983</td>
</tr>
<tr>
<td>Arterial</td>
<td>0.979</td>
<td>0.955</td>
<td>1</td>
<td>129</td>
<td>0.828</td>
<td>1</td>
</tr>
<tr>
<td>Venous</td>
<td>0.965</td>
<td>0.933</td>
<td>0.997</td>
<td>120.6</td>
<td>0.897</td>
<td>0.931</td>
</tr>
<tr>
<td>Delayed</td>
<td>0.957</td>
<td>0.916</td>
<td>0.998</td>
<td>125</td>
<td>0.966</td>
<td>0.879</td>
</tr>
</tbody>
</table>

Note: AUC indicates area under the curve; HU, Hounsfield unit; LCL, lower control limit; UCL, upper control limit.

Table 5: Area under ROC curve for multivariate analysis with P values compared with precontrast phase alone

<table>
<thead>
<tr>
<th>Phase</th>
<th>AUC</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precontrast</td>
<td>0.983</td>
<td>.036</td>
</tr>
<tr>
<td>Arterial</td>
<td>0.979</td>
<td>.8</td>
</tr>
<tr>
<td>Venous</td>
<td>0.965</td>
<td>.21</td>
</tr>
<tr>
<td>Delayed</td>
<td>0.957</td>
<td>.04</td>
</tr>
<tr>
<td>Precontrast, arterial</td>
<td>0.994</td>
<td>.26</td>
</tr>
<tr>
<td>Precontrast, venous</td>
<td>0.985</td>
<td>.28</td>
</tr>
<tr>
<td>Precontrast, delayed</td>
<td>0.983</td>
<td>NA</td>
</tr>
<tr>
<td>Arterial, venous</td>
<td>0.979</td>
<td>.8</td>
</tr>
<tr>
<td>Arterial, delayed</td>
<td>0.977</td>
<td>.73</td>
</tr>
<tr>
<td>Venous, delayed</td>
<td>0.964</td>
<td>.2</td>
</tr>
<tr>
<td>Precontrast, arterial, venous</td>
<td>0.995</td>
<td>.26</td>
</tr>
<tr>
<td>Precontrast, arterial, delayed</td>
<td>0.995</td>
<td>.29</td>
</tr>
<tr>
<td>Precontrast, venous, delayed</td>
<td>0.987</td>
<td>.57</td>
</tr>
<tr>
<td>Arterial, venous, delayed</td>
<td>0.979</td>
<td>.81</td>
</tr>
<tr>
<td>Precontrast, arterial, venous, delayed</td>
<td>0.996</td>
<td>.37</td>
</tr>
</tbody>
</table>

Note: NA indicates not applicable; AUC, area under the curve; P, P value.
such as re-operation. Another limitation is that patients with a multinodular thyroid as part of the diseased thyroid group could conceivably have had an undetected indolent thyroid cancer. In addition, the use of radioactive iodine for the treatment of thyroid cancer might alter the imaging appearance of benign thyroid tissue from what we observed. After radioactive ablation therapy, residual thyroid tissue undergoes fibrosis and appears sonographically as an avascular heterogeneously hypoechoic mass. Caution should be used when extrapolating work done in the thyroid gland that has been treated with radioactive iodine to our current findings in the postthyroidectomy population. Further investigation is needed in this area.

Approximately 27 mSv of radiation is administered during a 4D-MDCT examination, with a roughly equal contribution from each phase. Because approximately 3 mSv is the estimated yearly natural background radiation exposure, the 4D-MDCT dose in our study is equivalent to the dose received naturally in 9 years. The dose for a CT study could potentially be cut to about 6.75 mSv if only 1 phase is acquired (but see next paragraph).

**CONCLUSIONS**

Our data suggest that the precontrast imaging for characterizing residual tissue after thyroidectomy actually contains the most valuable information and should certainly be a component of any imaging protocol performed for this purpose. The choice of which phase or phases to include is informed but not dictated by our work. Clinical needs might demand that contrast-enhanced sequences be a part of the imaging protocol (for example to

---

**FIG 4.** Logistic regression analysis determined the optimal cutoff Hounsfield unit value in each of the 4 phases. The least overlap was present in the precontrast phase with a value of \(\approx 62\) HU.
demonstrate lymph nodes and vascular anatomy). Individual radiologists and practices can use the data found in this study to optimally select those imaging phases needed in their practice and balance the need for clinical information against radiation burden, as their patients’ needs might require. Prospective clinical imaging trials using these cutoff values are justified.

Disclosures: Mark E. Zafereo—UNRELATED: Grants/Grants Pending: Merck, Eli Lilly, Genetech, Comments: research funding for clinical trials. —Maria E. Cabanillas—UNRELATED: Consultancy: Blueprint Medicines, Ignita, Bayer Onyx, Loxo Oncology, Grants/Grants Pending: Eisai, Exelixis, Genetech, Kura Oncology; Payment for Lectures Including Service on Speakers Bureaus: various Continuing Medical Education lectures with multiple sponsors; Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed: various Continuing Medical Education lectures. *Money paid to the institution.

REFERENCES

MRI Signal Intensity and Electron Ultrastructure Classification Predict the Long-Term Outcome of Skull Base Chordomas


ABSTRACT

BACKGROUND AND PURPOSE: MR imaging is a useful and widely used evaluation for chordomas. Prior studies have classified chordomas into cell-dense type and matrix-rich type according to the ultrastructural features. However, the relationship between the MR imaging signal intensity and ultrastructural classification is unknown. We hypothesized that MR imaging signal intensity may predict both tumor ultrastructural classification and prognosis.

MATERIALS AND METHODS: Seventy-nine patients with skull base chordomas who underwent 95 operations were included in this retrospective single-center series. Preoperative tumor-to-pons MR imaging signal intensity ratios were calculated and designated as ratio on T1 FLAIR sequence (RT1), ratio on T2 sequence (RT2), and ratio on enhanced T1 FLAIR sequence (REN), respectively. We assessed the relationships among signal intensity ratios, ultrastructural classification, and survival.

RESULTS: Compared with the matrix-rich type group, the cell-dense type chordomas showed lower RT2 (cell-dense type: 1.90 ± 0.38; matrix-rich type: 2.61 ± 0.60 P < .001). The model of predicting cell-dense type based on RT2 had an area under the curve of 0.83 (95% CI, 0.75–0.92). In patients without radiation therapy, both progression-free survival (P = .003) and overall survival (P = .002) were longer in the matrix-rich type group than in the cell-dense type group. REN was a risk factor for progression-free survival (hazard ratio = 10.24; 95% CI, 1.73–60.79); RT2 was a protective factor for overall survival (hazard ratio = 0.33; 95% CI, 0.12–0.87); and REN was a risk factor for overall survival (hazard ratio = 4.76; 95% CI, 1.51–15.01).

CONCLUSIONS: The difference in MR imaging signal intensity in chordomas can be explained by electron microscopic features. Both signal intensity ratios and electron microscopic features may be prognostic factors.

ABBREVIATIONS: CDT = cell-dense type; MRT = matrix-rich type; OS = overall survival; PFS = progression-free survival; RT1, RT2, and REN = ratios of tumor-to-pons signal intensity in the T1 FLAIR sequence, T2 sequence, and enhanced T1 FLAIR sequence, respectively.

Chordoma is a rare malignant tumor with an overall incidence of 0.08 per 100,000 in the United States. These tumors occur mainly in the sacrum, followed by the skull base and spine. Skull base chordomas account for 1% of all brain tumors and 4% of all primary bone tumors. Chordomas generally grow slowly; however, recurrence is common because en bloc resection is difficult in the skull base. The recurrence intervals are distinctively variable among patients due to biologic variability among and within patients. This difference also impacts the treatment plan; for example, some patients should receive radiation therapy as early as possible, and others can wait. Previous studies by our group and others have demonstrated several risk factors related to long-term survival, such as the expression level of Ki-67, Platelet-derived growth factor receptor β (PDGFR-β) and SNF5, resection extent, age, postsurgery radiation therapy, and histopathologic findings. On the basis of our previous study on the electron microscopic ultrastructural characteristics of chordomas, we have classified chordomas into 2 different groups: cell-dense type (CDT) and matrix-rich...
We have shown that CDT is predictive of poorer outcomes, higher risk of recurrence, and shorter survival, though these findings were derived from a small sample size (n = 27).

MR imaging allows a differential diagnosis between chordomas and other tumors using signal intensity.14,15 Recent studies have suggested that MR imaging signal intensity ratios on different sequences (ratio of tumor-to-pons signal intensity on T1 FLAIR [RT1]; ratio of tumor-to-pons signal intensity on T2 [RT2]; ratio of tumor-to-pons signal intensity on enhanced T1 FLAIR [REN]) are useful in predicting outcomes. Tian et al16 revealed that higher RT2 predicted diminished tumor progression and higher REN predicted more rapid tumor progression; this correlation between REN and recurrence or progression was also shown by Lin et al17 However, neither of these studies showed evidence for the possible explanation of signal intensity difference.

We hypothesized that ultrastructural morphologic characters could explain the mechanism of different signal intensities. Hence, this study aimed to explore the relationships among MR imaging signal intensity ratio, the electronic microscopic feature, and survival, which may provide meaningful preoperative prognostic information and help apply better individual-based treatment strategies.

### MATERIALS AND METHODS

#### Patients

Data from 127 patients who underwent an operation by the same surgical team between July 2012 and July 2016 were reviewed retrospectively from the chordoma data base of the neuroendoscopic ward in Beijing Tiantan Hospital. The present series were classified into CDT and MRT according to the electronic features (Table 1).13 Overall, in contrast to the features in the MRT tumors, tumor density was high in CDT tumors; thus, the extracellular matrix was sparse.13 Only the patients with both electron microscopy images and preoperative MR imaging available were included for further analysis; thus, 79 patients were enrolled in the present study. Among these patients, 27 cases were previously reported.13 The relevant clinical information was collected from the inpatient digital archives, including sex, age, tumor status (primary or recurrent tumor), presurgery radiation therapy, tumor blood supply (rich or not rich, with a rich blood supply being defined by having a tumor resection surface that bled easily and was difficult to suction cleanly, while a blood supply was that was not rich was defined by having a tumor resection surface that bled less easily and was easy to suction cleanly), resection rate (≥90% or <90%), and histopathology (Ki-67 level, conventional chordoma, or chondroid chordoma). As a retrospective observational study of de-identified data without any additional therapy or monitoring procedures, this study did not need approval from the Ethics Committee of Beijing Tiantan Hospital.

#### MR Imaging Scans and Image Evaluation

The patients underwent preoperative MR imaging with a 3T MR imaging scanner (Siemens, Erlangen, Germany). Axial unenhanced T1 FLAIR and T2-weighted images were obtained with the following parameters, respectively: TR/TE range = 1850–2500/9.4–19.8 ms and 4500–6000/84–97 ms. Postcontrast scanning was performed 2 minutes after gadopentetate dimeglumine injection, with a concentration of 0.2 mL/kg (0.5 mol/L). The TR/TE ranges for the enhanced T1 FLAIR sequence were 1850–2500/9.4–19.8 ms. The section thickness was 3.0–50.5 mm.

The images were analyzed on our PACS. Operator-defined ROIs were placed in the tumor and the pons. The mean signal intensity of the tumor was normalized by dividing by the mean signal intensity of the corresponding pons on axial T1 FLAIR (RT1), T2 (RT2), and enhanced T1 FLAIR (REN), respectively.

#### Follow-Up

The follow-up period was calculated from the date of the first operation in Beijing Tiantan Hospital to the last follow-up in March 2019 or the date of death, whichever occurred first. The median follow-up time in this series was 52 months (range, 1–118 months), and 7 patients were lost to follow-up. Progression-free survival (PFS) was calculated from the date of the operation to the date of tumor recurrence or progression of residual tumor, which was confirmed by the imaging findings; and overall survival (OS) was defined as the time from diagnosis to disease-specific death or to the last follow-up time. Patients were censored at the last follow-up date if no death occurred or on the date of death if death was not disease-specific.5,12

#### Statistical Analysis

Statistical analyses were performed with SPSS software, Version 24 (IBM). T tests or Mann-Whitney U tests were used to analyze the continuous variables between groups. A χ² test was used in analyzing the distributions of categoric variables between groups. To evaluate the relationship between signal intensity ratio and electron microscopic classification, we included the signal intensity ratios (RT1, RT2, REN) in univariate analyses separately; multivariate analysis was performed using logistic regression. The area under the receiver operating characteristic curve was calculated to estimate the accuracy.

### Table 1: Key points for the differential diagnosis of cell-dense and matrix-rich types according to the ultrastructures of chordomas*

<table>
<thead>
<tr>
<th>Tumor cell content in a low-power field (×3000)</th>
<th>Densely arranged</th>
<th>Seldom or none</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extracellular matrix content</td>
<td>Relatively sparse</td>
<td>Abundant</td>
</tr>
<tr>
<td>Complex of rough endoplasmic reticulum and the mitochondria</td>
<td>Rich</td>
<td>Relatively seldom</td>
</tr>
<tr>
<td>Nuclei and chromatin</td>
<td>Abnormal nuclei with condensed, deeply staining heterochromatin</td>
<td>Regularly shaped nuclei, mainly with loosely arranged euchromatin</td>
</tr>
<tr>
<td>Cell junctions</td>
<td>Desmosomes are obvious</td>
<td>Few desmosomes</td>
</tr>
<tr>
<td>Collagen fibrils</td>
<td>Type III collagen fibrils can be found</td>
<td>Seldom</td>
</tr>
<tr>
<td>Exocytosis</td>
<td>Easily found</td>
<td>Seldom</td>
</tr>
</tbody>
</table>

---

*Reprinted with permission from Bai et al13 and Elsevier.*
OS and PFS were estimated using the Kaplan-Meier method, and survival curves for the 2 electron microscopic classifications were compared using the log-rank test. Statistically significant variables in the univariate analyses were included in the multivariate Cox proportional hazards regression model, using forward stepwise variable selection, to find significant predictors of recurrence and survival. A P value < .05 was considered statistically significant.

RESULTS
Clinical Characteristics and Electron Microscopic Classifications
Seventy-nine cases fit the inclusion criteria and were enrolled in the study (Table 2), which included 37 cases with newly diagnosed primary tumors and 42 cases with recurrent tumors or progression of residual tumors after prior operations. CDT chordomas were more common in recurrent tumors than in primary tumors (P = .01). There were 43 men and 36 women, and the mean age was 47.5 ± 16.2 years. Seventy-one cases were confirmed as conventional chordomas, and 8 were chondroid chordomas. Among them, 47 patients received adjuvant radiation therapy, while the other 32 patients received no radiation therapy. Because no medical treatment was approved for chordoma, no patients received any drug treatment.

According to the electron microscopic classification, 39 cases were classified as MRT and 40 cases were classified as CDT in the first operation. The MRT tumor cells were loosely arranged, the extracellular matrix was abundant, and the tumor cells showed bubble-like vacuoles. CDT tumors displayed relatively tightly arranged tumor cells, a less abundant matrix, and vacuoles not as rich as those in MRT cells (Fig 1). In the patients with primary tumors, patients with MRT tumors were younger than patients with CDT tumors (42.3 ± 18.2 years versus 58.2 ± 9.97 years, P = .006).

Sixty-five patients underwent 1 operation during follow-up, and 14 patients underwent repeat operations due to tumor recurrence or progression of residual tumor. No patients underwent a repeat operation because of an incomplete initial resection. Among those 14 patients, 12 patients underwent 2 operations and 2 patients underwent 3 operations. Therefore, a total of 79 patients with 95 operations were included in the following analysis (Table 3). The median PFS for the patients who underwent a repeat operation was 10.0 months (range, 2–42 months). Among the patients who underwent 2 operations, 6 patients were diagnosed with CDT tumors in both operations, 3 patients were diagnosed with MRT tumors in both operations, 2 patients were diagnosed with a MRT tumor in the first operation and a CDT tumor in the second operation, and 1 patient was diagnosed with a CDT tumor in the first operation and an MRT tumor in the second operation. Among the 2 patients who had 3 operations, one was diagnosed each time with CDT tumors, while the other was diagnosed twice with CDT tumors and once with an MRT tumor.

Relationship between Signal Intensity Ratio and Electron Microscopic Classifications
Presurgery MR images of 95 operations were pooled and analyzed. The CDT chordomas showed lower signal intensity than MRT chordomas on T2-weighted images (Table 4 and Fig 2). The $R_{T1}$ values of CDT chordomas were higher than those of MRT chordomas. The $R_{EN}$ value was also higher in patients with CDT than in those with MRT. When only 37 primary tumors were considered, the difference was still significant (Table 4).

---

**Table 2: Demographics of the 79 patients undergoing chordoma surgery in Beijing Tiantan Hospital between July 2012 and July 2016**

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>MRT</th>
<th>CDT</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of cases</td>
<td>79</td>
<td>39</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Tumor status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>37</td>
<td>24</td>
<td>13</td>
<td>.01</td>
</tr>
<tr>
<td>Recurrent</td>
<td>42</td>
<td>15</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>43</td>
<td>21</td>
<td>22</td>
<td>.918</td>
</tr>
<tr>
<td>Female</td>
<td>36</td>
<td>18</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Age (mean) (yr)</td>
<td>47.5 ± 16.2</td>
<td>43.9 ± 17.5</td>
<td>51.0 ± 14.3</td>
<td>.050</td>
</tr>
<tr>
<td>Histopathology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional</td>
<td>71</td>
<td>35</td>
<td>36</td>
<td>1.000</td>
</tr>
<tr>
<td>Chondroid</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Blood supply</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rich</td>
<td>21</td>
<td>7</td>
<td>14</td>
<td>.073</td>
</tr>
<tr>
<td>Not rich</td>
<td>55</td>
<td>31</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Resection grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥90%</td>
<td>43</td>
<td>23</td>
<td>20</td>
<td>.423</td>
</tr>
<tr>
<td>&lt;90%</td>
<td>36</td>
<td>16</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Survival status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alive</td>
<td>42</td>
<td>27</td>
<td>15</td>
<td>.018</td>
</tr>
<tr>
<td>Died</td>
<td>26</td>
<td>8</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>11</td>
<td>4</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Presurgery radiation therapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>17</td>
<td>5</td>
<td>12</td>
<td>.063</td>
</tr>
<tr>
<td>No</td>
<td>62</td>
<td>34</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Postsurgery radiation therapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>30</td>
<td>14</td>
<td>16</td>
<td>.667</td>
</tr>
<tr>
<td>No</td>
<td>43</td>
<td>21</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>PFS (mo)$^d$</td>
<td>40.3 ± 5.4</td>
<td>52.2 ± 5.7</td>
<td>18.7 ± 6.9</td>
<td>.003</td>
</tr>
<tr>
<td>OS (mo)</td>
<td>100.5 ± 7.1</td>
<td>113.9 ± 6.3</td>
<td>88.9 ± 9.3</td>
<td>.018</td>
</tr>
</tbody>
</table>

**Note:** NA indicates not applicable.

$^a$ Operative notes on 3 cases showed no specified blood supply.

$^b$ Including both patients lost to follow-up and patients who did not die of chordomas.

$^c$ Including the patients whose radiation therapy information was not available.

$^d$ Patients without radiation therapy (n = 32)
However, when only recurrent tumors were analyzed, both $R_{T1}$ and $R_{EN}$ were no longer statistically different, while $R_{T2}$ was still significantly different (Table 4).

Logistic regression analysis was then conducted to predict electron microscopic classification. A univariate analysis indicated that recurrent tumor ($P = .004$) and decreased $R_{T2}$ ($P < .001$) and increased $R_{T1}$ and $R_{EN}$ ($P = .019$ and 0.027, respectively) were significant variables for predicting CDT, while presurgery radiation therapy and histopathology subtype were not significant variables ($P > .05$). From the multivariate analysis, $R_{T2}$ was the only significant independent predictor of CDTs ($P < .001$). The area under the receiver operating characteristic curve was 0.83 (95% CI, 0.75–0.92) with only $R_{T2}$ in the model.

### Relationship between Electron Microscopic Classifications and PFS and OS

At follow-up, 26 patients (32.9%) had died of disease progression. Concerning the 32 patients without radiation therapy, 9 patients had died, including 2 patients in the MRT group and 7 patients in the CDT group. Among the 95 operations, 59 tumors (62.1%) had recurred.

When all 95 operations were analyzed together, the PFS was significantly longer in the MRT group than in the CDT group, with median PFS values of 36.0 months (interquartile range, 14–73 months) and 10 months (interquartile range, 6–42 months), respectively ($P = .003$). When the tumors without radiation therapy were considered, the PFS was 56 months (interquartile range, 34–undefined months) in the MRT group and 10 months (interquartile range, 5–24 months) in the CDT group ($P = .003$) (Fig 3A).

The OS was significantly longer in the MRT group than in the CDT group, with a median OS of 132.0 months (interquartile range, 82–undefined months) and 96.0 months (interquartile range, 48–114 months), respectively ($n = 79$; $P = .018$). When just the tumors without radiation therapy ($n = 32$) were considered, the MRT group showed a

### Table 3: Details for the electron microscopic classification

<table>
<thead>
<tr>
<th>Surgery No./Electron Microscopic Classification</th>
<th>Tumor Statusa</th>
<th>No. of Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Operation ($n = 65$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDT</td>
<td>Primary</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Recurrent</td>
<td>23</td>
</tr>
<tr>
<td>MRT</td>
<td>Primary</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Recurrent</td>
<td>14</td>
</tr>
<tr>
<td>2 Operations ($n = 12$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Both CDTs</td>
<td>Primary</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Recurrent</td>
<td>3</td>
</tr>
<tr>
<td>Both MRTs</td>
<td>Primary</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Recurrent</td>
<td>0</td>
</tr>
<tr>
<td>First MRT, second CDT</td>
<td>Primary</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Recurrent</td>
<td>1</td>
</tr>
<tr>
<td>First CDT, second MRT</td>
<td>Primary</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Recurrent</td>
<td>0</td>
</tr>
<tr>
<td>3 Operations ($n = 2$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 CDTs</td>
<td>Primary</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Recurrent</td>
<td>0</td>
</tr>
<tr>
<td>2 CDTs, 1 MRTb</td>
<td>Primary</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Recurrent</td>
<td>1</td>
</tr>
</tbody>
</table>

a The tumor status when the first operation was performed in our hospital.

### Table 4: Comparison of signal intensity ratios between CDT and MRT

<table>
<thead>
<tr>
<th>Signal Intensity Ratio</th>
<th>Electron Microscopic Classification</th>
<th>Entire cohort ($n = 95$)</th>
<th>Primary tumors ($n = 37$)</th>
<th>Recurrent tumors ($n = 58$)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_{T2}$</td>
<td>CDT</td>
<td>1.90 ± 0.38</td>
<td>1.99 ± 0.39</td>
<td>1.87 ± 0.38</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>MRT</td>
<td>2.61 ± 0.60</td>
<td>2.69 ± 0.53</td>
<td>2.50 ± 0.67</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>$R_{T1}$</td>
<td>CDT</td>
<td>0.63 ± 0.14</td>
<td>0.67 ± 0.18</td>
<td>0.62 ± 0.13</td>
<td>.019</td>
</tr>
<tr>
<td></td>
<td>MRT</td>
<td>0.56 ± 0.13</td>
<td>0.55 ± 0.13</td>
<td>0.58 ± 0.13</td>
<td>.024</td>
</tr>
<tr>
<td>$R_{EN}$</td>
<td>CDT</td>
<td>1.19 ± 0.42</td>
<td>1.17 ± 0.42</td>
<td>1.13 ± 0.43</td>
<td>.204</td>
</tr>
<tr>
<td></td>
<td>MRT</td>
<td>0.99 ± 0.41</td>
<td>0.89 ± 0.32</td>
<td>1.11 ± 0.49</td>
<td>.526</td>
</tr>
</tbody>
</table>

However, when only recurrent tumors were analyzed, both $R_{T1}$ and $R_{EN}$ were no longer statistically different, while $R_{T2}$ was still significantly different (Table 4).
significantly longer OS than the CDT group, with 2 deaths in the MRT group compared with 7 deaths in the CDT group ($P = .002$) (Fig 3B).

**Relationship between MR Imaging Signal Intensity Ratio and Expression of Ki-67**
The Ki-67 levels were available in 53 samples, which were collected from pathologic reports. The median Ki-67 level was 2% (range, 0%–20%) in 23 MRT chordomas, and the median Ki-67 level was 5% (range, 1%–20%) in 30 CDT chordomas. The expression of Ki-67 was significantly higher in the CDT group than in the MRT group ($P = .033$).

**DISCUSSION**
Skull base chordoma can be an aggressive malignancy with poor prognosis. We previously demonstrated that electron microscopic features of chordomas were variable and could be classified into 2 morphologic types, CDT and MRT.13 In the present study, we show that electron microscopic classification is a valid predictor of survival. Patients with CDT had a significantly shorter PFS and OS than those with MRT, even after adjusting for the adjuvant radiation therapy. We further demonstrated that the different electron microscopic features of chordomas exhibited different MR imaging signal intensity ratios, especially $R_{T2}$, and CDT chordomas had lower $R_{T2}$ than MRT chordomas. In a recent radiomics study of the differentiation among primary chordoma, giant-cell tumor, and metastatic tumor of the sacrum, the radiomics model constructed on the basis of T2-weighted and contrast-enhanced T1-weighted MR imaging provided good predictive values.18 Another radiomics...
study also found that T2-weighted and contrast-enhanced T1-weighted MR imaging provided more useful information in differentiating chordoma from chondrosarcoma, compared with T1-weighted MR imaging.\textsuperscript{19} An additional study reported that both T2-weighted and contrast-enhanced T1-weighted MR imaging are helpful in differentiating recurrent tumors from postoperative changes.\textsuperscript{3} In the present study, we additionally found that lower RT2 was a relatively accurate indicator for CDT (area under the receiver operating characteristic curve = 0.83; 95% CI, 0.75–0.92), which is applicable in both primary chordomas and recurrent chordomas.

In CDT chordomas, the tumor cells are rich and contact tightly with each other focally, and the extracellular matrix contents are relatively thin. Conversely, in MRT chordomas, the tumor cells are sparsely arranged and the extracellular matrix contents are abundant.\textsuperscript{13} These morphologic characteristics were also confirmed by hematoxylin-eosin staining under low-power microscopy.\textsuperscript{20} There were numerous vacuoles of various sizes in the cytoplasm of MRT chordomas, while the vacuoles were absent or less obvious in CDT chordomas. Furthermore, the karyoplasm ratio increased occasionally in CDT tumor cells,\textsuperscript{13} which led to reduced fluid content in tumor cells. Both high fluid content in the physaliphorous cells and extracellular matrix likely contribute to the high signal intensity in T2-weighted MR imaging in MRT chordomas.\textsuperscript{3,14,15} Similar findings have been reported in liposarcoma and breast tumors, in which histopathologic components of loose myxoid stroma generate high T2 signal.\textsuperscript{21,22} To our knowledge, this is the first study focusing on the MR imaging signal intensity and electron microscopic features. The results partially-but-directly explain the pathologic basis of the different MR imaging signal intensities among chordomas.

Previous studies\textsuperscript{15,23} demonstrated that poorly differentiated chordomas exhibited no extracellular myxoid stroma or very minimal matrix in pathologic slides and low signal intensity in T2-weighted MR imaging, and this subset of patients has decreased survival compared with those with conventional or chondroid chordomas. Although there were no poorly differentiated chordomas in our series, CDT chordomas do have a sparser extracellular matrix and higher tumor cell density, characteristics that make it reasonable to suggest that our results are comparable with those of these studies\textsuperscript{15,23} and that RT2 is also qualified in poorly differentiated chordomas.

In the patients who underwent 1 operation by our team, we found a higher ratio of CDT in recurrent tumors (27/42, 64.3%) than in primary tumors (13/37, 35.1%) (Table 2). This finding indicates that recurrent chordomas are more prone to be CDT tumors than primary chordomas. In the patients who underwent >1 resection by our team, some tumors recurred as a different electron microscopic type from their previous surgical result. These findings may indicate that the tumor cells were heterogeneously arranged,\textsuperscript{24} and the tiny amount of tissue in the electron microscopic sample may not exactly represent the whole tumor. With MR imaging analysis, the sampling error is greatly reduced because ROIs are selected by drawing the largest possible elliptic area within the tumor but avoiding areas of hemorrhage, necrosis, cyst, or calcification.\textsuperscript{15,17,25} This feature may also indicate the emergence of new malignant cell populations within the recurrent tumor when the MRT tumors recurred as CDT tumors in some cases. Additionally, we found that most of the tumors that underwent multiple resections eventually evolved into CDTs (10/14) (Table 3). A previous study reported that 2 primary sacral chordomas with high signal intensity in T2-weighted images recurred with poorly differentiated chordomas with intermediate signal intensity on T2-weighted images.\textsuperscript{14} This occurrence suggests that decreasing T2 signal intensity may represent increasing malignancy. Recently, we confirmed that Ki-67 was associated with poor prognosis.\textsuperscript{7} In the present study, we found higher Ki-67 levels in CDT tumors than in MRT tumors. The findings suggest that there is a more aggressive proliferation ability in CDT chordomas.\textsuperscript{13} Furthermore, we found that CDT tumors were more likely to recur.\textsuperscript{13} Taken together, we believe that the extent of malignancy increases in most chordomas when they recur, and this progression can be observed by both electron microscopic diagnosis and MR signal intensity.

MR imaging is important not only for diagnosing and differentiating disease but also in predicting prognosis.\textsuperscript{3,16,24} Among the 32 patients who did not undergo adjuvant radiation therapy in the present series, R_{EN} and tumor blood supply were demonstrated as independent factors for PFS. Tian et al\textsuperscript{16} and Lin et al\textsuperscript{17} have also had similar results, which suggested that R_{EN} was a risk factor for PFS. Tian et al\textsuperscript{16} also found that tumors with higher R_{EN} had more tumor blood supply. A recent radiomic analysis of 148 skull base chordomas showed that both radiomic signature (which was based on preoperative axial T1 FLAIR, T2-weighted, and enhanced T1 FLAIR) and blood supply were independent markers of tumor progression.\textsuperscript{25} These 4 independent studies provide convincing evidence to conclude that the high R_{EN} and rich blood supply of chordomas can be used as prognostic indicators for rapid recurrence or progression.

We also demonstrated that RT2 and postsurgery radiation therapy are protective factors for survival and that R_{EN} is a risk factor for OS. Radiation therapy has been recommended as a standard procedure in chordoma treatment,\textsuperscript{2} and many studies also showed that adjuvant radiation therapy improves outcomes.\textsuperscript{4,5,26} To our knowledge, the correlation between signal intensity ratio and OS in chordomas has not been studied previously, and the precise mechanism of our findings needs further research. However, considering that 79 cases are a relatively large sample size for the low incidence and prevalence of chordoma and the statistical significance of our findings, our results are meaningful. Therefore, MR imaging signal intensity ratios are useful clinical indicators for risk stratification of patients with chordomas or treatment-plan ranking.

**Limitations**

The main limitation of this study is that the time interval between the operation and adjuvant radiation therapy was not available in the present data base, and radiation therapy most likely positively influences recurrence intervals.\textsuperscript{4,5} Furthermore, radiation therapy has been shown to impact the MR imaging enhancement parameters.\textsuperscript{24} Therefore, we were unable to control for the effect of radiation therapy in the progression-free survival models. Another limitation is that this is a retrospective study for which some clinical information was missing; for example, the tumor blood supply of 3 tumors was not recorded in the archives. A study showed that lower ADC values may predict tumor
progression in postsurgery chordomas and may also predict tumor-control probability in skull base chordomas treated with carbon ion therapy. This probability suggests that the relationship between ADC values and electronic microscopic characteristics will also be worth investigating. However, diffusion-weighted MR imaging was previously not routinely performed in our hospital; therefore, ADC values were also not available in the present data base. These deficiencies will be remedied in our future study.

**REFERENCES**

Inter- and Intrareader Agreement of NI-RADS in the Interpretation of Surveillance Contrast-Enhanced CT after Treatment of Oral Cavity and Oropharyngeal Squamous Cell Carcinoma

ABSTRACT

BACKGROUND AND PURPOSE: The Neck Imaging Reporting and Data System was introduced to assess the probability of recurrence in surveillance imaging after treatment of head and neck cancer. This study investigated inter- and intrareader agreement in interpreting contrast-enhanced CT after treatment of oral cavity and oropharyngeal squamous cell carcinoma.

MATERIALS AND METHODS: This retrospective study analyzed CT datasets of 101 patients. Four radiologists provided the Neck Imaging Reporting and Data System reports for the primary site and neck (cervical lymph nodes). The Kendall's coefficient of concordance (W), Fleiss κ (κf), the Kendall's rank correlation coefficient (τB), and weighted κ statistics (κw) were calculated to assess inter- and intrareader agreement.

RESULTS: Overall, interreader agreement was strong or moderate for both the primary site (W = 0.74, κf = 0.48) and the neck (W = 0.80, κf = 0.50), depending on the statistics applied. Interreader agreement was higher in patients with proved recurrence at the primary site (W = 0.96 versus 0.56, κf = 0.65 versus 0.30) or in the neck (W = 0.78 versus 0.56, κf = 0.41 versus 0.29). Intrareader agreement was moderate to strong or almost perfect at the primary site (range τB = 0.67–0.82, κw = 0.85–0.96) and strong or almost perfect in the neck (range τB = 0.76–0.86, κw = 0.89–0.95).

CONCLUSIONS: The Neck Imaging Reporting and Data System used for surveillance contrast-enhanced CT after treatment of oral cavity and oropharyngeal squamous cell carcinoma provides acceptable score reproducibility with limitations in patients with post-therapeutic changes but no cancer recurrence.

ORAL CAVITY SQUAMOUS CELL CARCINOMA (OCSCC) is the most common malignancy of the head and neck but might soon be overtaken by oropharyngeal squamous cell carcinoma (OPSCC), whose incidence is rapidly rising, mainly because its occurrence is related to the human papillomavirus.1–3 Smoking and alcohol use are outstanding risk factors with synergistic effects.4 While some authors use OCSCC for cancers in both locations, we think it is important to separate them. The oral cavity is separated from the oropharynx by the junction of the hard and soft palates above and the circumvallate papillae located at the transition from the anterior two-thirds to the posterior third of the tongue below.5

After completion of curative treatment for OCSCC or OPSCC, patients are enrolled in a program of continuous surveillance imaging and clinical examinations. Surveillance imaging can be performed using CT, MR imaging, or PET/CT and PET/MR imaging.56 Radiologists interpreting posttherapeutic imaging studies in these patients typically focus on the detection of submucosal recurrence at the primary cancer site and the identification of suspicious lymph nodes in the neck. Mucosal recurrence might also be seen in surveillance imaging but is a domain of referring clinicians. Especially in patients who underwent high-dose radiation therapy, the best surveillance can be ensured with a combination of clinical examinations, high-resolution imaging, and possibly endoscopy.7

Interpretation of posttherapeutic neck imaging studies in these patients is often challenging for radiologists. In this setting, nonstandardized framing is the common way to rate the probability of cancer recurrence. Reporting and Data Systems (RADS) provide standardized terminology and guidance toward a final
score reflecting the probability of malignancy in patients enrolled in cancer surveillance programs. Following the introduction of such a system for breast imaging (BI-RADS) in 1997, several RADS for different organs and body regions (eg, PI-RADS for the prostate and LI-RADS for the liver) have been published and also become highly appreciated by referring clinicians, not in the least because they improve comparability and reproducibility.8-10

In 2016, the Neck Imaging Reporting and Data System (NI-RADS) was introduced by the American College of Radiology and has shown a promising initial performance.11-13 Defined features and findings lead to a numeric value that reflects the probability of cancer recurrence and is directly linked to recommendations for measures to be taken for further patient management.

The major motivation to perform this study was to test NI-RADS for its reliability in interpreting contrast-enhanced CT (CECT), which is, by far, the most common technique used for the surveillance of patients with head and neck cancer in our institution, to obtain evidence to support its implementation as a reporting standard for imaging studies and discussion of findings with referring physicians from the department of oral and maxillofacial surgery.

MATERIALS AND METHODS

Patient Population
This retrospective study was approved by our institutional review board, and written informed consent was obtained from all patients. In the records of our weekly interdisciplinary conferences (of radiologists and oral and maxillofacial surgeons) held between June 2017 and July 2019, we identified 123 consecutive patients for whom CECT studies performed at our department or by an external institution were available, and 101 patients (41 women, 60 men; median age, 64 years) were finally included in this study. A flow chart of participants is provided in Fig 1. A total of 202 target sites (primary cancer site and neck for each patient) were evaluated. Of the patients included, 72 had OCSCC localized in the mouth floor (n = 22), the anterior two-thirds of the tongue (n = 19), the hard palate (n = 3), and the gingival, labial, or buccal mucosa (n = 28). Twenty-nine patients had OPSCC localized in the posterior third (base) of the tongue (n = 13), the soft palate (n = 2), the palatine tonsils (n = 13), and the posterior oropharyngeal wall (n = 1).

Imaging
Of the 101 CECT studies included, 72 were performed in our department, and 29, by an external institution. In our department, we perform neck CECT scans on an 80-section CT scanner (Aquilion PRIME; Canon Medical Systems, Otawara, Japan). Our standard protocol includes scout-based automated selection of tube voltages between 100 and 130 kV and tube current modulation between 60 and 600 mA, a tube rotation time of 0.75 seconds, collimated section thickness of 0.5 mm, and a pitch factor of 0.813. Seventy-five milliliters of contrast medium (iopamidol, Imeron 400; Bracco, Milan, Italy) is injected as a split bolus: the first bolus of 50 mL at a flow rate of 2.5 mL/s and the second bolus of 25 mL 55 seconds later at a flow rate of 3.5 mL/s, followed by a 40 mL saline chaser at a flow rate of 2.5 mL/s. The helical scan starts with a delay of 18 seconds after the start of the second bolus injection.

Image quality of the CECT datasets was rated on a 4-point scale (1, excellent; 2, good; 3, acceptable; 4, not acceptable) to ensure that the dataset allows adequate assessment of the primary site, which is often and primarily affected by metal artifacts. A rating of 4 means that the primary site cannot be evaluated for cancer recurrence.

Inclusion Criteria
1) Status posttreatment of OCSCC or OPSCC and recorded case discussion in our weekly interdisciplinary conference (departments of radiology and of oral and maxillofacial surgery).
2) CECT within 3–12 months after treatment or prior surveillance imaging.
3) CECT imaging-quality requirements.

FIG 1. Flow chart of study participants.
a) Split bolus injection of contrast medium resulting in a combined vascular and delayed phase in 1 acquisition.
b) Arms positioned below the head and neck (next to the chest and abdomen).
c) Image quality rating of 1 (excellent), 2 (good), or 3 (acceptable).
4) Confirmation study either as:
   a) Subsequent surveillance imaging (CECT, MRI, PET) no earlier than 3 months after the CECT study included or
   b) Histopathologic study.

**Exclusion Criteria**

1) Failure to meet CECT quality requirements:
   a) Single bolus injection of contrast medium resulting in a single delayed phase.
   b) Arms positioned over the head.
   c) Image quality rating of 4 (not acceptable).
2) No subsequent confirmation study.

**Readers and Reporting Process**

Four radiologists with different levels of experience (A, 3 years and ∼300 prior reports of neck CECT; B, 4 years and ∼300 reports of neck CECT; C, 7 years and ∼700 reports of neck CECT; D, 15 years and ∼3300 reports of neck CECT) reviewed the 101 cases included in our analysis. Radiologists A and B were grouped as less
Radiologist D is specialized in imaging of the head and neck. At no time were any of the 4 radiologists involved in the interdisciplinary conferences from which patients were included in this study. Anonymized patients were reordered using random numbers assigned by Excel (Version 16.16.10; Microsoft, Redmond, Washington). Readers had access to previous imaging studies (before and after treatment, if available), and they were aware of clinical information to simulate a real reporting situation. Subsequent imaging findings, diagnoses, or clinical examination reports were not available to the 4 readers. After 3 months, radiologists A, B, C, and D were asked again to report on the CECT datasets of the same 101 patients now presented in a newly randomized order. Each of the 2 serial rating sessions was performed in 4 rounds with 25, 25, 25, and 26 patients and a break of 1 week between each round. Another radiologist who was not part of the NI-RADS reader group (E, 6 years of experience and ~400 CECT examinations of the neck) rated the image quality.

**NI-RADS Scoring System**

Reports of imaging findings were based on the NI-RADS White Paper published in 2018, which was well-studied and jointly discussed by our readers and the authors.\(^\text{11}\) NI-RADS scores between 1 and 4, reflecting increasing probabilities of cancer recurrence, are assigned separately for the primary site and for cervical lymph nodes ("neck"). NI-RADS 0 is only used as a preliminary score in cases in which prior images have been obtained but are not available at the time of reading and therefore were not required in our study design. NI-RADS 1 is assigned for expected posttherapeutic changes like the typical superficial diffuse linear contrast enhancement in the primary site and absence of residual abnormal, new, or

---

**FIG 3.** Pretreatment CT (A) of a patient with OCSCC located in the right glossopharyngeal sulcus. Posttreatment CT (B) of the same patient obtained 36 months after resection and neck dissection on the right side (B). A NI-RADS score of 1 was assigned in B for the primary site by all 4 readers. The white arrow indicates the cancer lesion in the primary site (A) and the fatty degenerated muscle flap after resection (B).

**FIG 4.** Posttreatment CTs of a patient with OPSCC located in the left mouth floor obtained 3 months (A) and 15 months (B) after resection and neck dissection on the left side. A NI-RADS score of 4 was assigned in B for the primary site by all 4 readers. Histopathology confirmed recurrence. The white arrow indicates a new enhancing mass in the mouth floor.

**FIG 5.** Posttreatment CTs of a patient with OCSCC located in the anterior mouth floor obtained 12 (A) and 24 (B) months after resection and bilateral neck dissection. The patient’s position differed slightly between the 2 posttreatment CT scans. NI-RADS scores of 2a, 2b, 1, and 1 reflect inconsistent interpretation of the primary site (indicated by the white arrows) in B. Histopathology revealed no malignancy.

**FIG 6.** Pretreatment CT (A) of a patient with OCSCC located in the buccal mucosa in the upper left quadrant. Posttreatment CT (B) of the same patient 3 months after resection and neck dissection on the left side shows an enlarged and necrotic parotid lymph node on the left side as indicated by the white arrows. A NI-RADS score of 3 was assigned in B for the neck by all 4 readers, and histopathology confirmed malignancy.
enlarged lymph nodes in the neck. NI-RADS 2 for the primary site is subdivided into 2a for focal superficial enhancement and 2b for deep, ill-defined enhancement. NI-RADS 2 for the neck indicates residual abnormal or new, enlarged lymph nodes without new necrosis or extranodal extension. NI-RADS 3 is assigned for discrete masses in the primary site and new necrosis or extranodal extension of lymph node involvement in the neck. NI-RADS 4 indicates definitive primary site or nodal radiologically or even histopathologically proved recurrence.

**Data Analysis**

Statistical analysis was performed using R Studio (Version 1.1.383; http://rstudio.org/download/desktop) with the “irr” package installed. The heatmap (Fig 2) was generated using R Studio and the “gplots” package. The flowchart was issued using draw.io (Version 10.8.0; JGraph, Northampton, UK).

Subgroups were formed according to readers’ experience (more-versus-less experienced), the results of the confirmation studies (no recurrence versus recurrence), and the probability of cancer recurrence based on the NI-RADS scores of most readers (NI-RADS 1 and 2 versus NI-RADS 3 and 4).

The Kendall’s W (W) and Fleiss κ (κF) were calculated to test interreader agreement. Calculation of W included a correction factor for tied ranks, and its statistical significance was assessed using the χ2 test. The Kendall's rank correlation coefficient τB and the Cohen weighted κ (κw) were computed to quantify either interreader agreement between 2 readers or intrareader agreement. Calculation of κw provided weighted disagreements according to their squared distance from perfect agreement.

W and τB were interpreted on the basis of the guidelines of Schmidt,15 proposing a 5-step classification: 0.10–0.29, very weak agreement; 0.30–0.49, weak agreement; 0.50–0.69, moderate agreement; 0.70–0.89, strong agreement; 0.90–1.00, very strong agreement. Interpretation of κF and κw followed the recommendations of Landis and Koch:16 < 0.20, slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; 0.81–1.00, (almost) perfect agreement.

Recurrence rates were calculated from the NI-RADS scores of most readers. In case of tied scores, the score assigned by the most experienced reader D was decisive.

**RESULTS**

Figure 2 provides an overview of rating distributions for all 101 patients in the form of a heatmap. It also includes results of the confirmation studies with arrows indicating exemplary cases with perfect or poor agreement among raters. Numbers next to the arrows indicate the figure in which the cases are presented (Figs 3–6).

Depending on the statistical tests used, overall interreader agreement (Table 1) was strong or moderate for both the primary site (W = 0.74, κF = 0.48) and the neck (W = 0.80, κF = 0.50). Less experienced readers showed higher interreader agreement for the primary site (τB = 0.82 versus 0.50, κw = 0.96 versus 0.80) and the neck (τB = 0.96 versus 0.60, κw = 0.99 versus 0.76). Other subgroups were formed according to the results of the confirmation studies. A total of 13 patients were diagnosed with cancer recurrence. Seven patients had simultaneous cancer recurrence at the primary site and in the neck, while 3 patients each had cancer recurrence at the primary site or in the neck. In patients without proved recurrence, interreader agreement was moderate or fair for the primary site (W = 0.56, κF = 0.30) and the neck (W = 0.56, κF = 0.29). By contrast, interreader agreement in patients with proved recurrence was very strong or substantial for the primary site (W = 0.96, κF = 0.65) and strong or moderate for the neck (W = 0.78, κF = 0.41). When forming merged NI-RADS categories according to high and low suspicion of cancer recurrence, we found higher interreader agreement for NI-RADS 3/4 than NI-RADS 1/2 for both the primary site (W = 0.85 versus 0.51, κF = 0.56 versus 0.23) and the neck (W = 0.59 versus 0.56, κF = 0.44 versus 0.26).

Intrareader agreement (Table 2) for the primary site ranged from moderate to strong (τB = 0.67–0.82) or almost perfect (κw = 0.85–0.96). Intrareader agreement for the neck was strong (τB = 0.76–0.86) or almost perfect (κw = 0.89–0.95).

All statistical analyses conducted to test inter- and intrareader agreement showed statistical significance (P < .05).

Recurrence rates (Table 3) were between 3.57% (NI-RADS 1) and 100% (NI-RADS 4) for the primary site and 0% (NI-RADS 1) and 83.33% (NI-RADS 4) for lymph nodes (Table 3). Patients without histopathology for confirmation of their diagnosis were followed up for a median of 351 days (range, 159–772 days), defined by the date of their last surveillance imaging study.

**DISCUSSION**

Inter- and intrareader agreement is important for estimating the reliability of any diagnostic test. To the best of our knowledge, a study investigating inter- and intrareader agreement of NI-RADS scores has not been published. However, we can discuss our results for NI-RADS with those other investigators’ results obtained for the reliability of RADS in other organs. Published data give a very diverse picture. A study similar to ours in terms

---

**Table 1: Interreader agreement**

<table>
<thead>
<tr>
<th>Group</th>
<th>Primary Site</th>
<th>Neck</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>W ϱB κF κw</td>
<td>W ϱB κF κw</td>
</tr>
<tr>
<td>Overall</td>
<td>0.74</td>
<td>0.50</td>
</tr>
<tr>
<td>0.80</td>
<td>0.60</td>
<td>0.76</td>
</tr>
<tr>
<td>More experienced readers</td>
<td>0.50</td>
<td>0.80</td>
</tr>
<tr>
<td>Less experienced readers</td>
<td>0.82</td>
<td>0.96</td>
</tr>
<tr>
<td>No recurrence</td>
<td>0.56</td>
<td>0.30</td>
</tr>
<tr>
<td>Recurrence</td>
<td>0.96</td>
<td>0.65</td>
</tr>
<tr>
<td>NI-RADS 1 and 2 merged</td>
<td>0.51</td>
<td>0.23</td>
</tr>
<tr>
<td>NI-RADS 3 and 4 merged</td>
<td>0.85</td>
<td>0.56</td>
</tr>
</tbody>
</table>

**Table 2: Intrareader agreement**

<table>
<thead>
<tr>
<th>Reader</th>
<th>Primary Site</th>
<th>Neck</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ϱB κF κw</td>
<td>ϱB κF κw</td>
</tr>
<tr>
<td>A</td>
<td>0.73</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>0.86</td>
<td>0.93</td>
</tr>
<tr>
<td>B</td>
<td>0.67</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>0.86</td>
<td>0.95</td>
</tr>
<tr>
<td>C</td>
<td>0.82</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>0.86</td>
<td>0.94</td>
</tr>
<tr>
<td>D</td>
<td>0.76</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>0.76</td>
<td>0.89</td>
</tr>
</tbody>
</table>
of statistical methods and results was published by Irshad et al,\textsuperscript{16} who assessed consecutive versions of BI-RADS including 5 readers and 104 mammographic examinations. They found an overall interreader agreement of 0.65 and 0.57 (Fleiss $\kappa$), while overall intrareader agreement was 0.84 and 0.78 (Cohen weighted $\kappa$). A study by Smith et al\textsuperscript{17} determined the reliability of PI-RADS in the interpretation of multiparametric MR imaging of the prostate, including 4 readers and 102 examinations, again similar to our study design. However, by contrast, they reported an overall interreader agreement of 0.24 (Fleiss $\kappa$) and an overall intrareader agreement of 0.43–0.67 (Cohen $\kappa$).

When we compared the 2 studies with our results, the difference in overall interreader agreement stood out first. Our results obtained with NI-RADS ($\kappa_F = 0.48$ and 0.50) are much better than findings reported by other investigators for PI-RADS but inferior to results achieved with BI-RADS. NI-RADS showed a very high intrareader agreement ($\kappa_w = 0.85–0.96$ and $\kappa_{wN} = 0.89–0.95$), especially against the poor values obtained in the PI-RADS study. Thus, our results are encouraging because they suggest that there is the potential for improving interreader agreement. Given that the NI-RADS lexicon and decision tree can only be used fully when interpreting PET/CT or PET/MR imaging, we expect that interreader agreement can be considerably improved using either of these modalities. Especially, NI-RADS categories 1 and 2 (2a and 2b) are defined more clearly when additional information on FDG uptake is available.

Apart from our findings regarding absolute overall agreement, our analysis also provides some interesting results regarding the subgroups formed. Unexpectedly, overall interreader agreement for both the primary site and the neck was higher between the 2 less experienced readers than between the 2 more experienced readers. Furthermore, interreader agreement for the absence of recurrence in lymph nodes was poorer than we expected. A possible explanation emerged from discussions with the readers after completion of the study: The definition for assigning a lymph node to NI-RADS 2 is “mildly enlarging without specific morphologically abnormal features such as new necrosis or extracapsular spread,” which was perceived as rather vague.\textsuperscript{11} Some kind of measurable threshold might significantly increase agreement among raters. Other results of our study suggest adequate sensitivity of NI-RADS. Interreader agreement was significantly higher in cases of proved cancer recurrence compared with patients without recurrence.

Coincidentally low recurrence rates in the group classified as NI-RADS 1 as well as high recurrence rates in groups with NI-RADS scores of 3 and 4 suggest that NI-RADS is a powerful tool for discrimination of patients with a low-versus-high risk of cancer recurrence. No patients assigned scores of 2a for the primary site had cancer recurrence, which might be attributable to the relatively small number of cases or greater variability in the interpretation of findings, as already discussed above. Recurrence rates calculated in our study are based on majority decision but align very well with initially published data.\textsuperscript{11,18,19}

While calculation of $\kappa$ coefficients is by far the most common statistical test to quantify inter- and intra-reader agreement,\textsuperscript{20,21} there are also more differentiated approaches addressing other aspects of inter- and intrareader agreement.\textsuperscript{22} Other investigators primarily recommend $\kappa$ statistics for testing nominal scaled data.\textsuperscript{23,24} From our standpoint, NI-RADS scores should be regarded as ordinal data because rising values represent a rising probability of cancer recurrence. Therefore, the Kendall's coefficient of concordance (used to determine interreader agreement for $>2$ readers) and the Kendall's rank correlation coefficient (interreader agreement with 2 raters or intrareader agreement) should be most appropriate.\textsuperscript{25} When we compared the result pairs of statistical methods in our study, it is apparent that values of $W$ are always higher than those of $\kappa_F$ but values of $\tau_W$ are always lower than those of $\kappa_w$, while their relationships stay basically constant. The intraclass correlation is also used to determine inter- and intra-reader agreement; however, it should only be used for underlying continuous data. We therefore chose not to calculate intraclass correlation statistics for the discrete data provided by NI-RADS.

This study, although retrospective, was designed to put readers in a real-world clinical reporting situation. This means that the readers had access to information on OCSCC/OPSCC localization as defined by the multidisciplinary cancer conference, surgical and radiotherapeutic procedures, and pre-existing illnesses. This information is available to reporting radiologists in the clinical setting and is important for appropriately and comprehensively interpreting imaging findings and assessing the patient’s condition. On the other hand, there were actions to reduce possible bias. Cases were presented in randomized order, and anonymization of patient data was performed to lower a possible detection bias. The 101 CECT datasets were split into 4 rating sessions (25, 25, 25, and 26) to minimize possible over- or underratings because of readers’ raised awareness and altered perception of similarities and differences when comparing cases with others they have recently seen in the artificial reading situation.

Clinically suspected OCSCC or OPSCC and posttherapeutic surveillance are the most frequent indications for neck imaging in our institution, with CECT being much more commonly used than MR imaging. Future studies should investigate inter- and intrareader agreement of NI-RADS, not only for other malignancies (eg, larynx and salivary glands) but also for different imaging modalities (CECT, MR imaging, PET/CT and PET/MR imaging). The role of PET/CT and PET/MR imaging in up- or downgrading lesions seen on CECT or MR imaging without PET should also be of interest in studies, especially prospectively designed, studies.

### Table 3: Score counts and recurrence rates for each category based on majority decision*

<table>
<thead>
<tr>
<th>NI-RADS Score</th>
<th>Score Count</th>
<th>Recurrence</th>
<th>No Recurrence</th>
<th>Recurrence Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary 1</td>
<td>84</td>
<td>3</td>
<td>81</td>
<td>3.57%</td>
</tr>
<tr>
<td>Primary 2a</td>
<td>6</td>
<td>0</td>
<td>6</td>
<td>0%</td>
</tr>
<tr>
<td>Primary 2b</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>33.33%</td>
</tr>
<tr>
<td>Primary 3</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>50%</td>
</tr>
<tr>
<td>Primary 4</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td>Node 1</td>
<td>83</td>
<td>0</td>
<td>83</td>
<td>0%</td>
</tr>
<tr>
<td>Node 2</td>
<td>7</td>
<td>1</td>
<td>6</td>
<td>14.29%</td>
</tr>
<tr>
<td>Node 3</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>80%</td>
</tr>
<tr>
<td>Node 4</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>83.33%</td>
</tr>
</tbody>
</table>

*The most experienced reader D was decisive in case of tied score counts.
**Limitations**

Four radiologists reported imaging findings in this study. While radiologists A and B were relatively close in terms of work experience (years and number of examinations), C and D were wider apart. Although C could easily be classified as more experienced than A and B, a work experience closer to D would have been desirable to ensure ideally balanced subgroups. Subdividing readers into 3 groups with an additional group of intermediate experience might also yield interesting additional results. Because we just started to integrate NI-RADS as a reporting system in our institution, future studies could address these limitations. As readers become more familiar with using NI-RADS and shared experience grows, common approaches might emerge and improve interreader agreement. Although all 4 radiologists were well-acquainted with the literature on NI-RADS, a joint discussion of exemplar cases from our department might have improved interreader and even intrareader agreement. Beyond that, in our opinion, more experience might also lead to higher rates of NI-RADS a/b scores being assigned because findings in this category are more difficult to express in prosaic reports because referring clinicians expect a clear decision between “suspected recurrence” versus “no suspected recurrence.” We determined recurrence rate as a secondary outcome. Although it attests to the good discriminatory power of NI-RADS, future studies investigating the validity of NI-RADS should define a longer follow-up period of at least 1 year.

**CONCLUSIONS**

NI-RADS used for interpreting CECT after treatment of OCSCC and OPSCC provides acceptable score reproducibility. A major strength of this standardized approach is the good interreader agreement in patients with proved cancer recurrence and overall intrareader agreement in general. At the same time, there are limitations in terms of interreader agreement in patients with post-therapeutic changes but no cancer recurrence. Although only determined as secondary outcomes, recurrence rates in our patients were similar to those in preliminary published data.


**REFERENCES**

15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74 CrossRef
Tumor Response Assessment in Diffuse Intrinsic Pontine Glioma: Comparison of Semiautomated Volumetric, Semiautomated Linear, and Manual Linear Tumor Measurement Strategies

L.A. Gilligan, M.D. DeWire-Schottmiller, M. Fouladi, P. DeBlank, and J.L. Leach

ABSTRACT

BACKGROUND AND PURPOSE: 2D measurements of diffuse intrinsic pontine gliomas are limited by variability, and volumetric response criteria are poorly defined. Semiautomated 2D measurements may improve consistency; however, the impact on tumor response assessments is unknown. The purpose of this study was to compare manual 2D, semiautomated 2D, and volumetric measurement strategies for diffuse intrinsic pontine gliomas.

MATERIALS AND METHODS: This study evaluated patients with diffuse intrinsic pontine gliomas through a Phase I/II trial (NCT02607124). Clinical 2D cross-product values were derived from manual linear measurements (cross-product = long axis x short axis). By means of dedicated software (mint Lesion), tumor margins were traced and maximum cross-product and tumor volume were automatically derived. Correlation and bias between methods were assessed, and response assessment per measurement strategy was reported.

RESULTS: Ten patients (median age, 7.6 years) underwent 58 MR imaging examinations. Correlation and mean bias (95% limits) of percentage change in tumor size from prior examinations were the following: clinical and semiautomated cross-product, \( r = 0.36, -1.5\% \) \((-59.9\%, 56.8\%\)); clinical cross-product and volume, \( r = 0.61, -2.2\% \) \((-52.0\%, 47.8\%\)); and semiautomated cross-product and volume, \( r = 0.79, 0.6\% \) \((-39.3\%, 38.1\%\)). Stable disease, progressive disease, and partial response rates per measurement strategy were the following: clinical cross-product, 82\%, 18\%, 0\%; semiautomated cross-product, 54\%, 42\%, 4\%; and volume, 50\%, 46\%, 4\%, respectively.

CONCLUSIONS: Manual 2D cross-product measurements may underestimate tumor size and disease progression compared with semiautomated 2D and volumetric measurements.

ABBREVIATIONS: CP = cross-product; DIPG = diffuse intrinsic pontine glioma; PD = progressive disease; PR = partial response; SD = stable disease; TRC = tumor response criteria

Diffuse intrinsic pontine gliomas (DIPGs) comprise 80% of brain stem gliomas, which, in turn, account for 10%–20% of central nervous system tumors in children.\(^1\)\(^2\) DIPG carries a dismal prognosis, with a mean survival of 11 months and 1-, 2-, and 5-year overall survival rates of 42\%, 10\%, and 2\%, respectively.\(^3\) Standard of care for DIPG consists of involved field radiation therapy, lengthening survival by an average of 3–4 months.\(^3\)\(^4\) Recently, it has been discovered that up to 80% of DIPGs have a pathognomonic point mutation in histone H3.3 (H3F3A) (65% of tumors) or histone H3.1 (HIST1H3B) (25% of tumors), the latter conferring longer survival in most studies assessing this.\(^5\) Chemotherapeutic agents have failed to demonstrate efficacy, and no improvement in survival has been achieved in the past 4 decades.\(^3\)\(^5\)\(^6\)\(^7\) Currently, there are 69 interventional research studies for DIPG listed on ClinicalTrials.gov.


Please address correspondence to James L. Leach, MD, Department of Radiology, MLC 5031, Cincinnati Children’s Hospital Medical Center, 3333 Burnet Ave, Cincinnati, OH 45229; e-mail: james.leach@ccchmc.org

http://dx.doi.org/10.3174/ajnr.A6555
Most therapeutic trials for DIPG have assessed treatment response with Response Assessment in Neuro-Oncology, MacDonald, or World Health Organization criteria (or modifications), which use a 2D measurement of tumor size on MR imaging, allowing comparison with historical data.8-11 While studies have demonstrated good correlation between 2D and volumetric measurements in high-grade gliomas, there is a lack of comparison data in tumor size measurement strategies in patients with brain stem gliomas.12-14 Furthermore, high interobserver variability among 2D measurements in DIPG has been observed.15 With increasing availability and capability of novel tumor segmentation software, semiautomated tumor volumetry is a potentially useful assessment tool that may be more sensitive to tumor response and enable earlier determination of treatment efficacy. However, research is needed to define therapeutic end point criteria before incorporation into clinical trials.

The purpose of this study was to compare methods of DIPG tumor measurement, including traditional manual 2D measurements and semiautomated, software-assisted 2D measurements and tumor volumes. A secondary aim was to explore the implications of using software-assisted 2D measurements and tumor volumes for response assessment, compared with the standard manual 2D method.

**MATERIALS AND METHODS**

The institutional review board at Cincinnati Children’s Hospital Medical Center approved this study as part of a Phase I/II drug trial of ribociclib following radiation therapy in patients with newly diagnosed DIPG (NCT02607124). Patients between 1 and 30 years of age with nonbiopsied DIPG, retinoblastoma gene mutation–positive DIPG, or retinoblastoma gene mutation–positive high-grade glioma were prospectively recruited to undergo MR imaging before and following drug therapy between April 2016 and November 2017 as part of the treatment protocol.16 Written informed consent was obtained. MR imaging examinations were typically obtained before cycles 3, 5, 7, 9, and 11.

**Imaging Protocol**

Imaging examinations were performed on 3T scanners (Signa Architect, GE Healthcare). Imaging protocol included the following: volumetric T1WI and T2-FLAIR (1-mm isotropic, axial, and coronal reformations; FOV = 256 mm, matrix = 256 × 256, slice thickness = 1 mm), axial T2WI (FOV = 220, matrix = 512 × 224, slice thickness = 3 mm contiguous), DTI and SWI sequences. IV gadoterate meglumine (Dotarem; Guerbet) was injected (rate = 1 mL/s, dose = 0.1 mmol/kg). Postcontrast T1WI (volumetric, 1 mm) and axial T1-FLAIR (FOV = 220 mm, matrix = 320 × 288, slice thickness = 3 mm contiguous) sequences were obtained.

**Tumor Measurements**

For clinical purposes per the study protocol, 2D tumor measurements were made by 1 of 7 fellowship-trained, pediatric neuroradiologists at Cincinnati Children’s Hospital Medical Center, including the largest diameter in the axial plane (long axis) and a measurement perpendicular to the long axis (short axis). Clinical cross-product (CP) was calculated by multiplying long and short axes (Fig 1).
For this study, a single reviewer (J.L.L.), a pediatric neuroradiologist with 25 years of experience, blinded to the clinical measurements, manually segmented each entire tumor by tracing tumor margins on axial T2WI and T2-FLAIR images using proprietary, clinically-available software (mint Lesion, Version 3.4.5; Mint Medical). From the outlined tumor, the largest axial diameter (long axis) and the measurement perpendicular to the long axis (short axis) were automatically derived. Semiautomated CP was calculated by multiplying the long and short axes. From the outlined tumor, tumor volume was also automatically derived (Fig 1).

**Classification of Tumor Response**

Per the study protocol, the following imaging-related tumor response criteria (TRC) were used to categorize each follow-up examination (in regard to tumor size): complete response, complete disappearance of all tumor and mass effect, maintained for 8+ weeks; partial response (PR), ≥50% decrease from baseline; stable disease (SD), <50% decrease from baseline size and <25% increase from the prior lowest tumor size; and progressive disease (PD), ≥25% increase from the prior lowest tumor size.¹¹,¹⁷ These criteria were applied to 2D and volumetric data. Additionally, we applied another TRC to the volume data only, following previously published recommendations: PR, ≥65% decrease from baseline; SD, <65% decrease from baseline or <40% increase from the prior lowest tumor size; and PD, ≥40% increase from the prior lowest tumor size.¹⁸,¹⁹ These published TRC for volumetric measurements are extrapolated from linear values using a spheric tumor model. Currently, there are no published prospective studies regarding TRC for tumor volumes in pediatric brain tumors.¹⁸

**Statistical Analysis**

Continuous data were summarized as means and SDs or medians and ranges; categoric data were summarized as counts and percentages. Correlation coefficients were used to compare clinical and semiautomated CP measurements and the percentage change in tumor size (compared with a prior examination) among the 3 measurement strategies. Bland-Altman analyses were performed to assess bias among measurement strategies. Descriptive statistics were used to describe percentage change in tumor size per TRC classification and measurement strategy.

P values < .05 were considered statistically significant for inference testing. Correlation coefficients were classified by the following definitions: 0–0.19, very weak; 0.20–0.39, weak; 0.40–0.59, moderate; 0.60–0.79, strong; and 0.80–1.0, very strong.²⁰ Analyses were performed using MedCalc for Windows (MedCalc Software).

**RESULTS**

Ten patients were included, all with DIPGs. The median age at baseline MR imaging was 7.6 years (range, 3.9–20.1 years), and 6 patients (60%) were female. The median number of follow-up

---

**FIG 2.** Sample case demonstrating clinical 2D measurements (clinical CP) and semiautomated 2D and volumetric measurements (semiautomated CP) during the treatment course. In this case, although there were differences in orientation of the measurements with the semiautomated process, response classification was the same compared with manual clinical CP measurements.

**FIG 3.** Sample case demonstrating clinical 2D measurements (clinical CP) and semiautomated 2D and volumetric measurements (semiautomated CP) during the treatment course. Note that in this case, although there were differences in orientation of the measurements with the semiautomated process, response classification was the same compared with manual clinical CP measurements.
follow-up MR imaging of 177 ± 98 days (range, 56–411 days). One patient (with 2 follow-up MRIs) had 2 discrete tumors, each measured and analyzed separately. Case examples of tumor measurements over the study course are illustrated in Figs 2 and 3.

**Correlation and Bias between Tumor Measurement Strategies**

Clinical and semiautomated tumor CP measurements were strongly correlated ($r = 0.74$, $P < .001$), with a mean bias of $-2.5$ mm$^2$ (95% limits of agreement, $-9.5$, $+4.4$ mm$^2$) (Figs 4 and 5). There were strong, statistically significant correlations between the percentage change in tumor size between clinical CP and volume ($r = 0.61$, $P < .001$) and semiautomated CP and volume-measurement strategies ($r = 0.79$, $P < .001$). Correlation and mean bias of percentage change in tumor size between clinical and semiautomated CP measurement strategies are depicted in Figs 6 and 7. There was a weak, statistically significant correlation ($r = 0.36$, $P = .011$) of percentage change from prior examinations comparing clinical and semiautomated CP measurements.
Comparison of Treatment Response Assessment between Tumor Measurement Strategies

Classification of percentage change of tumor size from baseline or nadir examination, per the protocol TRC, yielded 25/50 (50%) cases that were classified concordantly across all 3 tumor measurement methods. Concordance was 32/50 (64%) between clinical CP and volume, and 38/50 (76%) between semiautomated CP and volume. Frequencies of SD, PD, and PR by tumor-measurement strategy are reported in Table 1. No examination or method demonstrated a complete response. Of note, 34% (14/41) of time points classified as SD by clinical CP were classified as PD by semiautomated CP measurements, with a mean bias of $-1.5\%$ (95% limits of agreement, $-59.9\%, +56.8\%$).

**FIG 6.** Correlation of percentage change from prior examination in clinical CP versus semiautomated CP ($r = 0.36, P = .011$).

**FIG 7.** Bland-Altman plot demonstrating bias between the percentage change in clinical and semiautomated tumor CP from a prior examination. The solid line indicates mean bias between techniques. Dashed lines indicate ±2 SDs of the mean (95% limits of agreement). Overall, percentage change in clinical CP was smaller than the percentage change in semiautomated CP (mean bias, $-1.5\%$, 95% limits of agreement, $-59.9\%, +56.8\%$) between time points.
CP. The mean percentage change in tumor volume from follow-up examination to baseline or nadir per response category and tumor measurement strategy is reported in Table 2.

**DISCUSSION**

Recent research effort has aimed at improving and discovering imaging biomarkers of DIPG. Such work is relevant for the standardization of clinical trial end points and improved detection of treatment effect. Historically, DIPGs have been measured using MacDonald, World Health Organization, or Response Assessment in Neuro-Oncology criteria (or modifications). These criteria use a 2D measurement strategy in which two perpendicular measurements are made on the image with the largest observed cross-sectional area of tumor. However, these standard tumor measurements were not developed and may not be best-suited for DIPG, and there remains an overall lack of standardization in DIPG measurements.

2D and 3D DIPG tumor measurements have demonstrated poor interobserver agreement, possibly related to their infiltrative nature, indiscernible borders, and heterogeneous appearance.15 Furthermore, 1D, 2D, and 3D tumor measurements of DIPG have not correlated well with clinical outcomes.21-24 Other surrogate imaging biomarkers of DIPG have been explored, including metabolic ratios by MR spectroscopy, tumor perfusion by dynamic susceptibility contrast MR imaging, and pontine size by conventional MR imaging.24-28 Recently, studies have investigated tumor volume measurement strategies (automated, semiautomated, or manual) and have shown them to be a promising tool with improved inter- and intraobserver agreement.18,29,30 It has been hypothesized that a volume measurement strategy may be more appropriate for DIPG, given its often complex morphology.

In this study, we have demonstrated a strong correlation and relatively small bias between manual and semiautomated (from mint Lesion software) CP values. However, the correlation of percentage change in tumor size (the metric by which treatment response is assessed) between these strategies was weak. In some cases, there was a large discrepancy between the percentage change in tumor size across the clinical CP, semiautomated CP, and volumetric measurement strategies. Generally, the mint Lesion–derived measurements (semiautomated CP, volume) tended to classify tumors as PD compared with the clinical CP strategy, which tended to classify tumors as SD.

There are several potential explanations for these discrepancies. DIPG can have variable, nonspheric morphology, and some shapes are not well-approximated by 1 CP measurement. In addition, tumor growth is constrained by anatomic boundaries to some degree and may be less constrained in certain directions. More pronounced growth along the cranial-caudal axis, for example, may not be captured by performing measurements only in the transverse plane. Furthermore, clinical CP measurements are likely biased toward the radiologist by using a section location and measurement orientation similar to those used on prior examinations. This bias may result in underestimation of tumor size if greater interval growth occurred at a different axial section or in a different orientation than the manual measurements. This bias is also supported by our results: mint Lesion automatically derives the largest CP, which explains why these measurements were, on average, 2.5 mm² larger than the clinical CP measurements. Our results have important implications for the use of

---

**Table 1:** Response assessment classifications of 50 MR imaging time points for 3 tumor-management strategies

<table>
<thead>
<tr>
<th>TRC</th>
<th>Clinical CP</th>
<th>Semiautomated CP</th>
<th>Volume (25/50)</th>
<th>Volume (40/65)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PR</td>
<td>SD</td>
<td>PD</td>
<td>PR</td>
</tr>
<tr>
<td>PD</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SD</td>
<td>2</td>
<td>25</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>PD</td>
<td>0</td>
<td>2</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Concordance</td>
<td>64%</td>
<td>60%</td>
<td>76%</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2:** Descriptive statistics of the percentage change in tumor size from baseline or nadir to follow-up examination (assigned per clinical protocol) for each response assessment classification per tumor measurement strategy

<table>
<thead>
<tr>
<th>TRC</th>
<th>Mean (%)</th>
<th>Standard Deviation (%)</th>
<th>Minimum (%)</th>
<th>Maximum (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CPc</td>
<td>CPsa</td>
<td>Vol1</td>
<td>Vol2</td>
</tr>
<tr>
<td>PD</td>
<td>57</td>
<td>83</td>
<td>75</td>
<td>79</td>
</tr>
<tr>
<td>SD</td>
<td>6</td>
<td>14</td>
<td>16</td>
<td>8</td>
</tr>
</tbody>
</table>

---

*Note:*—CPc indicates clinical cross-product; CPsa, semiautomated cross-product; Vol, volume (25/50% tumor response criteria); Vol, volume (40/65% tumor response criteria).
segmentation software in clinical practice. Further study is needed to determine which method is a better indicator of clinical outcome.

We found the strongest correlation in percentage change of tumor size between the 2 methods by using mint Lesion software (semiautomatic CP and volume). This is not surprising because the same segmented images (derived from 1 reviewer who traced tumor margins in all slices) were used for these 2 strategies. Although we found that the bias between these techniques was very low in magnitude (0.6%), the limits of agreement remained wide (−39%, +38%), which could be related to our small sample size. Furthermore, we did not have adequate power to define new cutoff values for tumor response–assessment classes using tumor volume measurements. If volumetric measurements of DIPG should be used for assessing tumor response, identifying relationships with linear measurements (to inform comparisons with prior studies and treatment trials) and clinical progression is critical to ensure maximum utility. Volumetric end points derived from spheric/elliptic mathematic models using cranial-caudal dimensions may not be applicable to tumors with nonspheric growth patterns.30 Further research using larger groups of subjects and treatment time points is needed.

Correlations between true volumetric measurements and 2D measurements of brain tumors are few, particularly in children. One study of low-grade pediatric gliomas showed that 20% of MR imaging examinations demonstrated discordant response assessments between 2D and volumetric tumor measurement strategies.18 Shah et al,13 in a study of adult patients with glioblastomas, also demonstrated a 20% discordance rate in response assessments when comparing 2D and volumetric measurements. These discrepancy rates are quite similar to the 24% discrepancy rate between 2 techniques (clinical CP and volume) demonstrated in our study.

Our study was limited by several factors. First, we included a small sample size of patients with DIPG, inherent to the nature of this single-center study of a disease with low prevalence. Second, although the clinical trial protocol standardized the timing of MR imaging examinations, the actual timing and total number of MR imaging examinations varied across patients, often related to additional MR imaging examinations being performed when patients had a change in clinical status. Third, volumetric tumor measurements using the mint lesion software were made by only 1 reviewer; thus, we cannot draw conclusions regarding interobserver variability of the tumor measurements with this technique. Fourth, the software used for 2 of the measurement strategies in our study (semiautomatic CP, volume) may not be widely available, potentially limiting the application of our results. Finally, because the clinical outcome data for this trial remain unpublished, our study did not correlate tumor measurements or response assessment with clinical outcomes. This would be an important area of future investigation.

CONCLUSIONS

We have shown that correlation of change in DIPG tumor size among 3 measurement strategies is variable, with the strongest correlation observed between semiautomated 2D and volumetric strategies and the weakest correlation observed between clinical 2D and semiautomated 2D strategies. The conventional method of manual 2D (cross-product) tumor measurement likely underestimates tumor size and disease progression compared with semiautomated 2D measurements. Application of semiautomated 2D and volumetric measurements in therapeutic trials will alter response assessment compared with standard 2D measures in DIPG. Further research is needed to outline relationships among these methods, clinical signs of progression, and survival.


REFERENCES


Diffuse intrinsic pontine glioma (DIPG) is a malignant brain tumor that accounts for 75%–80% of brain stem tumors in children. Despite medical therapy, these tumors have a poor prognosis, with a median survival at 1, 2, and 3 years of 41%, 15.3%, and 7.2%, respectively. The 2016 World Health Organization classification of brain tumors introduced the term “diffuse midline glioma, H3 K27M mutant” to describe these neoplasms due to the high prevalence of unique histone protein mutations that differ from those of supratentorial (hemispheric) pediatric high-grade gliomas and adult high-grade gliomas. While H3 K27M is the most common variant detected, more recently, other distinct histone mutations have been identified in DIPG, which appear to correspond to subgroups with
different prognosis and pathologic phenotypes. Identification of these subgroups of tumors is important to appropriately counsel families and tailor treatment strategies.

Radiogenomic techniques are emerging as valuable tools to noninvasively characterize brain tumors and are increasingly important, given the evolving molecular landscape in pediatric neuro-oncology. Although most data available to date pertain to adult high-grade gliomas, radiogenomic techniques have been successfully implemented to characterize atypical teratoid/rhabdoid tumors and medulloblastomas. A challenge to molecular analysis of DIPGs is that historically, patients with these tumors did not routinely undergo biopsy due to a perceived high likelihood of morbidity associated with the procedure; consequently, most imaging-based studies do not have pathologic or molecular correlates. Despite this limitation, important imaging-based prognostic factors have been described for DIPG. Jansen et al. and Hoffman et al. reported that ring enhancement was associated with shorter survival. Chen et al. reported worse outcome in DIPGs with lower ADC values at baseline. Poussaint et al. reported that lower diffusion values on ADC histogram analysis, high skewness, and enhancement were associated with shorter survival. The association among these MR imaging features, the molecular subgroups, and various histone mutations has not been elucidated.

The purpose of this study was to investigate the relationship between the molecular subgroup and histone mutations of DIPGs and MR imaging features on anatomic sequences, and ADC histogram analysis, using the baseline MR imaging from a prospective clinical trial of children with newly diagnosed DIPG who underwent biopsy. 

**MATERIALS AND METHODS**

**Subjects**

Patients for this study were recruited as part of an institutional review board (Dana Farber Cancer Institute)–approved, Health Insurance Portability and Accountability Act–compliant, multicenter clinical trial: Molecularly Determined Treatment of Diffuse Intrinsic Pontine Gliomas (DIPG) (NCT01182350). Prospective subjects had MR imaging that showed classic clinical-imaging criteria of nonmetastatic DIPG, including an expansile and infiltrative T1-isointense or -hypointense and T2-hyperintense mass centered within the pons, encompassing at least 50% of the pons with no or little contrast enhancement, which could vary. Findings on clinical examination included multiple cranial nerve neuropathies, long tract signs, and ataxia. Detailed criteria for inclusion and exclusion are summarized in On-line Table 1. The patients enrolled in the trial underwent a biopsy before initiation of therapy; the biopsy target and approach were at the discretion of the neurosurgeon, with the goal of minimizing operative risk. All subjects received local radiation therapy (59.4 Gy) with adjuvant bevacizumab, as per the standard of care. Subsequently, patients were stratified into different treatment arms with erlotinib and/or temozolomide at the start of radiation therapy based on MGMT promoter methylation status and epidermal growth factor receptor (EGFR) expression in tumor tissue obtained from surgical biopsies.

Of 53 patients enrolled, 50 underwent biopsy and were included in the analytic cohort of this study. Subjects without biopsy were excluded from this analysis. We performed a retrospective review that included biopsy data (histone mutations, molecular groups), MR imaging features on baseline MR imaging, and clinical outcomes. The molecular analysis determined the MGMT methylation status and expression of EGFR; 4 categories were derived from molecular analysis: MGMT−/EGFR−, MGMT−/EGFR+, MGMT+/EGFR−, MGMT+/EGFR+. Three histone mutations were identified through whole-genome RNA sequencing: H3F3A, HIST1H3B, and HIST1H3C. Clinical end points of the study were determined with longitudinal follow-up. Overall survival (OS) was measured from the time of registration to death or to last follow-up if censored. Progression-free survival (PFS) was measured from the time of registration to progression or death (whichever was first) or to the last follow-up if censored.

**Image Analysis**

The MR imaging of the brain used for analysis corresponded to the baseline examination obtained before treatment and biopsy. Images were obtained using 3T (n = 17) and 1.5T (n = 33) scanners. A standard clinical protocol was used, which included the following: sagittal T1, axial T2, precontrast axial T2-FLAIR, axial DTI (TR = 6500 ms, TE = 88 ms, section thickness = 2 mm, b-values = 0 and 1000 s/mm², 30 gradient directions), and postgadolinium (PG) 3D T1.

Volumetric data of various components of the tumor were obtained by generating 3D ROIs on a Vitrea workstation (Vital Images) on the anatomic sequences. Specifically, T2-FLAIR/T2 volume represents the volume of tissue (milliliter) with abnormal T2 hyperintense signal on a 2D TSE T2-FLAIR sequence; for cases in which T2-FLAIR was unavailable, this parameter was estimated using a 2D T2 TSE sequence. Tumor volume enhancing represents the volume (milliliter) of enhancement within the tumor estimated on the PG 3D T1 sequence.

---

**Table 1. Baseline patient characteristics (n = 50)**

<table>
<thead>
<tr>
<th>Patient Characteristic</th>
<th>No. (%) or Median (Range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of MR imaging studies</td>
<td>50 (100)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>24 (48)</td>
</tr>
<tr>
<td>Female</td>
<td>26 (52)</td>
</tr>
<tr>
<td>Median age at registration (range) (yr)</td>
<td>6 (3.3–17.5)</td>
</tr>
<tr>
<td>Molecular subgroups</td>
<td></td>
</tr>
<tr>
<td>MGMT−/EGFR−</td>
<td>28/48 (58)</td>
</tr>
<tr>
<td>MGMT−/EGFR+</td>
<td>14/48 (29)</td>
</tr>
<tr>
<td>MGMT+/EGFR−</td>
<td>3/48 (6)</td>
</tr>
<tr>
<td>MGMT+/EGFR+</td>
<td>3/48 (6)</td>
</tr>
<tr>
<td>Unassigned</td>
<td>2</td>
</tr>
<tr>
<td>Mutational status</td>
<td></td>
</tr>
<tr>
<td>Any histone mutation</td>
<td>34/49 (69)</td>
</tr>
<tr>
<td>Wild-type</td>
<td>15/49 (31)</td>
</tr>
<tr>
<td>Histone mutation (among patients with any histone mutations)</td>
<td></td>
</tr>
<tr>
<td>H3F3A</td>
<td>23/34 (68)</td>
</tr>
<tr>
<td>HIST1H3B</td>
<td>8/34 (24)</td>
</tr>
<tr>
<td>HIST1H3C</td>
<td>3/34 (9)</td>
</tr>
<tr>
<td>Median follow-up time (range) (mo)</td>
<td>10.9 (0.4–33.4)</td>
</tr>
</tbody>
</table>

*One patient with both H3F3A and HIST1H3B histone mutations was excluded.**
To perform ADC histogram analysis, we registered the ADC maps to the anatomic sequence of interest (T2-FLAIR/T2 and PG T1) using tools from the FSL library (http://www.fmrib.ox.ac.uk/fsl),\textsuperscript{15} as described by Poussaint et al.\textsuperscript{12} Briefly, we transformed the $b=0$ (and subsequently the ADC map) into the space of the individual anatomic sequences. Subsequently, 3D ROIs were created using the thresholding feature in Fiji (http://fiji.sc/),\textsuperscript{16} an Open Source (https://opensource.org/) distribution of Java modules along with ImageJ software (National Institutes of Health); we then extracted the values from every pixel from the ADC map within the corresponding ROI.

We applied a uniform threshold to the ADC maps, ranging from 600 to $2600 \times 10^{-6}$ mm$^2$/s to differentiate solid tumor from cyst/necrosis. ADC histograms were then generated using a uniform bin width of 10 mm$^2$/s. Histogram metrics used for statistical analysis were mean, median, mode, skewness, and kurtosis of these histograms.

**Statistical Analysis**

Clinical and demographic variables were summarized using descriptive statistics. Continuous measures were summarized using medians and ranges, and categoric measures were summarized.
using frequencies and proportions. Kaplan-Meier curves were used to summarize OS and PFS for all patients. Univariate Cox proportional hazards regression was used to test the association of imaging predictors with OS and PFS. The Wilcoxon rank sum and Kruskal-Wallis tests were used to compare continuous imaging features among groups. The Fisher exact test was used to compare categoric imaging features among groups. When applicable, the presence of enhancement (defined as patients with tumor volume enhancing >0 mL) was recorded as a binary measure and used for subanalyses. Due to the low number of H3.1 mutations (HIST1H3B = 8, HIST1H3C = 3), we grouped them into a single category and compared them with H3F3A mutations. For each outcome tested, we calculated the false discovery rate (FDR) using the Benjamani-Hochberg method to account for multiple hypothesis tests. Tests with a calculated FDR of ≤ 10% were considered statistically significant after FDR adjustment.17 R version 3.5.0 (http://www.r-project.org) and SAS version 9.4 (SAS Institute) were used for statistical analyses.

RESULTS

Patients

Of the 50 patients who underwent biopsy, all had baseline MR imaging studies. The median age at time of enrollment was 6 years (range, 3.3–17.5 years). Twenty-six (52%) subjects were female, 48 (96%) were successfully assigned to a molecular subgroup (2 unassigned), and 34 (68%) patients had (single) identifiable histone mutations (Table 1). One subject was excluded from subsequent analysis due to the concurrent presence of 2 histone mutations.

Tumor Enhancement and Survival

The presence of enhancement (hazard ratio = 3.2; 95% CI, 1.5–7.0; \( P = .003 \)) was associated with decreased OS after FDR adjustment (On-line Fig A and On-line Table 2). OS was not associated (\( P \geq .07 \)) with T2-FLAIR/T2 volume, tumor volume cyst/necrosis, the presence of necrosis, tumor volume enhancing (subset of patients with enhancing tumor of >0 mL), ADC_FLAIR histogram metrics, or PG ADC enhancement histogram metrics after FDR adjustment.

Higher tumor volume enhancing (\( P = .02 \)) and the presence of enhancement (\( P = .06 \)) showed a strong trend with decreased PFS but were not significant after FDR adjustment (On-line Fig B and On-line Table 2). T2-FLAIR/T2 tumor volume (\( P = .07 \)) also showed a trend toward a significant association with decreased PFS.

Molecular Subgroup Analysis

There were differences in tumor volume enhancing across the 4 molecular subgroups within the subset of patients with tumor volume enhancing of >0 mL, but the test was not significant.

**FIG 2.** Differences in tumor volume enhancing and ADC histogram parameters between histone mutations. A, H3F3A-mutated tumor shows lower tumor volume enhancing (0.08 mL), lower ADC_FLAIR mode (1099 × 10^-6 mm^2/s), higher ADC_FLAIR skewness (1.41), and higher ADC_FLAIR kurtosis (4.39), relative to (B) a HIST1H3B-mutated tumor (tumor volume enhancing: 3.61 mL; FLAIR_ADC mode: 1558 × 10^-6 mm^2/s; FLAIR_ADC skewness: 0.039; and FLAIR_ADC kurtosis: 0.33).
after FDR adjustment \((P = .04; n = 30)\) (On-line Table 3). All MGMT+ tumors showed enhancement, most MGMT−/EGFR+ tumors \((77\%)\) showed enhancement, and 54% of MGMT−/EGFR− tumors showed enhancement (On-line Table 3). The 3 MGMT+/EGFR− tumors had the highest tumor volume enhancing values, occasionally an order of magnitude higher than those of other molecular groups, including the MGMT+/EGFR+ tumors. No other factor analyzed showed a significant association with molecular groups, including ADC histogram metrics.

**Histone Mutation Analysis**

Subjects with HIST1H3B/C tumors showed a trend toward improved OS that bordered on significance \((P = .0501)\) compared with subjects with H3FA3 tumors (Fig 1). There was no difference in PFS \((P = .22)\) between subjects with H3FA3 and HIST1H3B/C tumors.

Tumor volume enhancing (in a subset of patients with enhancing tumor volumes of \(> 0 \text{ mL}, \ P = .048\)), median ADC_{FLAIR} \((P = .03)\), and mode ADC_{FLAIR} \((P = .02)\) were significantly higher in tumors with HIST1H3B/C mutations. Skewness ADC_{FLAIR} \((P = .009)\) and kurtosis ADC_{FLAIR} \((P = .03)\) were significantly higher in tumors with H3FA3 mutations. H3FA3 mutations had lower ADC values compared with HIST1H3B tumors (Fig 2 and Table 2).

A comparison of ADC T2-FLAIR/T2 mean, median, and mode; skewness; and kurtosis did not show differences between wild-type tumors and grouped histone mutation (H3FA3 and HIST1H3B/C) tumors (Table 3). Kaplan-Meier analysis also failed to demonstrate significant differences in OS or PFS between wild-type and grouped histone-mutated tumors (Fig 3).

**DISCUSSION**

Even though MR imaging has traditionally played a central role in the evaluation and monitoring of patients with DIPG, little is known regarding the biologic and molecular correlates of imaging features such as the presence of enhancement, low diffusivity, and findings on ADC histogram analysis.12,18 Recently, more centers are performing biopsies for DIPG, which has permitted identification of molecular subgroups with distinct clinical characteristics.3,19-21 This work demonstrates the feasibility of a radiogenomic approach to investigate differences among various tumor groups. Specifically, we identified significant differences in volume of enhancing tumor between molecular subgroups and differences in volume of enhancing tumor and ADC histogram parameters across tumors with different histone mutations. We followed a previously validated approach for patients with DIPG, thus, minimizing the risk of introducing methodologic bias in our analysis.1,2

Molecular characteristics of brain tumors are now widely accepted as important diagnostic criteria for many entities as well as independent prognostic markers.22,23 Castel et al3 investigated the molecular profile of DIPG in 91 children and identified 2 distinct subgroups with mutations in H3.3 (H3FA3A) and H3.1 (HIST1H3B/C) histone variants. Tumors with H3FA3A mutations had a worse prognosis, with a poor response to radiation therapy and earlier relapse compared with HIST1H3B/C mutated tumors.20 Similar to the observations by Castel et al, we found a strong trend toward shorter OS in patients with H3FA3 mutations relative to HIST1H3B/C, though this did not reach statistical significance, likely due to the smaller sample size in our study.

ADC histogram analysis revealed differences across histone mutations, with H3FA3A-mutated tumors having lower ADC T2-FLAIR/T2, higher skewness, and higher kurtosis than HIST1H3B/C-mutated tumors. This finding suggests that histogram metrics
may help differentiate tumors that have the H3F3A mutation, which carries a worse prognosis, from those with HIST1H3B/C mutations. Poussaint et al12 found that tumors with lower preradiotherapy mean, median, and mode ADC FLAIR (analogous to T2-FLAIR/T2 in our study) had shorter PFS and that tumors with higher skewness and kurtosis had shorter OS and PFS. Although that study did not include pathologic or molecular analysis of the tumors, the findings consistently indicated that ADC histogram metrics can identify tumors with more aggressive biology; it is therefore possible that some of the tumors with these ADC histogram features in the study by Poussaint et al may have had H3F3A mutations.

The differences in ADC and ADC histogram parameters among histone-mutated tumors is likely multifactorial. As postulated by Poussaint et al12, Lober et al26 Zukotynski et al27 Chen et al28 and Lober et al28 on prior studies of DIPG, and by Nowosielski et al29 on studies of adult high-grade gliomas, the lower diffusivity and increased kurtosis possibly reflect higher cellularity, which could be related to the more aggressive biology of H3F3A-mutated tumors. Additionally, Castel et al3 reported that the H3.1 tumors had an enrichment in genes associated with glioblastoma with edema, which correlated with exuberant extracellular edema in their histologic analysis; these factors may also contribute to overall higher ADC and lower skewness/kurtosis in H3.1-mutated tumors. A difference between the study by Poussaint et al12 and the current study is that we did not find an association of ADC histogram metrics with PFS or OS, which could be related to a larger sample size in the former study, which analyzed approximately 3 times as many patients. However, the significant trends we observed showed an association of other adverse prognostic factors with the more aggressive (H3F3A) tumors.

We did not identify differences in mean, median, mode ADC histogram parameters, or survival between wild-type histone tumors and H3K27M-mutated tumors (collectively). Similarly, the study by Castel et al3 did not identify differences in survival between wild-type tumors and histone-mutated tumors (collectively). Chen et al31 found lower ADC in the tumor and peritumoral regions of patients with H3K27M mutations compared with wild-type tumors; however, the patient population analyzed in that study is vastly different, possibly accounting for the discrepancy. For instance, the study by Chen et al included H3K27M tumors outside of the brain stem (spinal cord) and included substantially older subjects (most of subjects were older than 18 years of age).31

In our cohort, higher tumor volume enhancing had a significant association with shorter OS and PFS, as has been shown in

FIG 3. Prognostic differences in subjects with any histone mutation (n = 34) versus wild-type (n = 15) tumors. A, Kaplan-Meier curve shows no significant difference in OS between mutant versus wild-type tumors. B, Kaplan-Meier curve shows no significant difference in PFS between mutant versus wild-type tumors.
Differences in tumor volume enhancing were also observed between tumors with H3F3A and HIST1H3B/C mutations, a finding that has not been previously reported in the literature.

Most interesting, higher tumor volume enhancing, which is thought to have a negative prognostic association, was higher in HIST1H3B/C tumors, which have a better prognosis compared with H3F3A tumors.7 We believe that this observation is likely related to the higher expression of genes responsible for angiogenesis in H3.1-mutated tumors, and these findings correlate with preliminary observations described by Castel et al3 in a subanalysis of their radiologic data. Our sample size was too small to perform a subgroup analysis between enhancing and nonenhancing tumors for each histone mutation; additional work is needed to further elucidate the implications of differences in enhancement in tumors with different histone mutations.24 This complex landscape suggests that a multiparametric approach may be useful when evaluating DIPG, because multiple genetic pathways (including MGMT, EGFR, histone) probably account for the pleomorphic MR imaging appearance and clinical course of these tumors, and these factors likely interact with each other to determine the clinical course.33

There are several limitations to this study. First, we conducted our analysis on a small sample size from a single prospective clinical trial. Consequently, our results are preliminary observations that need to be corroborated in a larger separate cohort. Due to the small sample size, we were unable to explore the prognostic implications of differences in enhancement and ADC histogram metrics within individual molecular subgroups and histone variants. Although we identified group differences in diffusion metrics, ADC histogram analysis, and enhancement, the exact biologic substrates for these were not explored. More detailed genetic analyses being performed on this cohort of patients may uncover biologic explanations for these differences. Finally, we analyzed only the baseline MR imaging and did not analyze the changes in any of the baseline parameters with treatment on subsequent MRIs. ADC histogram metrics and tumor enhancing volume could potentially serve as valuable markers of treatment response.

**CONCLUSIONS**

MR imaging features including enhancement and ADC histogram metrics are correlated with molecular subgroups and histone mutations in children with DIPG. Noninvasive markers that allow identification of molecular subgroups of DIPG could help prognosticate survival and guide therapy. Future studies in a larger cohort are required to verify these findings.

**REFERENCES**


Syntnete MRI of Preterm Infants at Term-Equivalent Age: Evaluation of Diagnostic Image Quality and Automated Brain Volume Segmentation

T. Vanderhasselt, M. Naeyaert, N. Watté, G.-J. Allemeersch, S. Raeymaeckers, J. Dudink, J. de Mey, and H. Raeymaekers

ABSTRACT

BACKGROUND AND PURPOSE: Neonatal MR imaging brain volume measurements can be used as biomarkers for long-term neurodevelopmental outcome, but quantitative volumetric MR imaging data are not usually available during routine radiologic evaluation. In the current study, the feasibility of automated quantitative brain volumetry and image reconstruction via synthetic MR imaging in very preterm infants was investigated.

MATERIALS AND METHODS: Conventional and synthetic T1WIs and T2WIs from 111 very preterm infants were acquired at term-equivalent age. Overall image quality and artifacts of the conventional and synthetic images were rated on a 4-point scale. Legibility of anatomic structures and lesion conspicuity were assessed on a binary scale. Synthetic MR volumetry was compared with that generated via MANTIS, which is a neonatal tissue segmentation toolbox based on T2WI.

RESULTS: Image quality was good or excellent for most conventional and synthetic images. The 2 methods did not differ significantly regarding image quality or diagnostic performance for focal and cystic WM lesions. Dice similarity coefficients had excellent overlap for intracranial volume (97.3%) and brain parenchymal volume (94.3%), and moderate overlap for CSF (75.6%). Bland-Altman plots demonstrated a small systematic bias in all cases (1.7%–5.9%).

CONCLUSIONS: Synthetic T1WI and T2WI sequences may complement or replace conventional images in neonatal imaging, and robust synthetic volumetric results are accessible from a clinical workstation in less than 1 minute. Via the above-described methods, volume assessments could be routinely used in daily clinical practice.

ABBREVIATIONS: BPV = brain parenchymal volume; SyMRI = synthetic MR imaging

Despite improvements in neonatal care, a high number of infants born very preterm develop sensorimotor, cognitive, and behavioral disabilities later in life. Identifying children at risk is essential for adequate parental counseling and may also facilitate early intervention strategies. Numerous studies have shown that structural MR imaging results at term-equivalent age are predictive of future motor and cognitive outcomes, and multiple volumetric MR imaging studies have identified associations between reduced brain volume and impaired neurodevelopment. Notably however, quantifying brain volumes in neonates is challenging.

A neonate’s head is only a quarter of the size of an adult’s brain, and the relative signal intensity of the GM and WM is reversed due to the high amount of unmyelinated WM. Neonatal scans often demonstrate changes in the brain, such as moderate-to-severe dilation or distortion of the ventricles, hyperintensity or hypointensity zones, and cystic or hemorrhagic lesions; therefore, they require substantial adjustments to the segmentation approaches used in adults. Several studies have demonstrated the feasibility of neonatal MR volumetry with custom-made software. These methods are often time-consuming and require specialized infrastructure, technical expertise, and powerful computing resources however and are thus not available in most centers. These difficulties hinder the use of this important biomarker in clinical practice.

Synthetic MR imaging (SyMRI; SyntheticMR, Linköping, Sweden) is an emerging imaging technique that simultaneously quantifies R1 and R2 relaxation rates, proton density, and B1 field based on a fast, multidelay, multiecho acquisition. Dedicated synthetic MR imaging (SyMRI) software (Synthetic MR, Linköping, Sweden) then facilitates the reconstruction of synthetic images with
a combination of virtually any TE, TR, and TI. This may save time because multiple synthetic sequences can be reconstructed from the same multidelay, multiecho sequence even after the patient has left the scanner. SyMRI also provides fully automated volumetric analysis based on the expected quantitative values for different brain tissue types. The software is integrated into the radiology PACS system, and brain volume analysis is available in <1 minute.

In 1 large, prospective, multicenter, multireader trial, the overall diagnostic quality of synthetic images in adults was reportedly similar to that of conventional imaging. Furthermore, synthetic MR imaging segmentation is reportedly rapid and robust and exhibits excellent repeatability. To date, only a few studies have examined the feasibility of synthetic MR imaging in children however, and there are very few data available on neonates.

Therefore, in the present study, total brain volume and intracranial values, the erroneous CSF is converted nearly entirely to GM. In the new Version 11.1, the CSF contribution in the immature brain is far more hydrated than an adult brain, which caused earlier versions of SyMRI to detect a CSF haze throughout the entire brain. In the new Version 11.1, the CSF contribution in the immature brain is suppressed. Because the WM definition remains at adult values, the erroneous CSF is converted nearly entirely to GM. Therefore, in the present study, total brain volume and intracranial volume were evaluated, but further distinction between GM and WM was not attempted.

The primary objective of the current study was to evaluate the image quality and diagnostic performance of synthetic MR imaging in a neonatal population compared with conventionally acquired images. A secondary objective was to assess the new and improved SyMRI algorithm for neonatal brain segmentation and its applicability in clinical practice. To validate these segmentations, the synthetic tissue volumes were compared with data derived from MANTiS (Morphologically Adaptive Neonatal Tissue Segmentation), a state-of-the-art research method for neonatal brain segmentation.

**MATERIALS AND METHODS**

**Study Population**

A total of 111 very preterm infants (born at <32 weeks’ gestation) were scanned at an average postmenstrual age of 40.0 ± 1.0 weeks at our institution (Universitair Ziekenhuis Brussels) as part of their routine clinical follow-up between February 2017 and June 2019. All patients who had both synthetic multidelay, multiecho, and conventional 2D or 3D MR imaging sequences were included in the study. Both 2D and 3D T2WI datasets were available for 59 neonates. Seventeen neonates had only the 2D T2WI dataset, while 32 neonates had only the 3D T2WI dataset.

Image quality and diagnostic performance were assessed on 76 subjects (mean gestational age at birth, 29.2 ± 2.6 weeks; 42 boys and 34 girls) who had both synthetic multidelay, multiecho, and conventional 3D T1WI and 2D T2WI TSE data available for comparison. Because evaluation of image artifacts and image quality was part of the experimental design, no subjects were excluded on this basis.

A conventional 3D T2WI dataset was required for segmentation using MANTiS. Ninety-one neonates (mean gestational age at birth, 29.4 ± 2.0 weeks; 49 boys and 42 girls) underwent both 3D T2WI TSE and multidelay, multiecho scanning and were included in the quantitative volumetric validation.

After visual quality control, 16 subjects were rejected due to the following: excessive movement during the multidelay, multiecho scan (n = 5), failed watershed skull-stripping by MANTiS (n = 4), segmentation error in MANTiS caused by a previously unreported bug (n = 6), or failed synthetic segmentation due to corrupted DICOM data (n = 1). The 75 remaining subjects included 40 boys and 35 girls with a mean gestational age at birth of 29.0 ± 2 weeks and a mean postmenstrual age at scanning of 40 ± 1 week.

**MR Imaging Acquisition**

All examinations were performed on a 3T Ingenia MR imaging scanner (Philips Healthcare, Best, the Netherlands) with a 16-channel head coil. The children were sedated using oral chloral hydrate (25–50 mg/kg) and placed in a MedVac vacuum immobilization device (CFI Medical Solutions, Fenton, Michigan) before scanning. Neonatal earmuffs (Minimuffs; Natus Medical, San Carlos, California) covered by headphones were used for ear protection. Informed parental consent was obtained before MR imaging and sedation.

First, conventional images were obtained in accordance with the standard dedicated neonatal MR imaging protocol used at our institution. This included, among other things, a 3D MPRAGE and a T2WI sequence (transversal T2WI TSE, 3D T2WI, or both when time allowed). A multidelay, multiecho acquisition covering the entire brain was then acquired to perform quantitative mapping of T1, T2, and proton density using synthetic MRI. Detailed imaging parameters are shown in On-line Table 1.

Synthetic T1WIs and T2WIs and their inversion recovery variants (T1 phase-sensitive inversion recovery and T2 STIR) were generated on the basis of R1, R2, and proton density maps by fitting the data to the analytic curves describing the signal intensity, using a least-squares approach. This method has been described in detail elsewhere. TE, TR, and TI were chosen to maximize tissue contrast (On-line Table 2).

**Image Quality and Diagnostic Performance**

The image quality and diagnostic performance of conventional and synthetic images were assessed by a blinded senior pediatric neuroradiologist (T.V.), a neuroradiologist (S.R.), and a fellow in neuroradiology (G.-J.A.) with 12, 7, and 3 years of experience in neuroradiology, respectively. For each rater, the order of the datasets was randomized, and datasets were divided into 2 sessions with a mix of conventional and synthetic datasets in each session. Only the synthetic or the conventional images of a subject were presented in the same session. Both sessions were conducted at least 2 weeks apart to prevent recall bias.

The overall image quality for each series was assessed on a 4-point Likert scale, in which poor indicates severe image-quality issues precluding diagnosis; sufficient, moderate image-quality issues but acceptable for diagnosis; good, only minor image-quality issues; and excellent, no noticeable image-quality issues. Artifacts were scored on a similar 4-point scale (severe, moderate, minor, and
Lesion conspicuity and the legibility of different anatomic structures were evaluated for each conventional and synthetic data-set on a binary scale, as indicated in Table 1. Assessment of diagnostic performance included the detection of focal WM lesions and cystic degeneration. Conventional and synthetic datasets were compared with the original radiology report on the basis of all available MR images and knowledge of the clinical findings.

**MR Volumetry Calculations**

Synthetic tissue volumes were calculated from R1, R2, and proton density maps using the automated segmentation tool in SyMRI, Version 11.1, optimized for segmenting neonatal brains. Segmentation was performed by comparing R1, R2, and proton density values of individual voxels with a lookup table. This process enables rapid segmentation, capable of coping with partial volume effects and anatomic distortions such as hydrocephalus or cysts. Synthetic MR imaging brain segmentation was validated by comparing the results with those of MANTiS, a state-of-the-art toolbox based on SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12) and ITK-SNAP 1.6 (www.itksnap.org) dedicated to segmentation of the neonatal brain. Brain extraction was performed before tissue segmentation using the watershed-based method provided in MANTiS.

**Statistics**

Only good or excellent images are desirable in daily clinical practice; therefore, we chose to dichotomize image quality and artifact scores into poor-sufficient and good-excellent groups. Group differences for conventional and synthetic imaging sets were assessed with the McNemar test for all binary results, using SPSS Statistics 23 (IBM, Armonk, New York). Descriptive statistics were used to assess the legibility of anatomic structures.

Bland-Altman plots for brain parenchymal volume (BPV), brain parenchymal fraction, intracranial volume, and CSF were generated using Matlab 2016b (MathWorks, Natick, Massachusetts) to compare the measured volumes and identify systematic differences between the proposed synthetic method and MANTiS. The synthetic T2WIs were coregistered to

---

**Table 1: Image quality**

<table>
<thead>
<tr>
<th>Category</th>
<th>Assessment Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall image quality</td>
<td>Poor, sufficient, good,</td>
</tr>
<tr>
<td>Artifacts</td>
<td>excellent</td>
</tr>
<tr>
<td>Visualization of anatomic</td>
<td>Severe, moderate, minor,</td>
</tr>
<tr>
<td>structures</td>
<td>none</td>
</tr>
<tr>
<td>Posterior cross-roads</td>
<td>Legible, illegible</td>
</tr>
<tr>
<td>(T2/T2 STIR only)</td>
<td></td>
</tr>
<tr>
<td>Central sulcus</td>
<td></td>
</tr>
<tr>
<td>Lentiform nucleus</td>
<td></td>
</tr>
<tr>
<td>Ventrolateral thalamus</td>
<td></td>
</tr>
<tr>
<td>Dorsal pons</td>
<td></td>
</tr>
<tr>
<td>Diagnostic performance</td>
<td>Present, absent</td>
</tr>
<tr>
<td>Focal WM lesions</td>
<td></td>
</tr>
<tr>
<td>Cystic degeneration</td>
<td></td>
</tr>
</tbody>
</table>

---

**FIG 1.** Qualitative comparison of T1WI and T2WI quality (A) and image artifacts (B). Horizontal lines within colored portions indicate scoring per rater. PSIR indicates phase-sensitive inversion recovery.
the conventional dataset using SPM12 running on Matlab 2016b. The coregistration parameters were applied to the synthetic segmentation maps. A Dice similarity coefficient (DSC in the formula below) was then calculated in Matlab 2016b for CSF, BPV, and intracranial volume to investigate differences between the segmentation methods.23 This score represents the overlap between 2 segments. For the component $x$, $DSC_x$ is defined as

$$DSC_x = 2 \frac{\text{Synt}_x \cap 3D_x}{|\text{Synt}_x| + |3D_x|}. $$

**RESULTS**

**Image Quality and Artifacts**

The results of the image-quality evaluation are presented in Fig 1. The number of scans rated as poor was negligible for both the conventional images and the synthetic images (0.0%–2.2%). The synthetic T1 phase-sensitive inversion recovery resulted in significantly more images with minimal or minor artifacts ($P = .017$, $P = .008$, and $P = .016$, respectively) than the conventional images. The most marked difference, however, was the higher percentage of synthetic images with no or minimal artifacts ($P < .001$ for all comparisons).

**Diagnostic Performance**

A representative sample of conventional and synthetic images is shown in Fig 2. The anatomic structures outlined in Table 1 were identified in all the conventional and synthetic sequences. The accuracy of the detection of WM lesions was slightly better in the conventional image sets, whereas the detection of cysts was better in the synthetic images; but neither of these differences was statistically significant (Table 2).

**Volumetric Validation**

The results of the intracranial volume, BPV, parenchymal fraction, and CSF volume measurements are shown in the Bland-Altman plots in Fig 3. There was no correlation in the Bland-Altman plots between the mean metric under consideration and the relative difference. There was a small systematic bias in all cases. SyMRI estimated the CSF volume to be 4.50% lower than in MANTiS, while intracranial volume, BPV, and parenchymal fraction were slightly larger (1.68% for parenchymal fraction, 4.19% for intracranial volume, and 5.87% for the BPV). Therefore, it is essential to consider which method is used for volumetry when interpreting the results. Except for the CSF measurement, the spread between both measures was small (<3.5%), indicating good agreement between the methods. For CSF, the spread was more substantial at 18%. This difference may be attributable to the smaller volume of CSF, for which minor differences in the segmentation threshold may have a greater impact. However, the spread of the relative CSF fraction (2.8%) was much lower.

Regarding overlap of segmentation maps, the average Dice similarity coefficients were $0.9731 \pm 0.0004$ for intracranial volume, $0.9428 \pm 0.0007$ for BPV, and $0.7566 \pm 0.0055$ for CSF. The very high scores for intracranial volume and BPV indicate substantial overlap. For the CSF maps, the overlap was moderate. Figure 4 shows a representative example of the overlap of 2 CSF segments.

**DISCUSSION**

A meticulous literature search suggests that this is the first dedicated study investigating the feasibility of synthetic MR imaging and brain volumetry in infants born very preterm, scanned at term-

---

**Table 2: Lesion conspicuity**

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WM lesions ($n = 23/76$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional</td>
<td>91%</td>
<td>86%</td>
<td>89%</td>
<td>1.00</td>
</tr>
<tr>
<td>Synthetic</td>
<td>89%</td>
<td>77%</td>
<td>86%</td>
<td></td>
</tr>
<tr>
<td>Cysts ($n = 7/76$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional</td>
<td>96%</td>
<td>57%</td>
<td>92%</td>
<td>.29</td>
</tr>
<tr>
<td>Synthetic</td>
<td>100%</td>
<td>64%</td>
<td>95%</td>
<td></td>
</tr>
</tbody>
</table>

FIG 2. Upper row, Conventional T1WI (A) and T2WI (B and C). Lower row, Synthetic T1WI (D) and T2WIs (E and F). Focal WM lesions (small arrows) are clearly visible on both T1WIs. Cystic periventricular leukomalacia (arrows) is seen around the ventricles in both T2WIs.
equivalent age. The quality of the synthetic images did not differ significantly from that of conventional images when the images were dichotomized into a good or excellent group and a poor or sufficient group. Both synthetic and conventional methods facilitated excellent visualization of normal brain structures. The accuracy of synthetic MR imaging for the detection of WM lesions was lower, but the difference was not significant (P = 1.00). Synthetic MR imaging sequences could therefore complement or possibly replace conventional T1WI and T2WI in neonatal imaging. Additionally, because the acquisition time remains approximately the same, contrasts that are currently omitted due to lack of time, such as single or double inverse recovery or proton density sequences, can be reconstructed without a time burden.

The inverted GM and WM contrast and the high amount of water in the neonatal brain render MR volumetrics in neonates challenging. Previous versions of SyMRI often mislabeled WM as CSF in neonates, resulting in

FIG 3. Bland-Altman plots comparing different volumetric metrics calculated with Synthetic MR imaging (labeled SyMRI) and MANTiS. The red line indicates the average relative difference, and the blue dashed lines indicate the average ±1.96.

FIG 4. Synthetic CSF segmentation maps: SyMRI algorithm without optimization for the immature neonatal brain and CSF haze in the WM (blue) (A); improved SyMRI (Version 11.1) at the same level (B). Example of CSF overlap to calculate the Dice similarity coefficients (C). The synthetic CSF segment is red, and the MANTiS segment is blue. Regions where the segments overlap are purple. Voxels of the SyMRI CSF segment containing <20% CSF are not shown.
inaccurate volumetric results.\textsuperscript{17,18} In the present study, a new version of SyMRI (Version 11.1) that features an optimized algorithm for neonatal brain segmentation was used.\textsuperscript{24} The volumetric measurements of synthetic MR imaging with this optimized algorithm were generally corroborated with those of MANTiS software analysis. The Dice similarity coefficients exhibited excellent overlap for intracranial volume and BPV and moderate overlap for CSF. The Bland-Altman tables also indicated minimal spread between the 2 methods for intracranial volume, BPV, and brain parenchymal fraction. The segmentation of SyMRI was more robust than that of MANTiS, with respective failure rates of 6.6\% and 11.0\%.

The added prognostic value of neonatal MR imaging volumetry has been demonstrated in numerous studies.\textsuperscript{6–8} To the best of our knowledge, SyMRI is the first commercially available product capable of accurately measuring brain volume in neonates. The tool is integrated into the radiology PACS system, and volume measurements are available in <1 minute.

The current study had several strengths, such as the evaluation of both the image quality and the volumetric segmentation results of synthetic MR imaging. The multidecal, multiecho sequence was part of the standard MR imaging protocol for all neonates born very preterm, which makes it an excellent representation for this population. All MR images were obtained very close to the term-equivalent age (40 ± 1 week), which is considered optimal for the assessment of structural abnormalities that are related to long-term outcome.\textsuperscript{25}

The study also had some limitations. Because the children were scanned only once, it was not possible to evaluate the intramethod repeatability of volume segmentations. Second, although all the scans were carefully blinded for imaging method, the trained neuroradiologists may have been able to discern the difference between synthetic and conventional images with the naked eye, and this may have introduced some bias. Third, even though MANTiS is state-of-the-art research software for neonatal brain segmentation, it cannot be regarded as a criterion standard like manual segmentation. In the present study, the correspondence between the 2 methods was excellent. The improved SyMRI segmentation algorithm performed well when segmenting the brain parenchyma and delineating the brain mask, but in contrast to MANTiS, it is still unable to differentiate between GM and WM in neonates.

**CONCLUSIONS**

The current study suggests that in neonates, synthetic MR imaging can complement or be an alternative to conventional T1WI and T2WI sequences because it had excellent image quality and diagnostic performance. Volumetric measurements can be obtained in <1 minute after processing and are comparable with measurements derived via current state-of-the-art research methods. Synthetic MR imaging may therefore facilitate the routine use of these biomarkers in daily clinical practice, bridging the existing gap between the past decade of research on volumetric measurements and long-term outcomes and clinical practice.

Disclosures: Tim Vanderhasselt—\textsuperscript{UNRELATED} Employment: Radiologist. Comments: Staff Member Radiology Department Universitair Ziekenhuis Brussels, Brussels, Belgium. Maarten Naeyaert—\textsuperscript{UNRELATED} Employment: paid PhD student at the Bio-Imaging lab, Universiteit Antwerpen, through the Interdisciplinary PhD grant BOF UA 202. Nina Watté—\textsuperscript{UNRELATED} Employment: radiology resident. Jeroen

**REFERENCES**


19. Andica C, Hagiwara A, Nakazawa M, et al. The advantage of synthetic MRI for the visualization of early white matter change in
Disrupted Functional and Structural Connectivity in Angelman Syndrome


ABSTRACT

BACKGROUND AND PURPOSE: This work investigated alterations in functional connectivity (FC) and associated structures in patients with Angelman syndrome (AS) by using integrated quantitative imaging analysis and connectivity measures.

MATERIALS AND METHODS: We obtained 3T brain MR imaging, including resting-state functional MR imaging, diffusion tensor imaging, and 3D T1-weighted imaging from children with AS (n = 14) and age- and sex-matched controls (n = 28). The brains of patients with AS were analyzed by measuring FC, white matter microstructural analysis, cortical thickness, and brain volumes; these were compared with brains of controls.

RESULTS: Interregional FC analysis revealed significantly reduced intra- and interhemispheric FC, especially in the basal ganglia and thalamus, in patients with AS. Significant reductions in fractional anisotropy were found in the corpus callosum, cingulum, posterior limb of the internal capsules, and arcuate fasciculus in patients with AS. Quantitative structural analysis also showed gray matter volume loss of the basal ganglia and diffuse WM volume reduction in AS compared with the control group.

CONCLUSIONS: This integrated quantitative MR imaging analysis demonstrated poor functional and structural connectivity, as well as brain volume reduction, in children with AS, which may explain the motor and language dysfunction observed in this well-characterized neurobehavioral phenotype.

ABBREVIATIONS: AS = Angelman syndrome; BG = basal ganglia; FA = fractional anisotropy; FC = functional connectivity; FDR = false discovery rate; MI = primary motor area; MD = mean diffusivity; RD = radial diffusivity; rs-fMRI = resting-state functional MRI; SMA = supplementary motor area; TBSS = Tract-Based Spatial Statistics; UBE3A = ubiquitin protein ligase E3A; VBM = voxel-based morphometry; MNI = Montreal Neurological Institute; TCE = threshold-free cluster enhancement

Angelman syndrome (AS) is a rare genetic disorder caused by the functional loss of ubiquitin protein ligase E3A (UBE3A), which arises from mutation of UBE3A or silencing of the maternal UBE3A allele. The distinctive clinical phenotypes of AS (ie, lack of expressive language development, frequent seizures, and gait abnormalities with ataxia) cause serious medical and social problems; moreover, approximately 10% of patients with AS never achieve self-ambulation. Despite their severe clinical features and well-known genetic causes, visual analysis of conventional brain MR imaging has revealed only nonspecific findings and minor abnormalities, such as mild cortical atrophy, dysmyelination, focal white matter signal abnormalities, cerebral atrophy of frontotemporal areas, and callosal thinning.

Recently, many functional MR imaging studies have investigated the abnormal neural connectivity in patients with speech disorders or ASD to identify the neural correlates of these functional deficits. However, there is no study to investigate an abnormal functional network in children with AS, which is a relatively homogeneous disease with a clear clinical and genetic delineation. Although WM pathway alteration and gray matter volume loss in patients with AS were previously reported by using diffusion tensor imaging and voxel-based morphometry...
VBM, it remains unclear whether alterations in functional connectivity (FC) are related to structural deformation in AS.

Based on previous studies, we hypothesized that the structural and FC of the fronto-striatal circuits, including language and sensorimotor domains, would be seriously compromised in patients with AS. Thus, we aimed to investigate the abnormal functional network in children with AS, which is associated with severe speech and motor deficits and to validate the key anatomic structures associated with functional loss of UBE3A by using multimodal quantitative MR imaging analysis.

METHODS
Participants
We enrolled patients who were clinically and genetically diagnosed with AS between July and November 2016. Eligibility criteria included (1) patients who were genetically confirmed with AS, (2) patients whose ages were less than 18 years old, and (3) patients who had no contraindications to brain MR imaging. Because most patients lack expressive language and have severe intellectual disability lower than 2 years of mental age, it was difficult to assess the patients’ developmental status by using a standardized mean. The patients’ motor and language functions were evaluated by an experienced pediatric neurologist (M.S.Y. with 15 years of clinical experience in pediatric neurology) at the time of their brain MR imaging. Patients were classified into 2 groups according to the severity of their verbal and motor functional status: 1) patients with no verbal output versus patients who can speak at least 1 word and 2) patients who are unable to walk unassisted versus patients who can walk unassisted.

Age- and sex-matched control subjects, with a patient to control ratio of 1:2, were retrieved from the retrospective database in our institution (Asan Medical Center) from May 2014 to July 2017. The institutional review board allowed the researchers to conduct the analysis of the MR data from this database without subject consent because the data were subject to anonymization. All patient data were anonymized when received before analysis. Children who were included in the control group were required to 1) have no alleged neurologic or psychiatric deficits, 2) be prescribed no regular medication, 3) have no abnormality on brain MR imaging, and 4) show clinically normal development. In total, 28 children (male:female = 14:14) were included, with a mean age of 7.7 ± 4.2 years (range, 1–16 years). They underwent brain MR imaging due to headache (n = 15), provoked seizures (n = 7), syncope (n = 3), dizziness (n = 1), tic disorder (n = 1), and minor head trauma (n = 1).

Written informed consent was obtained from the legal guardians of all patients. This study was performed in accordance with the ethical standards of the Declaration of Helsinki and was approved by the institutional review board of the Asan Medical Center (2016–0279, Seoul, Korea).

Imaging Protocol
All participants successfully underwent brain MR imaging by using a 3T MR imaging scanner (Achieva; Philips Healthcare, Best, The Netherlands) with an 8-channel, sensitivity-encoding head coil. 3D T1-weighted imaging sequences were acquired with the following image parameters: TR/TE, 9.9/4.6 ms; flip angle = 8°; FOV = 224 mm, matrix = 224 × 224 mm, section thickness = 1 mm with no gap; and scan time, 4 min, 41 seconds. DTI was performed by using single-shot echo-planar imaging with the following parameters: TR/TE, 5615/70 ms; section thickness, 3 mm; diffusion directions, 32; b-value, 1000 seconds/mm²; number of excitations, 1; matrix, 112 × 112; FOV, 224 mm; acquired voxel size, 2 × 2 × 3 mm; and scan time, 3 min, 17 seconds. Blood oxygen level–dependent contrast functional images at rest (resting-state functional MR imaging [rs-fMRI]) were acquired by using single-shot gradient-echo/EPI with the following parameters: TR/TE, 3000/30 ms; flip angle, 90°; FOV, 224 mm; matrix, 112 × 112; section thickness, 6 mm with no gap; number of slices, 17; number of dynamics, 150; and acquisition time, 7 min, 42 seconds. Sedation during MR imaging examination was performed by trained anesthesiologists in accordance with our standard institutional protocol. All participants in the AS group were sedated by using intravenous administration of propofol, and children in the control group were sedated by using intravenous administration of propofol or oral chloral hydrate. Ten children in the control group did not require sedation during MR imaging examination.

Seed-Based Resting-State Functional MR Analysis
Before statistical analysis, rs-fMRI data were preprocessed by using the Analysis of Functional Neuro Images (AFNI; http://afni.nimh.nih.gov/afni) basic ANATICOR program to remove artifacts and normalize the image data. In brief, we removed the first 3 EPIs to control for error, aligned the EPIs, corrected for section-time acquisition, realigned for motion correction, co-registered the EPI to anatomical T1WI, and bandpass filtered from 0.01 to 0.1 Hz.

Anatomic segmentations of 42 subjects were created by using FreeSurfer and were used to mask the following regions for FC analysis: thalami, basal ganglia (BG), primary motor area (M1), supplementary motor area (SMA), sensory area, Broca area (Brodmann area 44 and 45), and Wernicke area (superior temporal gyrus) in the left and right hemispheres. Previous studies with patients with AS indicated severe speech impairment and difficulty in motor controls as well as volume reduction of the BG. The associated anatomic areas were selected as seed regions based on these clinical phenotypes of patients with AS.

Pearson correlation coefficients were calculated by using the Matlab (Matlab 8.2.0.701; MathWorks, Natick, Massachusetts) corr function between the hypothesis-derived, selectively masked regions to indicate the strength of FC between the 2 regions, producing a 14-by-14 matrix for each subject. Individual subjects’ correlation matrices were then grouped into control and patient groups, and the patient group’s FC was compared with that of the control group by using 2-sample t-tests after testing for equal variance (false discovery rate [FDR] corrected; P < .05).

Tract-Based Spatial Statistics
The DTI original time-series data were corrected for head movement and eddy current distortions by using the “eddy_correct” function within the FMRIB Diffusion Toolbox (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT). The FMRIB DTIFit module (http://fsl.fmrib.ox.ac.uk/fsl-fdt.html) was used to independently fit the diffusion tensors to each voxel, creating an output of voxelwise maps of fractional anisotropy (FA) for each
subject. To limit the result to be strictly pertinent to the subject’s brain, we obtained a binary brain mask from each participant’s B0 image with the FSL Brain Extraction Tool’s (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET) bet2 function and applied the brain mask to mask-out structures external to the subject’s FA map.

Diffusion MR imaging data were analyzed by using FMRIB’s Diffusion Toolbox Tract-Based Spatial Statistics (TBSS; v1.2; https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/TBSS) module, which used the FA, mean diffusivity (MD), and radial diffusivity (RD) maps of each participant.15 All FA images were preprocessed using TBSS’s tbs_1_preproc command to further clean the FA images. Next, all FA images were registered to the MNI space by using nonlinear registration (tbs_2_reg), and the group mean FA image and WM skeleton were created by averaging all FA images (tbs_3_postreg -S), which represents the center of all WM tracts common to the group. For each individual, the local FA maxima were found along the perpendicular direction of the mean FA image and then projected onto the mean FA skeleton using a FA threshold of 0.2 (tbs_4_prestats 0.2). With the randomization function in the FMRIB Diffusion Toolbox, the skeletonized FA images were subjected to voxel-wise, cross-subject statistical analysis. MD and RD maps were measured by running the FSL tool tbsds_non_FA script. Here, the mean skeletons of the control and AS groups were compared using voxel-wise 2-sample unpaired t-tests with nonparametric permutation with threshold-free cluster enhancement (TFCE) corrections by using the T2 option and cluster-based thresholding at 0.95 ($P < .05$) in the randomize function (5000 permutations/design, TFCE-corrected $P < .05$ for significance).

**Cortical Thickness and Volume Measurements**

Entire cortex analyses were computed to explore local cortical thicknesses and cortical volumes in the AS and control groups. Statistical maps were generated with the Query, Design, Estimate, Contrast interface within FreeSurfer (version 5.3.1) used to perform group averaging and inference on cortical morphometric data. For each hemisphere, a general linear model was computed vertex by vertex for analysis of cortical thickness and cortical volume, respectively, and for regression to exclude the covariate of age. Cortical maps were smoothed with a 10-mm full width at half maximum Gaussian filter and multiple comparisons were corrected by FDR with a $P$ value set at $.05$. Comparisons of segmented brain volumes between subgroups in the AS group were analyzed using Mann-Whitney $U$ tests in commercial statistical software (SPSS, version 21; SPSS; IBM, Armonk, New York). A $P$ value less than 0.05 was considered statistically significant.

**Voxel-Based Morphometry**

VBM analysis in SPM8 software (http://www.fil.ion.ucl.ac.uk/spm/) was performed with the VBM technique by adhering to the standard VBM processing routine. Images of each subject were normalized and registered to MNI space; the resulting images were modulated (without the affine component) and smoothed with a full width at half maximum of 10 mm. Measurements for the transformed GM and WM images were then clustered with a threshold over 20 voxels. The final processed images were used for statistical inference; age as a covariate measure was excluded by regression with the GLM. Statistical nonparametric mapping with 5000 permutations, without variance smoothing, was used to compare voxel-wise differences in GM volumes between the AS and control groups. The level of significance in SPM was considered a $P$ value $< .001$, and multiple comparisons were corrected with family-wise error rate.

**RESULTS**

**Patient Characteristics**

A total of 14 patients (male:female = 7:7) were enrolled in this study. The mean age of the patient group was 7.6 ± 4.7 years (range, 3–18 years). All patients were genetically confirmed with AS; microdeletion within chromosome 15q11-q13 in 13 patients and paternal uniparental disomy in 1 patient. Most patients initially presented with developmental delay and hypotonia. Clinical phenotypes of the enrolled patients are shown in Table 1. Among 10 patients who had seizures, 2 had histories of myoclonic status in nonprogressive encephalopathies or intractable epilepsy.

**Interregional Functional Connectivity between Language, Motor, and Subcortical Areas**

Fourteen seeds were placed in bilateral thalami, BG, M1, SMA, sensory, Broca, and Wernicke areas. General reduction of interregional FC was found in the AS group compared with the control group (Fig 1). In terms of interhemispheric FC, interregional FC between the left thalamus and all seeds of the right hemisphere, as well as FC between the left BG and all seeds of right hemisphere except for the right sensory area, was significantly reduced in patients with AS compared with the control group. In addition, significant reduction of interhemispheric FC between the right thalamus and left M1, sensory, and Broca areas and of FC between the right BG and left Broca area were observed in patients with AS. Regarding intrahemispheric FC of the right hemisphere, interregional FC of thalamocortical networks (thalamus and M1, Broca, Wernicke), BG and cortical networks (M1, SMA, Broca,}

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Angelman Syndrome ($n = 14$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (months)</td>
<td>95.9 ± 57.8 (range, 3–18 years)</td>
</tr>
<tr>
<td>Sex (male:female)</td>
<td>7:7</td>
</tr>
<tr>
<td>Genetic mutation</td>
<td>UBE3A (5q11-q13 deletions)</td>
</tr>
<tr>
<td></td>
<td>Paternal UPD</td>
</tr>
<tr>
<td>Seizure</td>
<td>10 (71.4%)</td>
</tr>
<tr>
<td>Use of at least 1 meaningful word</td>
<td>5 (35.7%)</td>
</tr>
<tr>
<td>Walking alone</td>
<td>11 (78.6%)</td>
</tr>
<tr>
<td>Ataxia or tremor</td>
<td>10 (71.4%)</td>
</tr>
<tr>
<td>Strabismus</td>
<td>5 (35.7%)</td>
</tr>
<tr>
<td>Ataxia and unable to walk alone</td>
<td>3 (21.4%)</td>
</tr>
<tr>
<td>Ataxia but walking alone</td>
<td>7 (50.0%)</td>
</tr>
<tr>
<td>No ataxia (walking alone)</td>
<td>4 (28.6%)</td>
</tr>
</tbody>
</table>

Note: UPD indicates uniparental disomy.

**Table 1: Clinical characteristics of children with Angelman syndrome in this study**

Wernicke), and the language network (Broca and Wernicke areas) was significantly decreased in patients with AS. Among left intra-hemispheric interregional FC, there were significant FC reductions between thalamus and BG, between thalamus and M1, and between Broca and Wernicke areas in patients with AS (all $P < .05$, FDR-corrected).

To test the effect of sedation during rs-fMRI, the Pearson correlation coefficients between patients with AS (all sedated) and sedated controls patients and the Pearson’s correlation coefficients between sedated controls and nonsedated controls were compared (On-line Fig 1). Using unpaired 2-sample t-tests, we found no significant FC differences between the sedated and nonsedated controls, and the FC reduction in patients with AS showed similar results when compared with all controls (age and sex matched; $n = 28$; Fig 1; FDR-corrected, $P < .05$) or sedated controls (younger, age 3–13 years; sex matched; $n = 10$) (Online Fig 1; FDR-corrected, $P < .05$).

**Tract-Based Spatial Statistics**

TBSS showed a significant FA decrease in the medial superior corpus callosum, cingulum, bilateral posterior limb of the internal capsules, corticospinal tract, cerebral peduncle, and arcuate fasciculus in the patient group compared with the control group (threshold $P < .005$ TFCE corrected, obtained using the FMRIB Software Library tool tbss_fill; Fig 2). The anatomic locales with the reduced FA were identified by using AFNI’s whereami function with matching MNI coordinates, of which the MNI atlas was used for image registration. MD and RD were also analyzed in the bilateral anterior corpus callosum, posterior limb of the internal capsules, and arcuate fasciculus and posterior corpus callosum (On-line Table). All but the mean MD in bilateral posterior limb of internal capsule and posterior corpus callosum showed significant increases in the AS group.
Altered Cortical Thickness in Patients with Angelman Syndrome. Compared with the control group, the AS group showed significantly smaller cortical thickness in bilateral superior temporal, bilateral postcentral, left transverse temporal, right supramarginal, and right superior frontal gyrus when using Mann-Whitney U tests (Table 2; FDR-corrected, \( P < .05 \)).

### Table 2: Cortical thicknesses of children with Angelman syndrome compared with the control group

<table>
<thead>
<tr>
<th>Area Label</th>
<th>Brodmann Area</th>
<th>Size (mm²)</th>
<th>Stereotactic Coordinates (Talairach Space)</th>
<th>T (-Log(10)P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left hemisphere</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transverse temporal</td>
<td>22</td>
<td>154.89</td>
<td>-51.1 -16.4 4.4</td>
<td>-4.9827</td>
</tr>
<tr>
<td>Postcentral</td>
<td>4</td>
<td>77.17</td>
<td>-60.1 -6.7 10.8</td>
<td>-4.7438</td>
</tr>
<tr>
<td>Superior temporal</td>
<td>21</td>
<td>86.13</td>
<td>-50.7 -14.6 -11.9</td>
<td>-4.2409</td>
</tr>
<tr>
<td>Right hemisphere</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior temporal</td>
<td>10</td>
<td>39.51</td>
<td>55.8 2.8 -12.3</td>
<td>-4.1946</td>
</tr>
<tr>
<td>Postcentral</td>
<td>1</td>
<td>9.20</td>
<td>63.6 -8.9 12.6</td>
<td>-3.8439</td>
</tr>
<tr>
<td>Supramarginal</td>
<td>40</td>
<td>37.79</td>
<td>48.0 -23.9 20.5</td>
<td>-3.8319</td>
</tr>
<tr>
<td>Superior frontal</td>
<td>8</td>
<td>13.04</td>
<td>11.3 21.6 37.7</td>
<td>-3.8079</td>
</tr>
</tbody>
</table>

Altered Brain Volume in Patients with Angelman Syndrome. In patients with AS, the global GM volume (mean ± SD of patients with AS versus control group, 600,625.8 ± 62,103.7 mm³ versus 723,046.5 ± 53,590.5 mm³; \( t = -4.83; df = 40; P < .0001 \), cerebral WM volume (306,694.5 ± 45,254.5 mm³ versus 390,684.4 ± 60,945.5 mm³; \( t = -4.05; df = 40; P < .0001 \), brain segmentation volume (941,639.3 ± 94,951.1 mm³ versus 1,155,856.0 ± 103,920.5 mm³; \( t = -5.08; df = 40; P < .0001 \), total intracranial volume (1,212,398.8 ± 132,858.0 mm³ versus 1,436,258.0 ± 150,240.0 mm³; \( t = -5.25; df = 40; P < .0001 \).
mm³ versus 1,442,983.0 ± 174,949.8 mm³; t = 3.90; df = 40; P < .001) and cerebellar volume (99,621.8 ± 15,085.7 mm³ versus 127,993.0 ± 12806.0 mm³; t = 5.69; df = 40; P < .001) decreased significantly compared with those of the control group. No significant difference was found in the global CSF volume between the 2 groups (14,668.07 ± 5,554.20 mm³ versus 15,084.69 ± 5,661.31 mm³; t = 0.22; df = 40; P = .821) (Fig 3). In subgroup analysis, volumes of cerebral WM and corpus callosum were smaller in patients with no verbal output compared with patients who can speak at least 1 word (Fig 4, upper 5 boxplots). In addition, patients who are unable to walk unassisted showed smaller volumes of total GM, cortex, central portion of corpus callosum, right cerebellar cortex, and left amygdala compared with patients who can walk unassisted (Fig 4, lower five boxplots).

Voxel-Based Morphometry
Gray Matter Volume Loss in Patients with Angelman Syndrome. Regional GM volume loss was found in the AS group compared with the control group (Table 3, On-line Fig 1). Using statistical nonparametric mapping with age regressed out as a covariate measure using the general linear model (5000 permutations; P < .001 corrected for family-wise error rate), we found regional volumes of the bilateral putamen, globus pallidus, caudate nucleus, rectus gyrus, left hippocampus, and left cerebellar cortex were significantly reduced in the AS group compared with the control group.

White Matter Volume Loss in Patients with Angelman Syndrome. In the AS cohort, we found significant WM volume loss in the pons, anterior commissure, splenium of the corpus callosum, bilateral subinsular WM, right frontal subcortical WM, and bilateral parietal periventricular WM (On-line Fig 2) compared with the control group. These results were also found by using statistical nonparametric mapping with 5000 permutations and age excluded as a covariate measure by using the general linear model (P < .001, corrected for family-wise error rate).

DISCUSSION
This study used a multimodal imaging approach to study patients with AS with severe speech impairment and motor deficits. With combined use of rs-fMRI and quantitative structural imaging analysis, we attempted to define functional brain network changes in AS and compare these functional changes with the structural changes observed in these children. Children with AS exhibited significantly decreased FC in multiple intra- and interhemispheric networks; this altered FC substantially correlated with specific volume deficits in patients with AS.

In rs-fMRI analysis, children with AS showed a general decrease in interregional FC when we focused the seeds in areas associated with primary motor and language systems. Especially, intra- and interhemispheric FC associated with the thalamus and BG of patient with AS were significantly decreased compared with controls. Considering the role of the BG and the thalamus network in motor control and balance,16,17 altered FC mainly involving thalamus and BG might be related to the impaired motor function in patients with AS. Furthermore, BG volume was smaller in the AS group than the control group in VBM analysis. In a murine AS model, ubc3a reinstatement could alleviate the AS-relevant phenotypes.18 However, cerebellum-specific reinstatement of ubc3a in AS mice alone could not rescue the locomotor deficits, which suggests the role of other structures, such as the BG, in locomotor dysfunctions of AS.19

Regarding the language network, interregional FC between the Broca and Wernicke areas in each hemisphere was significantly reduced in the AS group. In TBSS analysis, the AS group also showed significant reductions in FA in the arcuate fasciculus, which connects the Broca and Wernicke areas; this was consistent with previous reports.12 This association between functional and structural alteration in the language network may be associated with impairment of language function in patients with AS. In addition, the language network demonstrated decreased interregional FC with the thalamus and BG, intra- and
FIG 4. Upper 5 plots: boxplots for segmented volume (mm$^3$) of the cerebral white matter and segments of the corpus callosum. Patients with no verbal output (gray box) show significantly reduced volume of aforementioned areas compared with patients who can speak at least a word (black box). Lower 5 plots: boxplots for segmented volume (mm$^3$) of the total gray matter, cortex, corpus callosum (central), right cerebellum cortex, and left amygdala. Patients who were unable to walk alone (light gray box) have significantly reduced volume of aforementioned areas compared with patients who can walk alone (gray box). Asterisks indicate significant difference ($P < .05$). y-axis = mm$^3$. 
neurons and glial cells,24 which tracts, as in a previous report.12 Additional analysis of other FA in the corpus callosum, arcuate fasciculus, and corticospinal impairment of language function in patients with AS. Similarly, research involving the murine model of AS showed that dysfunctional behavioral phenotypes in AS. 

Together with this global reduction of functional and structural connectivity, as well as WM volume reduction in AS, VBM analysis also showed consistent cerebral and cerebellar volume reductions in patients with AS. In the human cortex, *UBE3A* is expressed in both glutamatergic, GABAergic neurons and glial cells,24 which suggests that *UBE3A* dysregulation might affect the whole neuronal circuit. Considering the symptoms of AS (ie, severe speech impairment, unique puppet-like movement, autistic features), the global reduction of brain volume as well as the abnormal FC is deemed reasonable. When we divided the patients into subgroups according to motor and language function to test whether anatomic change can be used as an imaging biomarker for these neurobehaviors, patients with poor language or motor function showed smaller cortical WM or GM volumes. These subgroups consisted of a smaller number of subjects, and further investigation is required. However, these findings provide further evidence that quantitative brain MR imaging may demonstrate functional deficits in AS.

When compared with the age- and sex-matched children with normal development, patients with AS exhibited cortical thinning of the superior temporal and transverse temporal gyri, the area of primary auditory cortex involved in auditory processing, including language sense and social cognition.25 The thinning of the right superior frontal gyrus, which mediates the inhibitory control of impulsive behavior,26 and the supramarginal gyrus, a part of the somatosensory association cortex with a mirror neuron system,27,28 were also observed by cortical thickness analysis.

Table 3: Regional gray matter volume reductions in patients with Angelman syndrome compared with controls based on voxel-based morphometry analysis

<table>
<thead>
<tr>
<th>Anatomic Region</th>
<th>Number of Voxels in Each Cluster</th>
<th>Peak Z Score</th>
<th>Coordinates of Voxel of Maximal Statistical Significance (MNI Space)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left putamen, pallidum, caudate nucleus</td>
<td>18,317</td>
<td>7.73</td>
<td>−21, 2, 12</td>
</tr>
<tr>
<td>Right caudate nucleus</td>
<td>469</td>
<td>6.16</td>
<td>9, 34, −26</td>
</tr>
<tr>
<td>Right putamen, pallidum</td>
<td>7.68</td>
<td>9, 11, 15</td>
<td></td>
</tr>
<tr>
<td>Right rectal gyrus</td>
<td>7.66</td>
<td>21, 5, 12</td>
<td></td>
</tr>
<tr>
<td>Right superior orbital gyrus</td>
<td>159</td>
<td>5.96</td>
<td>−20, −34, −0</td>
</tr>
<tr>
<td>Left rectal gyrus</td>
<td>156</td>
<td>5.95</td>
<td>−51, −48, −36</td>
</tr>
<tr>
<td>Left hippocampus</td>
<td>47</td>
<td>5.67</td>
<td>−46, −43, −50</td>
</tr>
<tr>
<td>Left thalamus</td>
<td>28</td>
<td>5.58</td>
<td>−6, 36, −30</td>
</tr>
<tr>
<td>Left parahippocampal gyrus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left lingual gyrus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left cerebellum (Crus 1, 2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left cerebellum (VII, VIII, crus 2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left rectal gyrus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right rectal gyrus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left superior orbital gyrus</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Whole-brain analysis (FWE-corrected at P < .001; cluster threshold >20 voxels).*

Nonetheless, this study has several limitations. First, the number of patients was small. This syndrome is quite rare; furthermore, recruitment of patients for MR imaging examination is difficult because many parents or guardians of the patients tend to be reluctant to make their children undergo MR imaging because of sleep and behavioral problems, which are closely associated with each other. After sedation, sleep disturbances in patients with AS can be exacerbated by sleep cycle changes, which can lead to behavioral deterioration. Second, different sedation protocols were applied in the control and AS groups. It has been shown that the FC of several brain networks can be influenced by sedative drugs used for MR imaging.29,30 Sedation was inevitable for all patients with AS because of severe communication difficulties; older children in the control group did not need sedation during MR imaging. However, it is unethical to make normal children undergo brain MR imaging under sedation. These heterogeneities in sedation protocol may affect our results, especially rs-fMRI data. Although the pattern of FC difference was similar when we excluded the nonsedated patients, the effect of age, sex, and sedation should be carefully considered. Last, the control group in this study had minor neurologic symptoms, such as headache or febrile seizures, which might also affect our results. Further studies with a homogeneous sedation protocol and control group without any neurologic symptoms are required to confirm our preliminary results.

**CONCLUSIONS**

We found reduced brain connectivity with associated anatomic structural changes in children with AS. Functional and structural underconnectivity, as well as diffuse brain volume reduction, shown by MR imaging may provide striking imaging evidence for the impairment of motor and speech ability seen in individuals with AS.

**ACKNOWLEDGMENTS**

We thank all of the patients and families involved in this study and the biomedical computing core facility at the ConveRgence
mEDicine research cenTer (CREDIT), Asan Medical Center for support and instrumentation.

Disclosures: Woo-Hyun Shim—RELATED: Grant: This research was supported by Basic Science Research Foundation of Korea (NRF) funded by the Ministry of Education [207RDIAB0303731]. Mi-Sun Yum—RELATED: Grant: This research was supported by Asan Institute for Life Sciences, Asan Medical Center [2018-55].* *Money paid to institution.

REFERENCES

ABSTRACT

BACKGROUND AND PURPOSE: Identification of a partial/complete chemotherapy response in pediatric patients with intracranial germ cell tumors is clinically important for radiation treatment and management. Partial/complete response is conventionally determined on postcontrast MR imaging sequences. The purpose of this study was to assess the diagnostic utility of a balanced steady-state free precession sequence as an adjunct to standard MR imaging sequences for the detection of residual tumor in pediatric patients on postchemoreduction pre–radiation planning MR imaging.

MATERIALS AND METHODS: This was a retrospective study of pediatric patients with intracranial germ cell tumors undergoing postchemotherapy, preradiotherapy MR imaging. Patients underwent 1.5T or 3T MR imaging with pre- and postcontrast T1WIs, T2WIs, and a balanced steady-state free precession sequence. Two neuroradiologists independently reviewed standard MR imaging sequences without the balanced steady-state free precession sequence, then with the balanced steady-state free precession sequence 1 week later. Assessment for partial/complete response was determined using Response Assessment in Neuro-Oncology criteria. A 5-point Likert scale scored the diagnostic confidence of the neuroradiologist rating each study without/with the balanced steady-state free precession sequence. Rates of residual disease concordance and diagnostic confidence levels without/with the balanced steady-state free precession sequence were calculated.

RESULTS: Thirty-nine patients were included with 31 males and 8 females (mean age, 14.15 ± 4.26 years). Thirty-one patients had single-site disease; 8 patients had multisynchronous disease (47 sites in total). Compared to review of the standard MR sequences alone, the addition of the balanced steady state free precession sequence resulted in higher rates of tumor partial response categorization and greater diagnostic confidence levels (P < .001, P < .001).

CONCLUSIONS: The balanced steady-state free precession sequence improves detection of residual chemotherapy-reduced intracranial germ cell tumors and increases diagnostic confidence of the neuroradiologist. The balanced steady-state free precession sequence may be an important adjunct to the standard MR imaging protocol for radiation planning.

ABBREVIATIONS: BRAVO = axial MRI 3D brain volume; bSSFP = balanced steady-state free precession; CR = complete response; GCT = germ cell tumor; PR = partial response; SPGR = spoiled gradient echo

Standard of care treatment for CNS germ cell tumors (GCTs) is chemotherapy followed by radiation therapy.1 Treatment is based on the amount of residual tumor after chemotherapy. Patients with a complete response (CR) are defined by response assessment in neuro-oncology (RANO) criteria as having a complete radiographic disappearance of all tumor, whereas those with a partial response (PR) are defined as having a ≥50% decrease (compared with baseline) in the sum of the products of the perpendicular diameters of the tumor.1,2 MR imaging is considered the standard of care for assessment of residual tumor and thus is used to determine CR versus PR.

The balanced steady-state free precession (bSSFP) technique is a heavily water-weighted 3D sequence used with isotropic or near-isotropic spatial encoding, which has been shown to be useful in neuro-oncologic imaging, given its high signal-to-noise ratio and superior discrimination of soft tissue from CSF within the ventricular system or cisterns.3,8 Recent consensus recommendations on MR imaging for the pediatric brain include axial T2-weighted images, axial diffusion-weighted images, and 3D T1 pre- and postcontrast sequences but do not specifically include bSSFP sequences.2
The assessment of CR versus PR is clinically relevant for the management and treatment of these patients. In patients with either CR or PR with residual disease of <1 cm, radiation therapy protocols typically include 18 Gy to the whole brain with a boost of 12 Gy to the primary tumor site. In patients with PR but with residual disease measuring between 1 and 1.5 cm, a higher whole-ventricular radiation dose to 24 Gy with a 12-Gy boost is delivered. Furthermore, if patients with a PR have residual disease measuring >1.5 cm, a second-look surgery is considered to assess for the presence of residual GCT such as mature teratoma versus scar.9,12

The purpose of our study was to evaluate whether the bSSFP sequence improves the detection of residual chemoreduced germ cell tumor over the standard neuro-oncologic MR imaging sequences.

**MATERIALS AND METHODS**

**Study Design**

This retrospective, institutional review board–approved study included pediatric patients who presented to Massachusetts General Hospital between 2010 and 2019 with an intracranial GCT. Patients were identified through a search of the institutional radiology information system and an in-house clinical neuro-oncology data base. Inclusion criteria for this study were patients 21 years of age or younger; a clinical diagnosis of intracranial GCT based on biopsy or CSF; and paired prechemotherapy MR imaging examinations and postchemotherapy, preradiation MR imaging examinations. Standard sequences included pre- and postcontrast T1WI, diffusion-weighted sequences, and T2WI as well as a bSSFP sequence. Patients with prior radiation treatment and nondiagnostic MRIs related to technical factors and motion artifacts were excluded.

**Clinical Data**

The electronic medical record for each patient was reviewed for the following: 1) demographic information, including age and sex; 2) tumor histology (eg, germinoma, yolk sac tumor, mature teratoma, or embryonal carcinoma); 3) tumor site (pineal, suprasellar, multiple midline, posterior fossa, purely intraventricular); and 4) treatment history, including a prior operation (date and biopsy versus debulking), chemotherapy (date and type), and radiation planning dose.

**Scanner Hardware and Sequence Parameters**

Imaging was performed on either a 1.5T (Signa Excite HDx, GE Healthcare; Avanto, Siemens) or 3T scanner (Discovery MR750, GE Healthcare; Tim Trio and Skyra, Siemens). The scans were obtained with an 8-channel brain array coil or a 16-channel head, neck, and spine coil for the Signa Excite HDx; a 20-channel head coil for the Avanto; a 16-channel head, neck, and spine coil for the Discovery MR750; and a 32-channel head coil for the Tim Trio.

**Imaging Protocol**

General anesthesia, child life specialists, and video goggles were used when deemed appropriate and based on our institutional protocols.

The standard acquired sequences obtained included a precontrast 3D T1WI, axial diffusion-weighted imaging, and axial T2WI. Following an intravenous bolus injection of gadoterate meglumine at 0.2 mL/kg (Dotarem; Guerbet), a postcontrast 3D T1WI was performed. Scan parameters such as TE, TR, and flip angle were optimized per patient.

The scan parameters for the 1.5T Avanto scanner for the T1WI (MPRAGE) pre- and postgadolinium were TR/TE = 2140/4.75 to 2200/4.76 ms, NEX = 1, matrix = 256 × 256, flip angle = 15°, FOV = 250 × 250 to 256 × 256 mm, section thickness = 1 mm, gap = 0 mm, with an acquisition time of approximately 5 minutes, 30 seconds. For T2WI, the parameters were TR/TE = 3000/104 to 5510/114 ms, NEX = 2, matrix = 256 × 256 to 448 × 448, flip angle = 180°, FOV = 160 × 160 to 175 × 220 mm, section thickness = 3 mm, gap = 3 mm, with an acquisition time of approximately 4 minutes, 30 seconds.

The scan parameters for the GE 1.5T Signa Excite HDx scanners for the T1WI (axial MRI 3D brain volume [BRAVO; GE Healthcare] and echo-spoiled gradient echo [SPGR]; GE Healthcare) pre- and postgadolinium were TR/TE = 14.66/6.36 to 15.59/6.64 ms for BRAVO and 32/8 ms for SPGR, NEX = 1, matrix = 260 × 260 for BRAVO and 256 × 192 for SPGR, flip angle = 13° for BRAVO and 25° for SPGR, FOV = 260 × 260 mm for BRAVO and 220–340 mm for SPGR, section thickness = 1 mm for BRAVO and 1.5 mm for SPGR, gap = 1 mm for BRAVO and 1.5 mm for SPGR, with an acquisition time of approximately 3 minutes, 30 seconds for BRAVO and 5.9–7.5 minutes for SPGR. For the T2WI, the parameters were TR/TE = 4150/95.5 to 6700/97.2 ms, NEX = 2, matrix = 320 × 256, flip angle = 90°, FOV = 220 × 220 to 240 × 240 mm, section thickness = 3 mm, gap = 3 mm, with an acquisition time of approximately 3–5 minutes, 30 seconds.

The scan parameters for the 3T Siemens Tim Trio scanner and the 3T Skyra scanner for the T1WI (MPRAGE) pre- and postgadolinium were TR/TE = 2530/7.22 ms, NEX = 1, matrix = 256 × 256, flip angle = 7°, FOV = 256 × 256 mm, section thickness = 1 mm, gap = 0 mm, with an acquisition time of approximately 4–6 minutes, 18 seconds. For the T2WI, they were TR/TE = 3890/97 to 8760/97 ms, NEX = 2, matrix = 512 × 294 to 512 × 336, flip angle = 150°, FOV = 256 × 256 mm, section thickness = 3 mm, gap = 3 mm, with an acquisition time of 3 minutes, 6 seconds–4 minutes, 30 seconds.

The scan parameters for the 3T Discovery MR750 scanner for the T1WI (SPGR) pre- and postgadolinium were TR/TE = 8.219/3.22 ms, NEX = 1, matrix = 256 × 256, flip angle = 20°, FOV = 240 × 240 mm, section thickness = 1.2 mm, gap = 1.2 mm, with an acquisition time of approximately 2 minutes, 12 seconds. For the T2WI, they were TR/TE = 4747/84.8 ms, NEX = 2, matrix = 384 × 288, flip angle = 111°, FOV = 220 × 220 mm, section thickness = 3 mm, gap = 3 mm, with an acquisition time of approximately 1 minute, 48 seconds.

The bSSFP sequence on the 3T Tim Trio and Skyra scanners (CISS) was performed using the following parameters: TR/TE = 11.26/5.85 to 12.55/6.28 ms, NEX = 1, matrix = 448 × 269, flip angle = 50°, FOV = 160 × 160 to 199 × 199 mm, section thickness = 0.8 mm, gap = 0 mm, with an acquisition time of approximately 8 minutes.
Table 1: Difference between the recordings with and without the bSSFP sequence for all cases and separately for different locations*

<table>
<thead>
<tr>
<th></th>
<th>Without bSSFP</th>
<th>With bSSFP</th>
<th>Without bSSFP</th>
<th>With bSSFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cases (47)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28 CR (60%)</td>
<td>4 CR (9%)</td>
<td>22 CR (47%)</td>
<td>2 CR (4%)</td>
<td></td>
</tr>
<tr>
<td>19 PR (40%)</td>
<td>43 PR (97%)</td>
<td>25 PR (53%)</td>
<td>45 PR (96%)</td>
<td></td>
</tr>
<tr>
<td>Pineal location (27)</td>
<td>3 CR (11%)</td>
<td>10 CR (37%)</td>
<td>1 CR (4%)</td>
<td></td>
</tr>
<tr>
<td>15 CR (56%)</td>
<td>24 PR (89%)</td>
<td>17 PR (63%)</td>
<td>26 PR (96%)</td>
<td></td>
</tr>
<tr>
<td>Suprasellar location (17)</td>
<td>1 CR (6%)</td>
<td>11 CR (65%)</td>
<td>1 CR (6%)</td>
<td></td>
</tr>
<tr>
<td>12 CR (71%)</td>
<td>16 PR (94%)</td>
<td>6 PR (35%)</td>
<td>16 PR (94%)</td>
<td></td>
</tr>
<tr>
<td>5 PR (29%)</td>
<td>0 CR</td>
<td>2 CR (67%)</td>
<td>0 CR</td>
<td></td>
</tr>
<tr>
<td>Other location (3)</td>
<td>3 PR (100%)</td>
<td>1 PR (33%)</td>
<td>3 PR (100%)</td>
<td></td>
</tr>
</tbody>
</table>

*The recordings of the intraventricular and posterior fossa locations are described with the label “other location.”

The bSSFP sequence on the 3T Discovery MR750 scanners (FIESTA) was performed using the following parameters: TR/TE = 6.878/2.78 ms, NEX = 1, matrix = 320 × 256, flip angle = 65°, FOV = 160 × 160 mm, section thickness = 1 mm, gap = 1 mm, with an acquisition time of approximately 4 minutes, 6 seconds.

The bSSFP sequence on the 1.5T Avanto scanner (CISS) was performed using the following parameters: TR/TE = 5.72/2.4–11.94/5.97 ms, NEX = 1, matrix = 320 × 224, flip angle = 70–80°, FOV 160 × 160 to 170 × 170 mm, slice thickness = 0.7–0.8 mm, gap = 0 mm, with an acquisition time of approximately 5 minutes, 24 seconds.

The bSSFP sequence on the 1.5T Sigma Excite HDx scanner (FIESTA) was performed using the following parameters: TR/TE = 6.048/2.33 to 12.47/2.43 ms, NEX = 1–6, matrix = 448 × 256, flip angle = 65°, FOV = 180 × 180 to 220 × 220 mm, section thickness = 0.8–1.0 mm, gap = 0.4–0.5 mm, with an acquisition time of approximately 2 minutes, 54 seconds–8 minutes, 54 seconds.

**Image Analysis**

The presentation, pretreatment MRIs and the postchemotherapy, preradiation MRIs were reviewed by 2 neuroradiologists (W.A.M. with >5 years’ postfellowship experience and P.C. with >15 years’ postfellowship experience), independently, at 2 separate sessions. During the first session, informed only of the location of the tumor, the 2 neuroradiologists reviewed the cases with standard neuro-oncologic sequences that included axial T2WI and postgadolinium SPGR/MPRAGE/BRAVO sequences, but without the bSSFP sequence. During the second session, at least 1 week later, the same 2 neuroradiologists independently reviewed all sequences, but this time with the addition of the bSSFP sequence. An interval of at least 1 week was chosen between the first and second readout sessions to avoid memory bias.

At both the first and second readout sessions, the 2 neuroradiologists evaluated the primary treatment bed as per Kretschmar et al for CR, defined as no radiographic evidence of disease at the site of the original tumor; PR, defined as ≥50% decrease in the sum of the products of perpendicular diameters of the lesion compared with the presentation MR imaging; stable disease, defined as no appreciable change in tumor volume; and progressive disease, defined as ≥25% increase in the sum of the products of perpendicular diameters.

Cubic tumor volumes (anteroposterior × transverse × superoinferior) were measured on all prechemotherapy, pre-radiation therapy MRIs on the volumetric T1 postgadolinium and bSSFP sequences and on all postchemotherapy, pre-radiation therapy MRIs for which a PR was recorded. Cubic tumor volumes were calculated for all cases on both the volumetric T1 postgadolinium and bSSFP sequences.

A 5-point Likert scale confidence score (1 = not confident at all, 2 = slightly confident, 3 = somewhat confident, 4 = fairly confident, 5 = completely confident) was recorded for each neuroradiologist’s assessment of CR versus PR during both review sessions without-versus-with the inclusion of the bSSFP sequence.

**Statistical Analysis**

A 2-sample test of proportions and a Fisher exact test were performed to analyze differences in confidence scores and assessment of CR versus PR in the bSSFP and non-bSSFP groups. The χ² score was used to assess the level of interrater agreement between the 2 radiologists with SPSS statistics 25 (IBM).

**RESULTS**

**Cohort Description**

A total of 74 patients were considered for this study. Thirty-five patients were excluded, with 31 due to lack of a bSSFP sequence, 2 due to prior radiation therapy, and 2 due to motion artifacts. In this cohort, a total of 39 pediatric patients were included, 31 males and 8 females, ranging from 4 to 21 years of age (mean, 14.15 ± 4.26 years). Twenty-four patients were diagnosed with a pure germinoma, and 15 patients were diagnosed with a nongerminomatous germ cell tumor. All patients had documented negative serum and CSF tumor markers before evaluation for radiation therapy. The nongerminomatous germ cell tumor group contained 1 yolk sac tumor, 1 mature teratoma, 1 embryonal carcinoma, and 12 mixed tumors with germ cell and mature/immature or yolk sac components. Of the 39 patients, 19 had pineal tumor location only, 9 had suprasellar tumor location only, 2 had intraventricular location only, 1 had posterior fossa location only and 8 patients had multiple midline germ cell tumors, including the suprasellar and pineal regions. There were 47 tumor locations in total: Twenty-seven tumors were pineal, 17 tumors were suprasellar, 2 tumors were intraventricular, and 1 tumor was located in the posterior fossa.

**Imaging Protocol**

Twenty-seven patients were scanned on a 1.5T scanner, and 12 patients were scanned on a 3T scanner.
Evaluation of CR and PR

Evaluation of Response during First Review (without the bSSFP). Reader 1 detected 28 instances of CR, 19 PR, 0 stable disease, and 0 progressive disease during the first review in all cases, and reader 2 recorded 22 CR, 25 PR, 0 stable disease, and 0 progressive disease.

Evaluation of Response during Second Review (with the bSSFP). During the second review, with the addition of the bSSFP sequence, reader 1 recorded 4 CR, 43 PR, 0 stable disease, and 0 progressive disease in all cases. Reader 2 recorded 2 CR, 45 PR, 0 stable disease, and 0 progressive disease.

Table 1 shows the difference between the recordings with and without the bSSFP sequence for all cases and separately for different locations.

There were 24 cases (reader 1) and 20 cases (reader 2) that were initially interpreted as CR during the first session of the evaluation and were recategorized as PR during the second review with the addition of the bSSFP sequence (Figs 1 and 2). There was a significant difference between the number of CR and PR evaluations between the recordings with and without the bSSFP sequence ($P < .001$, Fisher exact test performed for both readers).

Tumor Size and Volume Measurements

For the presentation MRIs, the mean cubic tumor volume on the volumetric postcontrast T1WI was 8.88 ± 12.85 cm$^3$ (range, 0.06–68.94 cm$^3$). The mean cubic tumor volume for all PR evaluations on the volumetric postcontrast T1WI was 0.97 ± 1.67 cm$^3$ (range, 0.004–7.54 cm$^3$). The mean cubic tumor volume for the PR evaluations on the bSSFP sequence was 1.31 ± 1.88 cm$^3$ (range, 0.016–8.09 cm$^3$). Although the mean cubic tumor volume on the postchemotherapy preradiation planning MRIs was slightly higher on the bSSFP sequence compared with the postcontrast T1WI, the difference was not significant ($P = .33$).

Of the cases recategorized from CR to PR after review of the bSSFP sequence, the mean tumor volume was 0.24 ± 0.55 cm$^3$ (range, 0.002–2.84 cm$^3$), and 1 of these cases had a volume of >1 cm$^3$.

Of the cases recategorized from CR to PR after review of the bSSFP sequence, 6 tumors had a single plane measurement of ≥1 cm (average single plane measurement, 1.4 ± 3.2 cm).

Confidence Scores

The confidence scores of both sessions and the difference between the interval per location are shown in Table 2. Reader 1 noted a confidence interval of 2–5 (mean, 4.17 ± 0.82) for all cases during

**FIG 1.** A 10-year-old boy with a history of chemotherapy-treated suprasellar germ cell tumor. Sagittal 3D T1 postcontrast SPGR image through the sella (A), interpreted as complete response in the first readout session without the bSSFP sequence. Sagittal FIESTA image through the sella (B) from the same patient interpreted as partial response during the second readout session with small nodules identified along the floor of the third ventricle (arrow).

**FIG 2.** A 19-year-old woman with a history of chemotherapy-treated pineal germ cell tumor. Sagittal 3D T1 postcontrast MPRAGE image through pineal region (A) interpreted as complete response in the first readout session without the bSSFP sequence. Sagittal CISS image through the pineal region (B) from the same patient interpreted as partial response during the second readout session with a >1-cm nodule identified within the pineal gland (arrows).
cases had tumors that demonstrated a residual maximal single-plane measurement of ≥1 cm, which, based on the Children’s Oncology Group ACNS1123 protocol, would have stratified 4 patients for a radiation boost and 2 patients for second-look surgery.9–12 Another important finding is that the addition of a bSSFP sequence improves reader confidence in detecting residual tumor and determining PR-versus-CR status. The average confidence score for the 2 neuroradiologists was 4.17 with the standard MR images and 4.77 with the addition of the bSSFP sequence (P < .001). The number of discrepant interpretations decreased from 12 with the standard MR images to 4 with the inclusion of the bSSFP sequence. Furthermore, interrater agreement also improved from poor (0.128) to excellent (0.872) with the bSSFP sequence.

Both the improved sensitivity for residual tumor and higher reader confidence with the bSSFP sequence when interpreting postchemotherapy-induction MR imaging studies are of paramount importance for therapeutic decision-making. Recategorizing patients from CR to PR is clinically relevant and may influence radiation therapy dosing and treated tumor volumes as well as help determine whether a second-look surgery may be necessary.1,2,9–12

Our findings are similar to those in other previous studies that have demonstrated the benefits of adding bSSFP sequences to standard neuroimaging protocols for CNS tumor detection, including high accuracy for vestibular schwannoma detection and spinal drop metastases.4,13–16

Although the inclusion of a bSSFP sequence increased the total scan length by 6 minutes on average, we found that the improved sensitivity and confidence in the interpretation justify the additional imaging time. Furthermore, there is diagnostic synergy between the bSSFP and the standard MR images when interpreted in conjunction. At our institution, for several years now, we have routinely performed a bSSFP sequence as part of the neuro-oncology MR imaging protocol, given that the pediatric neuroradiologists prefer this sequence for its high spatial resolution, isotropic images, and ability to characterize cystic tumor components. One pitfall of bSSFP, however, is a banding artifact caused by destructive interference between the 2 primary coherent signal pathways in off-resonance regions of B₀ inhomogeneities. This generates dark lines on the images that may at times obscure anatomic structures and lesions.17 CISS is a 2-pass bSSFP technique that attempts to minimize these artifacts. Nevertheless, in 1 case in our sample, there were a large amount of banding artifacts obscuring the pineal region on the bSSFP sequence; this issue was surmounted by simultaneously correlating with the postcontrast T₁-weighted images.

This study has several limitations. The sample size was relatively small. Also, patients were imaged on different MR imaging scanners, which introduced heterogeneity between magnet field strengths (1.5T versus 3T) as well as between the vendor-specific bSSFP sequences (CISS for Siemens versus FIESTA for GE Healthcare). However, each patient’s standard MR images and bSSFP sequence were performed on the same scanner. In no case was a comparison made in the same patient between MR images obtained at different magnetic field strengths or between different vendors. The increased rate of detection of tissue nodularity at the primary tumor site with the addition of the bSSFP sequence may not be clinically important in all patients, particularly given

### Table 2: Confidence scores recorded by the 2 readers for all cases and for individual locations on the bSSFP sequence compared with postcontrast T₁WI

<table>
<thead>
<tr>
<th></th>
<th>Reader 1</th>
<th>Reader 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cases (47)</td>
<td>4.77 ± 0.43 (4–5)</td>
<td>4.47 ± 0.28 (3–5)</td>
</tr>
<tr>
<td>T₁WI</td>
<td>4.36 ± 0.61 (3–5)</td>
<td>4.55 ± 0.58 (3–5)</td>
</tr>
<tr>
<td>bSSFP</td>
<td>4.77 ± 0.43 (4–5)</td>
<td>4.47 ± 0.28 (3–5)</td>
</tr>
<tr>
<td>Pineal location (27)</td>
<td>4.35 ± 0.61 (3–5)</td>
<td>4.29 ± 0.69 (3–5)</td>
</tr>
<tr>
<td>Without bSSFP</td>
<td>4.41 ± 0.67 (3–5)</td>
<td>4.55 ± 0.58 (3–5)</td>
</tr>
<tr>
<td>With bSSFP</td>
<td>4.47 ± 0.28 (3–5)</td>
<td>4.55 ± 0.58 (3–5)</td>
</tr>
<tr>
<td>Suprasellar location (17)</td>
<td>4.36 ± 0.61 (3–5)</td>
<td>4.29 ± 0.69 (3–5)</td>
</tr>
<tr>
<td>Without bSSFP</td>
<td>4.36 ± 0.61 (3–5)</td>
<td>4.29 ± 0.69 (3–5)</td>
</tr>
<tr>
<td>With bSSFP</td>
<td>4.47 ± 0.28 (3–5)</td>
<td>4.55 ± 0.58 (3–5)</td>
</tr>
<tr>
<td>Other location (3)</td>
<td>4.07 ± 0.96 (2–5)</td>
<td>4.41 ± 0.67 (3–5)</td>
</tr>
<tr>
<td>Without bSSFP</td>
<td>4.41 ± 0.67 (3–5)</td>
<td>4.55 ± 0.58 (3–5)</td>
</tr>
<tr>
<td>With bSSFP</td>
<td>4.55 ± 0.58 (3–5)</td>
<td>4.55 ± 0.58 (3–5)</td>
</tr>
</tbody>
</table>

*Data are mean and range.*

the first recording. Reader 2 noted a confidence interval of 3–5 (mean, 4.36 ± 0.61).

During the second review session, with the bSSFP sequence, the mean confidence score for both readers was higher. Reader 1 showed a confidence interval of 4–5 (mean, 4.77 ± 0.43) for all cases. Reader 2 noted a confidence interval of 3–5 (mean, 4.55 ± 0.58).

There was a statistically significant difference between the confidence scores between the reviews performed with the bSSFP sequence compared with reviews performed without bSSFP sequence (P < .001, 2-tailed t test).

### Interrater Reliability

The number of discrepant cases between readers 1 and 2 during the first review, without the bSSFP sequence, was 12. The number of discrepancies between readers 1 and 2 during the second review, with the bSSFP sequence, was 4. For the first review session, without the bSSFP, the calculated $k$ score between reader 1 and reader 2 was 0.128. For the second review session, with the bSSFP, the calculated $k$ score between reader 1 and 2 was 0.872.

### Discussion

Our study demonstrates that the addition of a bSSFP technique to a neuro-oncology MR imaging protocol increases the detection of residual germ cell tumor and improves the interpreting neuroradiologist’s confidence in reporting the presence of residual disease.

Compared with the neuroimaging standard of T₁-weighted postcontrast imaging for assessing tumor response, the bSSFP sequence adds value through improved contrast and higher spatial resolution, which facilitates the detection of nonenhancing or minimally enhancing tumor components. These features allow improved residual germ cell tumor detection, particularly in the typical locations in the suprasellar and pineal regions, which are surrounded by CSF, thus allowing superior contrast resolution between hypointense tumor and hyperintense CSF.

The sensitivity of bSSFP sequences for the detection of small volumes of residual tumor is exemplified by the 24 (reader 1) and 20 (reader 2) cases reclassified from CR to PR by the addition of bSSFP sequences. The residual tumor volumes were, on average, 0.25 cm³, which are small and could be potentially overlooked on standard T₁ postcontrast and T₂/FLAIR sequences. A total of 6

<table>
<thead>
<tr>
<th>Location</th>
<th>Without bSSFP</th>
<th>With bSSFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pineal location</td>
<td>4.36 ± 0.61</td>
<td>4.29 ± 0.69</td>
</tr>
<tr>
<td>Suprasellar location</td>
<td>4.41 ± 0.67</td>
<td>4.55 ± 0.58</td>
</tr>
<tr>
<td>Other location</td>
<td>4.07 ± 0.96</td>
<td>4.41 ± 0.67</td>
</tr>
</tbody>
</table>

CISS is a 2-pass bSSFP technique that attempts to minimize these artifacts. Nevertheless, in 1 case in our sample, there were a large amount of banding artifacts obscuring the pineal region on the bSSFP sequence; this issue was surmounted by simultaneously correlating with the postcontrast T₁-weighted images.

This study has several limitations. The sample size was relatively small. Also, patients were imaged on different MR imaging scanners, which introduced heterogeneity between magnet field strengths (1.5T versus 3T) as well as between the vendor-specific bSSFP sequences (CISS for Siemens versus FIESTA for GE Healthcare). However, each patient’s standard MR images and bSSFP sequence were performed on the same scanner. In no case was a comparison made in the same patient between MR images obtained at different magnetic field strengths or between different vendors. The increased rate of detection of tissue nodularity at the primary tumor site with the addition of the bSSFP sequence may not be clinically important in all patients, particularly given

**Table 2: Confidence scores recorded by the 2 readers for all cases and for individual locations on the bSSFP sequence compared with postcontrast T₁WI**

- **All cases (47)**
  - **T₁WI**: 4.36 ± 0.61 (3–5)
  - **bSSFP**: 4.77 ± 0.43 (4–5)
- **Pineal location (27)**
  - Without bSSFP: 4.07 ± 0.96 (2–5)
  - With bSSFP: 4.74 ± 0.45 (4–5)
- **Suprasellar location (17)**
  - Without bSSFP: 4.35 ± 0.61 (3–5)
  - With bSSFP: 4.94 ± 0.24 (4–5)
- **Other location (3)**
  - Without bSSFP: 4.00 ± 0.00 (4)
  - With bSSFP: 4.00 ± 0.00 (4)

**Data are mean and range.**
that most residual tumors detected were <1 cm. Given the investigational nature of our current study, additional studies are needed to determine whether these areas of presumed residual tumor identified on bSSFP imaging impact patient management and outcomes. Furthermore, there was a lack of histopathologic confirmation in most cases. Postsurgical and postchemotherapy scar tissue may have been misinterpreted as residual tumor, thus resulting in false-positives. However, histopathologic confirmation is not standard of care.

CONCLUSIONS
Adding a bSSFP sequence to the standard neuro-oncology MR imaging protocol improves the neuroradiologist’s detection of presumed residual chemotherapy-reduced CNS germ cell tumor and increases confidence in distinguishing partial response from complete response disease status. This distinction could have important therapeutic consequences for the patient with regard to treatment options from the Response Assessment in Pediatric Neuro-Oncology committee. The contribution of 3D-CISS and contrast-enhanced MR cisternography in detecting cerebrospinal fluid leak in patients with rhinorrhea. Addition of contrast-enhanced MR cisternography to the evaluation of the aqueductal stenosis. The contribution of 3D-CISS and contrast-enhanced MR cisternography in detecting cerebrospinal fluid leak in patients with rhinorrhea. Br J Radiol 2019;83:225–32 CrossRef Medline


REFERENCES
ABSTRACT

BACKGROUND AND PURPOSE: Constitutional mismatch repair deficiency is a hereditary childhood cancer predisposition syndrome characterized by brain tumors and colorectal and hematologic malignancies. Our objective was to describe the neuroimaging findings in patients with constitutional mismatch repair deficiency.

MATERIALS AND METHODS: This retrospective study included 14 children with genetically confirmed constitutional mismatch repair deficiency who were referred to 2 tertiary pediatric oncology centers.

RESULTS: Fourteen patients from 11 different families had diagnosed constitutional mismatch repair deficiency. The mean age at presentation was 9.3 years (range, 5–14 years). The most common clinical presentation was brain malignancy, diagnosed in 13 of the 14 patients. The most common brain tumors were glioblastoma (n = 7 patients), anaplastic astrocytoma (n = 3 patients), and diffuse astrocytoma (n = 3 patients). Nonspecific subcortical white matter T2 hyperintensities were noted in 10 patients (71%). Subcortical hyperintensities transformed into overt brain tumors on follow-up imaging in 3 patients. Additional non-neoplastic brain MR imaging findings included developmental venous anomalies in 12 patients (85%) and nontherapy-induced cavernous hemangiomas in 3 patients (21%).

CONCLUSIONS: On brain MR imaging, these patients have both highly characteristic intra-axial tumors (typically multifocal high-grade gliomas) and nonspecific findings, some of which might represent early stages of neoplastic transformation. The incidence of developmental venous anomalies is high in these patients for unclear reasons. Awareness of these imaging findings, especially in combination, is important to raise the suspicion of constitutional mismatch repair deficiency in routine diagnostic imaging evaluation or surveillance imaging studies of asymptomatic carriers because early identification of the phenotypic “gestalt” might improve outcomes.

ABBREVIATIONS: CMMRD = constitutional mismatch repair deficiency; DVA = developmental venous anomaly; WHO = World Health Organization.

DNA replication is a highly controlled process during cell division. The DNA mismatch repair system primarily functions to correct errors arising during DNA replication. Hence, mutations in mismatch repair genes and their dysfunctional protein products lead to accumulation of other “unrepaired” mutations in each cell division, which could cause neoplastic (malignancies, premalignancies, and benign tumors) or non-neoplastic phenotypic features to develop. In humans, germline mutations in the mismatch repair genes result in distinct cancer predisposition syndromes. Heterozygous germline mutations in 1 of the 4 mismatch repair genes (mutL homolog 1 [MLH1], PMS1 homolog 2, mismatch repair system component [PMS2], mutS homolog 2 [MSH2], or mutS homolog 6 [MSH6]) result in Lynch syndrome, which is the most frequent form of inherited colorectal cancer. Lynch syndrome manifests in early adulthood and is characterized by an increased incidence of endometrial and gastrointestinarian cancer. Biallelic germline mutations in the mismatch repair genes result in a distinct phenotypically...
defined constitutional mismatch repair deficiency syndrome (CMMRD).\(^5,6,8\) Individuals with biallelic mismatch repair gene mutations are at risk of developing a vast spectrum of malignancies carrying a dismal prognosis during childhood. Typical malignancies include brain tumors, digestive tract tumors, hematologic malignancies, and Lynch syndrome–associated tumors other than in the digestive tract.\(^2,5,6,9\) Synchronous and metachronous cancers are common.\(^10,11\) A proposed scoring system based on clinical features was developed by the European Consortium “Care for CMMRD”\(^5,15\) (Table 1), indicating which patients should be further tested for CMMRD.

Early diagnosis of CMMRD could have considerable clinical importance, not only for surveillance of the index patient but also for screening in the entire pedigree.\(^5\) CMMRD should be considered in children with multiple malignancies and café au lait spots without the genetic diagnosis of neurofibromatosis type 1.\(^12\) Remarkably, in children with CMMRD, both parents almost always have Lynch syndrome. Because genetic testing has considerably improved in recent years, our ability to confidently diagnose CMMRD has increased. Surveillance guidelines are also emerging (eg, European Consortium “Care for CMMRD,”\(^9\) US Multi-Society Task Force on Colorectal Cancer,\(^4\) and American Association of Cancer Research).\(^8\) Recently, a remarkable response to immune checkpoint inhibitors has been reported in patients with CMMRD with glioblastoma.\(^11\) Hence, early detection of brain tumors and precancerous conditions might offer opportunities to improve outcomes.\(^5,7,13\)

Our objective is to describe neuroimaging findings in children with both CMMRD and brain tumors, thus increasing neuroradiologists’ awareness of this newly discovered cancer predisposition syndrome.

## MATERIALS AND METHODS

### Study Design and Patient Eligibility

Children with CMMRD diagnosed in Sheba Medical Center, Tel Hashomer, Ramat Gan, Israel and St. Jude Children’s Research Hospital, Memphis, Tennessee between 2014 and 2019 were included in this retrospective study. Review of the patients’ medical records provided demographic, clinical, and genetic data. Brain MR imaging studies were reviewed by board-certified neuroradiologists in each participating institution. This retrospective study was approved by the institutional review boards in both hospitals.

### Imaging Studies

MR imaging was performed on the following platforms: 1.5T (Magnetom Avanto or Magnetom Essenza, Siemens; Optima MR450w, GE Healthcare) or 3T (Ingenia, Philips Healthcare; Signa, GE Healthcare; Magnetom Prisma or Skyra, Siemens). Routine brain imaging protocols were different at each institution but always included multiplanar nonenhanced T1-weighted images, T2-weighted images, diffusion-weighted images, and postcontrast T1-weighted sequences, which were occasionally and inconsistently complemented by precontrast T2-FLAIR, SWI/T2*, and postcontrast T2-FLAIR sequences. Acquisition parameters differed according to the specific MR imaging vendors.

### Statistical Analysis

Descriptive statistics were used to analyze data. Continuous variables are presented as mean and range.

## RESULTS

### Clinical Data and Tumor Types

Fourteen patients from 11 families with diagnosed CMMRD were included in this retrospective case series (5 boys and 9 girls). A summary of the patients’ clinical and genetic features is presented in the On-line Table. Twelve patients (86%) had MSH6 mutations and 2 (14%) had PMS2 mutations. Consanguinity was found in 4 of the 11 families. The mean age at diagnosis was 9.3 years (range, 5–14 years). Multiple synchronous and metachronous malignancies were common. Eighteen brain tumors were found in 13 of 14 patients (93%). Colorectal adenocarcinoma was diagnosed in 4

### Table 1: Indication criteria for CMMRD testing in patients with cancer*  

<table>
<thead>
<tr>
<th>Indication for CMMRD testing</th>
<th>Criteria</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malignancies/premalignancies: 1 is mandatory; if &gt;1 is present in the patient, add the points</td>
<td>Carcinoma from the LS spectrum at younger than 25 years of age</td>
<td>3 points</td>
</tr>
<tr>
<td>Multiple bowel adenomas at younger than 25 years of age and absence of APC/MUTYH mutations or a single high-grade dysplasia adenoma at younger than 25 years of age</td>
<td>WHO grade III or IV glioma at younger than 25 years of age</td>
<td>3 points</td>
</tr>
<tr>
<td>Any malignancy at younger than 18 years of age</td>
<td>NHL of T-cell lineage or sPNET at younger than 18 years of age</td>
<td>2 points</td>
</tr>
<tr>
<td>Additional features: optional; if &gt;1 of the following is present, add the points</td>
<td>Clinical sign of NF1 and/or ≥2 hyperpigmented and/or hypopigmented skin alterations Ø 1 cm in the patient</td>
<td>2 points</td>
</tr>
<tr>
<td>Diagnosis of LS in a first-degree, second-degree, and third-degree relative</td>
<td>Carcinoma from LS spectrum before 60 years of age in first-degree, second-degree, and third-degree relatives</td>
<td>1 point</td>
</tr>
<tr>
<td>A sibling with carcinoma from the LS spectrum, high-grade glioma, sPNET, or NHL</td>
<td>A sibling with any type of childhood malignancy</td>
<td>2 points</td>
</tr>
<tr>
<td>Multiple pilomatrixomas in the patient</td>
<td>One pilomatrixoma in the patient</td>
<td>1 point</td>
</tr>
<tr>
<td>Agnogenicity of the corpus callosum or nontherapy-induced cavernoma in the patient</td>
<td>Consanguineous parents</td>
<td>1 point</td>
</tr>
<tr>
<td>Deficiency/reduced levels of immunoglobulin (lg)G2/4 and/or lgA</td>
<td></td>
<td>1 point</td>
</tr>
</tbody>
</table>

Note: LS indicates Lynch syndrome; NHL, non-Hodgkin lymphoma; sPNET, supratentorial primitive neuroectodermal tumor; NF1, neurofibromatosis type I.  
Reprinted with permission from Wimmer et al.\(^5\)  
*Colorectal, endometrial, small bowel, ureter, renal pelvis, biliary tract, stomach, bladder carcinoma.
patients (29%), and 2 patients had multiple colorectal polyps with no high-grade features. Hematologic malignancies were diagnosed in 3 of 14 patients (21%), 2 with non-Hodgkin lymphomas of the T-cell lineage and 1 with B-cell acute lymphoblastic leukemia. Café au lait spots were observed on physical examination on all patients in our cohort. The mean surveillance period was 4.6 years (range, 1–9 years).

**Neoplastic Neuroimaging Findings**

Glioblastoma (World Health Organization [WHO] IV) was the most common brain malignancy in our cohort and was diagnosed in 7 patients (50%). All these tumors were supratentorial. On MR imaging, the lesions were irregular, poorly marginated, and resulted in marked mass effect. Typically, these tumors showed low-T1 and heterogeneous high-T2 or T2-FLAIR signal, with thick irregular peripheral enhancement surrounding central necrosis (Fig 1). The enhancing margins showed some degree of diffusion restriction. In 1 patient, glioblastoma showed irregular patchy areas of enhancement without characteristic central necrosis (Fig 2). Prominent T2 flow voids (ie, prominent feeding vessels) were also noted.

Supratentorial diffuse astrocytoma (WHO II) was diagnosed in 3 patients. On imaging, infiltration and expansion of the overlying cortex were noted. No contrast enhancement or restricted diffusion was identified in these cases. Supratentorial anaplastic astrocytoma (WHO III) was diagnosed in another 3 patients. One tumor of the latter group was partly intraventricular. Diffuse brain stem tumors were found in 2 patients (1 with diffuse intrinsic pontine glioma and 1 with glioblastoma). Pleomorphic xanthoastrocytoma (WHO II) involving the temporal lobe was diagnosed in 1 patient. Sonic hedgehog (SHH)-activated medulloblastoma (WHO IV) was diagnosed in 1 patient.

Spinal imaging, performed routinely as part of the surveillance protocol, did not show any structural or neoplastic spinal cord lesions.

**Nonspecific, Subcortical T2/T2-FLAIR Hyperintensities**

In our cohort, 10 of 14 patients (71%) had single or multiple nonspecific patchy subcortical white matter T2-FLAIR hyperintensities, most commonly in the frontal and parietal lobes, but occasionally in the cerebellum and/or the brain stem (Table 2). No mass effect, enhancement, or restricted diffusion were noted in these lesions (Figs 1 to 3). Most of the lesions were stable or showed very modest changes during the surveillance period (Fig 4). In 3 patients, a gradual transformation toward overt brain tumors was observed on surveillance studies (2 with glioblastomas and 1 with diffuse astrocytoma; Figs 2 and 3, respectively). Occasionally, additional new lesions emerged during the follow-up period, too.

**Vascular Abnormalities and Focal Areas of Signal Intensity**

Developmental Venous Anomaly. Twelve of the 14 patients (86%) presented with a network of prominent medullary veins converging into a single draining vein (caput medusae), meeting the conventional MR imaging criteria of developmental venous anomaly (DVA). These DVAs drained into the dural sinuses or ependymal veins. The average number of DVAs in patients who had them was 2.5 (range, 1–5). These DVAs were already present at the initial diagnostic imaging evaluation. DVAs were either supra- (mostly frontoparietal) or infratentorial (cere-
bellum and pons) or in both locations. No spatial association was found between DVAs and other parenchymal abnormalities (overt brain tumors, subcortical T2/T2-FLAIR hyperintensities) (Fig 5). These findings were seen on initial imaging with no new DVAs identified during follow-up.

Non–Therapy-Related Cavernous Hemangioma. Popcorn-like nonenhancing lesions with prominent T2* SWI blooming artifacts suggestive of cavernous hemangiomas were identified in 3 of 14 patients (21%). These findings were seen on initial imaging with no new nontherapy-induced cavernomas identified during follow-up. After radiation therapy, several patients showed multiple foci of microhemorrhage, yet these were not considered to be primary cavernous hemangiomas. Not all patients had T2* or SWI sequences included in the MR imaging protocol.

Table 2: Nonmalignant neuroradiologic findings

<table>
<thead>
<tr>
<th>Patient</th>
<th>T2-FLAIR Subcortical Hyperintensities</th>
<th>Focal Areas of T2-FLAIR Hyperintense Signal Intensity (Basal Ganglia, Thalamus, Pons, or Cerebellum)</th>
<th>Developmental Venous Anomalies</th>
<th>Cavernous Hemangioma</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Parietal</td>
<td>—</td>
<td>Cerebellum</td>
<td>Frontal periventricular</td>
</tr>
<tr>
<td>2</td>
<td>—</td>
<td>—</td>
<td>Frontal, cerebellum</td>
<td>—</td>
</tr>
<tr>
<td>3</td>
<td>Frontal</td>
<td>—</td>
<td>Frontal, cerebellum</td>
<td>Pons</td>
</tr>
<tr>
<td>4</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>Frontal, parietal</td>
<td>+</td>
<td>Frontoparietal</td>
<td>—</td>
</tr>
<tr>
<td>6</td>
<td>Frontal, parietal, occipital</td>
<td>—</td>
<td>Insula</td>
<td>Temporal</td>
</tr>
<tr>
<td>7</td>
<td>Frontal, temporal</td>
<td>—</td>
<td>Pons</td>
<td>—</td>
</tr>
<tr>
<td>8</td>
<td>—</td>
<td>—</td>
<td>Frontal</td>
<td>—</td>
</tr>
<tr>
<td>9</td>
<td>—</td>
<td>—</td>
<td>Frontal, temporal</td>
<td>—</td>
</tr>
<tr>
<td>10</td>
<td>Frontal, parietal, temporal, occipital</td>
<td>+</td>
<td>Frontal, parietal</td>
<td>—</td>
</tr>
<tr>
<td>11</td>
<td>Temporal, cerebellar peduncle</td>
<td>+</td>
<td>Frontal, temporal, cerebellum</td>
<td>—</td>
</tr>
<tr>
<td>12</td>
<td>Cerebellum, cerebellar peduncle, frontal brain stem</td>
<td>+</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>13</td>
<td>Frontal, parietal</td>
<td>—</td>
<td>Frontal, temporal</td>
<td>—</td>
</tr>
<tr>
<td>14</td>
<td>Frontal</td>
<td>—</td>
<td>Frontal</td>
<td>—</td>
</tr>
</tbody>
</table>

Note: — indicates absent; +, present.

FIG 3. Surveillance brain MR imaging studies of a boy with CMMRD (patient 2). Serial T2-FLAIR imaging shows nonspecific subcortical hyperintensities (arrows) slowly increasing in size (10 years of age, A; 15 years of age, B). At 18 years of age (C), there is an infiltrative lesion involving the right insula (asterisk) and base of the right frontal lobe. Biopsy results indicated diffuse astrocytoma (WHO II). Three months later (D), marked increase in the left frontotemporal lesion was obvious along with central necrosis (arrowhead). This lesion was resected, and pathologic testing indicated glioblastoma.

FIG 4. MR images of a girl with CMMRD (patient 6) and glioblastoma (not shown). A gradual increase in frontoparietal subcortical T2-FLAIR hyperintensities (arrows) is noted during 2 years of surveillance (7 years of age, A; 8 years of age, B; and 9 years of age, C).
Focal Areas of Signal Intensity. Four patients showed nonenhancing T2/T2-FLAIR hyperintense white matter lesions within the supratentorial deep gray matter nuclei (globus pallidus and thalamus), similar to patchy T2/T2-FLAIR hyperintensities commonly seen in neurofibromatosis type 1 (Fig 6). All 4 patients were younger than 15 years of age. All lesions persisted during follow-up.

DISCUSSION

Here, we report the neuroimaging findings in a series of patients with CMMRD. As previously published, brain tumors, particularly high-grade gliomas, are common in patients with CMMRD. In children, high-grade gliomas are usually primary, rarely originating from pre-existing low-grade tumors. Recent data based on molecular abnormalities found in patients with cancer predisposition syndromes suggest that high-grade gliomas in that setting might be secondary (ie, following malignant transformation of low-grade gliomas), as seen typically in young adult patients. Subcortical T2/T2-FLAIR hyperintensities described in patients with CMMRD in our study might support these findings because some of the high-grade gliomas eventually emerged from these lesions, possibly as a result of the relentlessly increasing mutation burden during the disease course. Therefore, we hypothesize that accumulation of somatic mutations in patients with CMMRD leads to these low-grade/precancerous lesions, similar to low-grade dysplastic polyps found on gastrointestinal endoscopy in patients with CMMRD, which may evolve to malignancies. Unfortunately, histopathologic analysis of these ill-defined subcortical lesions is lacking because no biopsies of those lesions were performed in our patients.

High-grade glioma was the most prevalent brain tumor in our patients with CMMRD, as in previous reports. Nevertheless, low-grade gliomas were not uncommon (found in 21% of our patient cohort). Supratentorial embryonal tumors, which have been reported previously in patients with CMMRD, were not found in our series. The median age of our patients at diagnosis of brain tumors was 9.3 years (range, 5–14 years of age), as has been previously reported. The imaging characteristics of various brain tumors in our series resembled those in brain tumors in patients without CMMRD. Nevertheless, recent genetic evaluation of pediatric brain tumors shows that in patients with CMMRD, mutations in mismatch repair genes may lead to accumulation of multiple somatic mutations, resulting in considerably higher tumoral mutation burden than that in non-CMMRD childhood gliomas.

In humans, most CMMRD cases are due to MSH6 or PMS2 mutation. Similarly, in our cohort, 12 patients had MSH6 mutations (86%) and 2 (14%) had PMS2 mutations. Individuals with these mutations in the context of Lynch syndrome exhibit low penetrance, which might partially explain the negative family history in our cohort. Additionally, patients with PMS2 and MSH6 mutations have a higher reported prevalence of brain tumors than do those with other mismatch repair mutations, and these patients are more likely to survive their first tumor and develop a metachronous tumor.

Individuals with CMMRD frequently display neurofibromatosis type 1 stigmata, such as café au lait spots. All the patients in our cohort presented with café au lait spots. All the patients in our cohort presented with café au lait spots. Brain lesions similar to the focal areas of signal intensity, commonly seen in prepubertal patients with neurofibromatosis type 1, were identified in 4 of our patients. All 4 showed lesions before 15 years of age, in a timeline similar to that of neurofibromatosis type 1, yet all persisted during the available follow-up. The associations between CMMRD and neurofibromatosis type 1 have been described by Wimmer et al, and a presumed explanation is that common somatic mutations in the neurofibromin 1 (NFI) gene are acquired as a consequence of the mismatch repair malfunction.

Structural brain abnormalities were reported by Baas et al in a case series of 4 patients with CMMRD with agenesis of the corpus callosum and gray matter heterotopia. Such malformative abnormalities were not found in our cohort.

Additional non-neoplastic brain MR imaging findings in CMMRD include DVAs and nontherapy-induced cavernous hemangiomas, as have been recently reported. Remarkably, 12 of
our 14 patients had DVAs (typically multiple DVAs), indicating a considerably higher prevalence than that reported in healthy populations (6%–7%). Because DVAs are considered a normal variant, under-reporting is probably common. However, patients with primary brain tumors have a higher prevalence of DVAs than the healthy population (10.2% versus 5.3%, respectively). In our series of children with brain tumors and CMMRD, there was no apparent association between the tumor location and the location of the DVA. DVAs have a well-recognized association with cavernomas, which was also noted in some of our patients. The pathogenesis of these vascular anomalies is not clearly understood. DVAs might arise as a response to compromise of the normal venous drainage by stenosis, thrombosis, or maldevelopment. Cavernomas might occur in both sporadic and familial forms. Patients with familial cavernomas typically have multiple malformations. Familial cavernomas have been linked to 3 specific cerebral cavernous malformation genes KRIT1/CCM1 (Krev interaction trapped 1), MGC4607/CCM2 (cerebral cavernous malformation 2), or PDCD10/CCM3 (programmed cell death protein 10). Almost 80% of patients with the hereditary form have a heterozygous germline mutation in 1 of these 3 genes. Development of sporadic cavernomas is thought to be secondary to microhemorrhages and angiogenic growth factor activation in a cause-and-effect relationship with adjacent DVA. Recognizing vascular abnormalities in the brain, particularly DVAs, is important; hence, they seem to contribute to a potentially suggestive imaging gestalt in children with CMMRD.

CMMRD should be clinically suspected in children with brain tumors who have café au lait spots, a family history of Lynch syndrome, or a sibling with childhood cancer manifesting as brain tumors and several nonspecific findings, some of which might be precancerous; however, others, including non-neoplastic vascular anomalies, may also contribute to a suggestive imaging gestalt. Awareness of these imaging findings is important at both the initial diagnostic imaging evaluation of children with suspected brain tumors and at that of those known to carry the genetic predisposition and requiring regular surveillance studies because early detection of brain neoplasms may improve clinical outcomes.

CONCLUSIONS

On brain imaging, patients with CMMRD present with overt tumors and several nonspecific findings, some of which might be precancerous; however, others, including non-neoplastic vascular anomalies, may also contribute to a suggestive imaging gestalt. Awareness of these imaging findings is important at both the initial diagnostic imaging evaluation of children with suspected brain tumors and at that of those known to carry the genetic predisposition and requiring regular surveillance studies because early detection of brain neoplasms may improve clinical outcomes.

Disclosures: Amar Gajjar—UNRELATED: Consultancy; Roche Pharmaceuticals. Kim E. Nichols—UNRELATED: Grants/Grants Pending: Incyte. Comments: small research grant to support work ongoing in basic laboratory.

REFERENCES

Utility of Pre-Hematopoietic Cell Transplantation Sinus CT Screening in Children and Adolescents


ABSTRACT

BACKGROUND AND PURPOSE: The clinical benefit of pre-hematopoietic cell transplantation sinus CT screening remains uncertain, while the risks of CT radiation and anesthesia are increasingly evident. We sought to re-assess the impact of screening sinus CT on pretransplantation patient management and prediction of posttransplantation invasive fungal rhinosinusitis.

MATERIALS AND METHODS: Pretransplantation noncontrast screening sinus CTs for 100 consecutive patients (mean age, 11.9 ± 5.5 years) were graded for mucosal thickening (Lund-Mackay score) and for signs of noninvasive or invasive fungal rhinosinusitis (sinus calcification, hyperattenuation, bone destruction, extrasinus inflammation, and nasal mucosal ulceration). Posttransplantation sinus CTs performed for sinus-related symptoms were similarly graded. Associations of Lund-Mackay scores, clinical assessments, changes in pretransplantation clinical management (additional antibiotic or fungal therapy, sinonasal surgery, delayed transplantation), and subsequent development of sinus-related symptoms or invasive fungal rhinosinusitis were tested (exact Wilcoxon rank sums, Fisher exact test, significance P<.05).

RESULTS: Mean pretransplantation screening Lund-Mackay scores (n=100) were greater in patients with clinical symptoms (8.07 ± 6.00 versus 2.48 ± 3.51, P<.001) but were not associated with pretransplantation management changes and did not predict posttransplantation sinus symptoms (n=21, P=.47) or invasive fungal rhinosinusitis symptoms (n=2, P=.59).

CONCLUSIONS: Pre-hematopoietic cell transplantation sinus CT does not meaningfully contribute to pretransplantation patient management or prediction of posttransplantation sinus disease, including invasive fungal rhinosinusitis, in children. The risks associated with CT radiation and possible anesthesia are not warranted in this setting.

ABBREVIATIONS: ENT = ear, nose, and throat; GVHD = graft versus host disease; HCT = hematopoietic cell transplantation; IFRS = invasive fungal rhinosinusitis

Due to prolonged immunosuppression, children undergoing hematopoietic cell transplantation (HCT) are at increased risk of opportunistic infections, including potentially lethal invasive fungal rhinosinusitis (IFRS).1-5 Children at St. Jude Children’s Research Hospital, therefore, undergo rigorous pretransplantation evaluation, including ear, nose, and throat (ENT) examination, infectious disease risk assessment, fungal serologies, and noncontrast sinus CT.

However, the clinical utility of screening sinus CT remains unclear, and the practice is not universal. According to the American Academy of Pediatrics and the American College of Radiology, uncomplicated sinusitis should be a clinical diagnosis,6,7 and several large studies (n≥100) have failed to identify pre-HCT imaging features predictive of post-HCT sinusitis.8-10 Fewer studies have evaluated the contribution of pretransplantation CT to patient management or prediction of posttransplantation IFRS. A small study in adults found that screening sinus CT did not contribute to a pretransplantation diagnosis or management or predict posttransplantation sinusitis or IFRS, but it was limited by a small sample size.11 A study of 187 children reported an association between pre-HCT sinus CT findings and management changes, but while 119 (64%) had abnormal sinus CT findings, only 29 had symptoms and were treated, suggesting symptoms, not imaging, drove treatment.12

Received June 3, 2019; accepted after revision March 3, 2020.

From the Departments of Diagnostic Imaging (J.H.H., R.A.K.), Biostatistics (G.K.), Infectious Diseases (G.M.), Bone Marrow Transplantation and Cellular Therapy (W.M., A.S.), and Surgery (J.W.T.), St. Jude Children’s Research Hospital, Memphis, Tennessee; and Department of Otolaryngology (J.W.T.), University of Tennessee Health Sciences Center, Memphis, Tennessee.

This work was supported, in part, by grant No. CA21765 from the National Cancer Institute and by the American Lebanese and Syrian Associated Charities.

Please address correspondence to Julie H. Harreld, MD, Department of Diagnostic Imaging, St. Jude Children’s Research Hospital, 262 Danny Thomas Place, MS 220, Memphis, TN 38105; e-mail: Julie.harreld@stjude.org

http://dx.doi.org/10.3174/ajnr.A6509

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

While the benefits of pre-HCT sinus CT remain unclear, the risks associated with sinus CT are well-established and include a low-but-real risk of radiation-induced cancers\textsuperscript{13,14} and early cataract formation.\textsuperscript{15} Evidence is also emerging that anesthe sia, required by some children to complete CT, may negatively impact cognitive development.\textsuperscript{16,17} On the other hand, advances in fungal prophylaxis have significantly reduced the incidence of invasive fungal infections in immunocompromised children.\textsuperscript{18}

Given these potential shifts in the risk-benefit ratio, we sought to re-evaluate the clinical utility of pre-HCT screening sinus CT for IFRS risk assessment and its role in pre-HCT management of pediatric patients.

**MATERIALS AND METHODS**

With institutional review board approval and waiver of consent, imaging, clinical, ENT risk assessment, and laboratory data were prospectively recorded for 100 sequential participants who had pre-HCT screening sinus CT at our institution between June 2015 and April 2017. A sample size of 100 was chosen to detect at least 1 case of IFRS, which has an incidence of approximately 2\% in patients with hematologic disorders.\textsuperscript{2} The patients’ medical records were reviewed for microbiologic diagnosis of IFRS for at least 100 days posttransplantation. With additional institutional review board approval and waiver of consent, medical records were retrospectively reviewed for subsequent clinical, imaging, and laboratory data until the patient’s death or November 2017.

**Clinic and Laboratory**

Recorded data included primary diagnosis; transplant donor type; absolute neutrophil count at imaging; pretransplantation imaging indication; development of posttransplantation graft versus host disease (GVHD) grades II–IV (moderate-to-life-threatening); pretransplantation galactomannan (Aspergillus antigen); sinonasal symptoms (rhinorrhea, congestion, nasal/facial pain, swelling, visible nasal lesion), or fever at pretransplantation evaluation; and changes to pretransplantation management (changes in fungal prophylaxis regimen, addition of antibiotics, sinonasal operation, or delay of HCT) attributable to screening sinus CT.

ENT pretransplantation risk assessment included visualization of the nasal septum, palate, and back of throat under magnification; and complete allergy and sinusitis histories. These findings were summarized as “at-risk” or “not-at-risk” for IFRS.

For patients requiring post-HCT sinus imaging for symptoms, we recorded symptoms, galactomannan serologies, endoscopy findings, fungal organism (if applicable), and consecutive days of neutropenia (absolute neutrophil count < 500/mm\textsuperscript{3}) immediately preceding imaging. The date of acquisition of tissue positive for IFRS by microbiologic/pathologic examination was considered the date of diagnosis.

**Imaging**

CTs were axially acquired with a section thickness of \(\leq 2.5\) mm on a LightSpeed VCT or Revolution CT scanner (GE Healthcare) with bone and soft-tissue reconstructions. The average effective dose for all ages was \(0.8\) mSv; the eye/lens dose was \(0.7–1.2\) mGy.

Images were reviewed in consensus by a board-certified pediatric radiologist with 40 years’ experience (R.A.K.) and a board-certified neuroradiologist with 10 years’ experience (J.H.H.) for the presence or absence of bone destruction, extrasinus inflammation, and nasal mucosal ulceration, associated with IFRS.\textsuperscript{4,19-22} Because noninvasive fungal disease could theoretically predispose to IFRS with pre-HCT conditioning/immunodepletion, the presence of calcifications and sinus hyperdensity was also noted.\textsuperscript{23}

Each anterior and posterior ethmoid, maxillary, frontal, and sphenoid sinus was graded as clear = 0, partially opacified = 1, or completely opacified = 2; each ostiomeatal unit was graded as clear = 0 or occluded = 2; and the numbers were added per Lund and Mackay.\textsuperscript{24} To account for age-related differences in sinus development, we calculated an adjusted Lund-Mackay score as\textsuperscript{25}

\[
\text{Adjusted Lund Score} = \frac{\text{Raw score} \times \text{Maximum Lund for No. Pneumatized Sinuses}}{24}
\]

**Statistical Analysis**

Sinus CT scores, laboratory values, symptoms, changes in management, and ENT risk were compared using the Fisher exact test or the Pearson \(\chi^2\) test for categoric variables and a 2-sample \(t\) test or (exact) Wilcoxon rank sum test for continuous variables, based on the normality assumption. The Lund-Mackay scores between pre- and posttransplantation were compared using a 1-sample \(t\) test or Wilcoxon signed rank test, depending the normality assumption. All \(P\) values are 2-sided and were considered statistically significant if <.05. Statistical analyses were performed with R-3.6.1 (R statistical and computing software; http://www.r-project.org/).

**RESULTS**

One hundred participants (60 males, 40 females; 8 months to 24 years of age; mean, 11.9 ± 5.5 years; males, 11.6 ± 5.7 years; females, 12.5 ± 5.2 years) had screening sinus CT an average of 24.6 ± 9.6 days before transplantation for IFRS risk assessment. Patient characteristics, imaging, and clinical findings are summarized in Table 1.

Follow-up for survivors \((n = 82)\) ranged from 189 to 889 days post-HCT. During the follow-up period, 9 patients died of relapsed disease from 28 to 519 days post-HCT. Nine died of complications between 23 and 557 days post-HCT, none related to sinus disease.

**Clinical Symptoms and Provided Imaging Indications**

At the time of the screening sinus CT, 18 patients had sinonasal symptoms documented in the medical record: Thirteen had rhinorrhea, 2 had nasal congestion, and 3 had both. No patients had overt signs or symptoms of IFRS or complicated rhinosinusitis—localized facial pain or inflammation, nasal ulcer/eschar, fever, or altered mental status—documented in the medical record at the time of CT.
The provided indication for 98/100 screening sinus CTs was “pretransplantation evaluation.” Symptoms mentioned for 2 patients did not match the medical record, possibly being erroneous or outdated. One request indicated fever without localizing symptoms, though the clinical examination documented congestion and rhinorrhea without fever. The other request indicated possible periorbital edema, which was not present on CT or the physical examination. On examination, the patient had clear rhinorrhea and nasal wash positive for enterovirus/rhinovirus and was determined eligible for transplantation the same day.

**Pretransplantation Screening and Management**

Mucosal thickening was present (Lund-Mackay score of ≥1) at screening in 66/100 patients, including 48/82 (58.5%) asymptomatic patients and all 18 symptomatic patients. No asymptomatic patients received antibiotics based on sinus CT.

The Lund-Mackay score was higher for patients with symptoms (4.67 ± 0.36) than symptomatic patients without IFRS (3.87 ± 0.73), but this finding did not reach statistical significance (P = .058) (Table 2).

All 5 patients treated with pretransplantation antibiotics had clinical sinusitis. Nasal wash was positive for enterovirus/rhinovirus in 3/5 (60%). Two received antibiotics primarily for pneumonia and secondary coverage of possible bacterial sinusitis. None developed IFRS before or after transplantation.

One patient had a smooth nasal septal perforation, which was negative for IFRS at screening endoscopy. Two others had sinus calcifications. All 3 were asymptomatic, and none developed IFRS or symptoms posttransplantation. No patients (0/100) had sinus hyperdensity, nasal ulceration, extrasinus inflammation, or bone destruction at screening.

No patients underwent a pretransplantation sinus operation, received new antifungal therapy, or had HCT delayed as a result of screening sinus CT.

**Posttransplantation Imaging**

Symptoms prompted posttransplantation evaluation of 21 patients whose Lund-Mackay scores had increased (7.43 ± 6.36) from baseline screening CT (3.87 ± 4.99; P = .007).

Two of the 21 (9.5%) symptomatic patients or 2% (2/100) of the entire cohort developed IFRS, 615 days (Bipolaris spp) and 248 days (Fusarium spp) posttransplantation. Pretransplantation, neither had symptoms. These 2 patients had lower Lund-Mackay scores (4.5 ± 3.54) than symptomatic patients without IFRS (n = 19, 7.74 ± 6.57), but this finding was not statistically significant (P = .59).

**Prediction of Posttransplantation Sinus Disease**

The degree of mucosal thickening (Lund-Mackay score) on pretransplantation CT did not predict posttransplantation symptoms (P = .47) or posttransplantation IFRS (P = .58) (Table 3).

The pretransplantation test for galactomannan was negative in all patients at screening. The mean absolute neutrophil count at screening was 2510 ± 5568 (range, 0–50,200). There was no association between screening absolute neutrophil count and pre-HCT symptoms, post-HCT symptoms, or post-HCT IFRS. Disease status, transplant donor, T-cell depletion, and posttransplantation GVHD grades II–IV did not predict development of IFRS (P = 1 for all; Table 4).

Of 96 participants evaluated by ENT, 29 were considered at risk for IFRS. ENT risk assessment did not predict development of IFRS (P = .61). There was no association between the Lund-Mackay score and the ENT risk stratum (P = .99).

The presence of symptoms before transplantation (n = 18) did not predict posttransplantation symptoms (n = 21); only 3 patients were symptomatic both before and after transplantation (P = .62).

**DISCUSSION**

In this study, we found no association between pre-HCT screening sinus CT and changes in pretransplantation

---

**Table 1: Patient characteristics**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>No.</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>100</td>
<td>100%</td>
</tr>
<tr>
<td>Male</td>
<td>60</td>
<td>62%</td>
</tr>
<tr>
<td>Female</td>
<td>40</td>
<td>38%</td>
</tr>
<tr>
<td>Age (mean) (yr)</td>
<td>11.9 ± 5.5</td>
<td></td>
</tr>
<tr>
<td>Primary diagnosis (No.) (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hematologic malignancy</td>
<td>65, 65%</td>
<td></td>
</tr>
<tr>
<td>Acute myeloid leukemia</td>
<td>34, 34%</td>
<td></td>
</tr>
<tr>
<td>Acute lymphoblastic leukemia</td>
<td>17, 17%</td>
<td></td>
</tr>
<tr>
<td>Hodgkin lymphoma</td>
<td>8, 8%</td>
<td></td>
</tr>
<tr>
<td>Myelodysplastic syndrome</td>
<td>5, 5%</td>
<td></td>
</tr>
<tr>
<td>Anaplastic large-cell lymphoma</td>
<td>1, 1%</td>
<td></td>
</tr>
<tr>
<td>Hematologic disorder</td>
<td>5, 5%</td>
<td></td>
</tr>
<tr>
<td>Severe aplastic anemia</td>
<td>3, 3%</td>
<td></td>
</tr>
<tr>
<td>Sickle cell disease</td>
<td>2, 2%</td>
<td></td>
</tr>
<tr>
<td>Solid tumor</td>
<td>28, 28%</td>
<td></td>
</tr>
<tr>
<td>Neuroblastoma</td>
<td>15, 15%</td>
<td></td>
</tr>
<tr>
<td>Ewing sarcoma</td>
<td>9, 9%</td>
<td></td>
</tr>
<tr>
<td>Germ cell tumor</td>
<td>3, 3%</td>
<td></td>
</tr>
<tr>
<td>Desmoplastic small round cell tumors</td>
<td>1, 1%</td>
<td></td>
</tr>
<tr>
<td>Immune disorder</td>
<td>2, 2%</td>
<td></td>
</tr>
<tr>
<td>Severe combined immunodeficiency</td>
<td>1, 1%</td>
<td></td>
</tr>
<tr>
<td>Wiskott-Aldrich syndrome</td>
<td>1, 1%</td>
<td></td>
</tr>
<tr>
<td>Transplant donor (No.) (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allogeneic</td>
<td>66, 66%</td>
<td></td>
</tr>
<tr>
<td>Autologous</td>
<td>34, 34%</td>
<td></td>
</tr>
<tr>
<td>Disease status at transplantation (No.) (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complete remission</td>
<td>82, 82%</td>
<td></td>
</tr>
<tr>
<td>Active disease</td>
<td>18, 18%</td>
<td></td>
</tr>
<tr>
<td>T-cell depletion (No.) (%)</td>
<td>2, 2%</td>
<td></td>
</tr>
<tr>
<td>GVHD II–IV (No.) (%)</td>
<td>15, 15%</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2: Pretransplantation sinus CT Lund-Mackay scores, sinonasal symptoms, and management**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No. (Frequency)</th>
<th>Pre-HCT Lund-Mackay Score (Mean)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All at screening</td>
<td>100 (100%)</td>
<td>3.48 ± 4.58</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>No symptoms at screening</td>
<td>82 (82%)</td>
<td>2.48 ± 3.51</td>
<td></td>
</tr>
<tr>
<td>Symptoms at screening</td>
<td>18 (18%)</td>
<td>8.07 ± 6.00</td>
<td></td>
</tr>
<tr>
<td>With symptoms, received antibiotics</td>
<td>5 (28%)</td>
<td>12.80 ± 5.72</td>
<td>.058</td>
</tr>
<tr>
<td>With symptoms, no antibiotics</td>
<td>13 (72%)</td>
<td>6.25 ± 5.25</td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Pre-HCT Lund-Mackay scores versus posttransplantation sinus outcomes

<table>
<thead>
<tr>
<th></th>
<th>No. (Frequency)</th>
<th>Pre-HCT Lund-Mackay Score (Mean)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All at screening</td>
<td>100 (100%)</td>
<td>3.48 ± 4.58</td>
<td></td>
</tr>
<tr>
<td>No post-HCT symptoms</td>
<td>79 (79%)</td>
<td>3.38 ± 4.49</td>
<td>.47</td>
</tr>
<tr>
<td>Post-HCT symptoms</td>
<td>21 (21%)</td>
<td>3.86 ± 5.00</td>
<td></td>
</tr>
<tr>
<td>Post-HCT symptoms, IFRS</td>
<td>2 (9.5%)</td>
<td>4.00 ± 4.24</td>
<td>.58</td>
</tr>
<tr>
<td>Post-HCT symptoms, no IFRS</td>
<td>19 (90.5%)</td>
<td>3.84 ± 5.17</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Clinical findings at screening for patients who did and did not develop IFRS after transplantation

<table>
<thead>
<tr>
<th>No. (Frequency)</th>
<th>All at Pre-HCT Screen</th>
<th>No Post-HCT IFRS</th>
<th>Post-HCT IFRS</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENT risk estimate (No.)</td>
<td>100</td>
<td>98</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>At risk</td>
<td>29 (29%)</td>
<td>28 (29%)</td>
<td>1 (50%)</td>
<td>.61</td>
</tr>
<tr>
<td>Low risk</td>
<td>67 (67%)</td>
<td>62 (63%)</td>
<td>1 (50%)</td>
<td></td>
</tr>
<tr>
<td>Not evaluated</td>
<td>4 (4%)</td>
<td>8 (8%)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>ANC at imaging (mean)</td>
<td>2510 ± 5568</td>
<td>2524 ± 5622</td>
<td>1800 ± 1697</td>
<td>.86</td>
</tr>
<tr>
<td>Allogenic</td>
<td>66 (66%)</td>
<td>65 (66%)</td>
<td>1 (50%)</td>
<td>1</td>
</tr>
<tr>
<td>Autologous</td>
<td>34 (34%)</td>
<td>33 (34%)</td>
<td>1 (50%)</td>
<td></td>
</tr>
<tr>
<td>Disease status at transplantation (No.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complete remission</td>
<td>82 (82%)</td>
<td>80 (80%)</td>
<td>2 (100%)</td>
<td>1</td>
</tr>
<tr>
<td>Active disease</td>
<td>18 (18%)</td>
<td>18 (18%)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>T-cell depletion (No.)</td>
<td>2 (2%)</td>
<td>2 (2%)</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>GVHD II–IV (No.)</td>
<td>15 (15%)</td>
<td>15 (15%)</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: ANC indicates absolute neutrophil count.

in 2000, Billings et al found that children with severe sinus disease on pre-bone marrow transplantation CT were more likely to have sinusitis after bone marrow transplantation and reported a trend toward decreased survival in these children, but these findings were not statistically significant ($P = .750$). In a study of 252 adults, Won et al found that asymptomatic radiographic sinus abnormalities did not increase the risk of post-HCT sinusitis, but they also reported a trend toward reduction of post-HCT sinusitis with the treatment of asymptomatic abnormalities, which did not reach statistical significance ($P = .057$). To our knowledge, no investigators have identified pre-transplantation imaging predictors of IFRS.

It has been suggested that CT might be a surrogate for inflammatory symptoms in immunocompromised patients, who may not be able to mount an inflammatory response. However, evidence shows that clinical sinusitis and imaging severity go hand in hand, even in immunocompromised children. This was also true in our study, in which Lund-Mackay scores were greater in the presence of symptoms. The American Academy of Pediatrics and American College of Radiology guidelines dictate that diagnosis and treatment of acute sinusitis should be clinical, with imaging reserved for suspicion of orbital or CNS complications. IFRS typically presents with a combination of fever, facial pain, nasal congestion, headache, eye symptoms, and/or facial swelling, which should also prompt imaging, and likely endoscopy. On pretransplantation physical examination, none of the 100 patients in our series had findings suggestive of orbital, CNS, or fungal involvement warranting imaging evaluation.

On the other hand, isolated radiographic abnormalities, common in asymptomatic children, may be confusing or lead to overtreatment. In our series, 58.8% of 82 asymptomatic children had mucosal thickening on pre-HCT CT. As suggested by other authors, a thorough history and examination are likely to be more meaningful than CT, without the associated risk of radiation-induced cancers and complications. Our findings suggest that there is insufficient benefit to justify the risks of radiation exposure, and potentially of anesthesia, associated with pre-HCT screening CT. The As Low As Reasonably Achievable principle dictates the use of the lowest possible radiation dose necessary for diagnosis. Most pediatric imaging centers, like ours, use targeted protocols with dose-reduction techniques like iterative reconstruction and low tube voltage wherever possible.
However, the best way to reduce the pediatric radiation dose is to eliminate imaging that does not meaningfully contribute to diagnosis or management. The preponderance of the evidence, including ours, suggests that pre-HCT screening sinus CT falls into this category and that the practice should be discontinued. This study had several weaknesses. Clinical evaluation was rarely performed on the same day as screening sinus CT, and symptoms were recorded from the medical record because they were frequently omitted from the imaging requisitions. These features could result in a mismatch between symptoms and imaging manifestations. Although a trend toward greater Lund-Mackay scores in symptomatic patients receiving antibiotics was not statistically significant, the sample size was small. It is possible that the degree of mucosal thickening on sinus CT unduly influenced the decision to treat with antibiotics, which should be based on history and examination. Mucosal thickening on CT does not differentiate between bacterial and viral pathogens, and 3 of 5 patients receiving antibiotics in this study had nasal washes positive for viral pathogens. Next, only patients undergoing imaging work-up were considered symptomatic post-HCT. It is possible that some patients were treated on the basis of clinical symptoms without any imaging, though observation of our current clinical practice suggests that this is unlikely. Similar to other studies, only 2% (2/100) of this study population ultimately developed IFRS, limiting statistical evaluation. However, neither had clinical or imaging findings suspicious for IFRS pretransplantation. Although a single-institution sample size of 100 children undergoing HCT may be considered large, given the rarity of the condition, it was nonetheless insufficient to identify clinical or imaging features associated with IFRS. A retrospective case-control study would be more practical for such analysis.

CONCLUSIONS

Pre-HCT sinus CT does not meaningfully contribute to pretransplantation patient management or prediction of posttransplantation sinus disease, including IFRS, in children. The risks associated with CT radiation and possible anesthesia are not warranted in this setting.

ACKNOWLEDGMENTS

The authors wish to thank Edwina Anderson and Cindy Morris for assistance with data collection.

REFERENCES


The Perirolandic Sign: A Unique Imaging Finding Observed in Association with Polymerase γ-Related Disorders


ABSTRACT

SUMMARY: Pathogenic variants in the polymerase γ gene (POLG) cause a diverse group of pathologies known as POLG-related disorders. In this report, we describe brain MR imaging findings and electroencephalogram correlates of 13 children with POLG-related disorders at diagnosis and follow-up. At diagnosis, all patients had seizures and 12 had abnormal MR imaging findings. The most common imaging findings were unilateral or bilateral perirolandic (54%) and unilateral or bilateral thalamic signal changes (77%). Association of epilepsy partialis continua with perirolandic and thalamic signal changes was present in 86% and 70% of the patients, respectively. The occipital lobe was affected in 2 patients. On follow-up, 92% of the patients had disease progression or fatal outcome. Rapid volume loss was seen in 77% of the patients. The occipital lobe (61%) and thalamus (61%) were the most affected brain regions. Perirolandic signal changes and seizures may represent a brain imaging biomarker of early-onset pediatric POLG-related disorders.

ABBREVIATIONS: ASL = arterial spin-labeling; EEG = electroencephalogram; EPC = epilepsy partialis continua; MELAS = mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes; mtDNA = mitochondrial DNA; Pol γ = polymerase γ; POLG-RD = DNA polymerase γ-related disorder

Polymerase γ (Polγ) is the only DNA polymerase active during human mitochondrial DNA (mtDNA) genome replication.1 Pathogenic variants in the Polγ gene (DNA polymerase γ [POLG]) cause a group of clinical syndromes known as POLG-related disorders (POLG-RDs).1 Patients with POLG-RDs fall into a heterogeneous clinical spectrum. At the least severe end of the spectrum, patients present in adulthood with ptosis and ophthalmoplegia, whereas those most severely affected present with progressive and severe neurologic impairment and liver involvement in early childhood.2

Pathogenic variants of POLG are the most frequently detected genetic forms of mitochondrial epilepsy.3 Seizures are described as the first clinical manifestation in up to 50% of patients. Seizure types include myoclonus, focal motor seizures, generalized seizures, status epilepticus, refractory febrile seizures, and epilepsy partialis continua (EPC).4

Neuroimaging findings in POLG-RDs have been described primarily in the later stages of the disease,5 with the occipital lobe being the most commonly involved region.6 Brain MR imaging findings in the early onset of pediatric POLG-RDs are not well-known. The primary goal of this study was to describe the early brain MR imaging findings in children, including the "perirolandic sign," defined as signal changes in the brain parenchyma surrounding the central sulcus of POLG-RDs, and to correlate them with electroencephalogram (EEG) findings. Our secondary goal was to describe the evolution of brain MR imaging findings on follow-up imaging.

Case Series

This retrospective institutional review board–approved study was conducted in a single academic pediatric hospital, Children’s Hospital of Philadelphia. Medical records were searched for primary mitochondrial disorders from January 2001 to July 2018. Patients with brain MR imaging and a confirmed molecular diagnosis of POLG-RDs (confirmed biparental inheritance for autosomal recessive disease when possible) were included. Exclusion criteria consisted of the unavailability of brain MR imaging or confirmed molecular diagnosis.

Signal changes around the central sulcus were variable with varying degrees of conspicuity. T2WI. Signal changes are subtle and focal, evident only in the left precentral gyrus (open arrows). B, T2WI. Signal changes are subtle and focal, evident in the left pre- and postcentral gyrus (open arrows), but more conspicuous in the DWI (open arrows). C, D, DWI. Linear signal changes involving mainly the cortex surrounding the right central sulcus (open arrows). E and F, DWI and ADC map, respectively. Marked signal changes in both right pre- and postcentral gyri.

Demographic data, clinical history, brain MR imaging at diagnosis, and the EEG most contemporaneous with the MR imaging at diagnosis were reviewed. On the basis of prior observation of an index case (which was included in the analysis), we also explicitly assessed the presence of MR imaging signal changes around the central sulcus involving the pre- or postcentral gyri perirolandic sign. Brain MR imaging was reviewed (at diagnosis and follow-up) by 3 pediatric neuroradiologists in consensus (F.G.G., A.V., and G.Z.). T1 and T2 imaging and DWI were available in at least 1 plane in all examinations. T2* and SWI, MR spectroscopy, and arterial spin-labeling (ASL) were partially available across the studies. When obtained, MR spectroscopy was acquired with the point resolved spectroscopy protocol, using at least a voxel in the right basal ganglia with the following parameters: TR/TE = 1700/20 ms.

From a cohort of 117 patients with primary mitochondrial disorders, 13 met the criteria for a POLG-related disorder. Demographic information, age at onset, brief relevant history at presentation, EEG findings, elapsed time between the first brain MR imaging and the most contemporaneous EEG, and pathogenic variants of each patient’s pathogenic variants are shown in On-line Table 1. The presence of the perirolandic sign involving either the pre- or postcentral gyri, MR imaging signal changes in the thalami, additional brain MR imaging findings, and the specific sequences in which the signal changes were depicted in each patient are shown in On-line Table 2. All patients had follow-up imaging, which varied significantly in number among patients. Thus, instead of being mentioned individually, their follow-up findings were pooled and summarized in 1 single column in On-line Table 2, called “Pooled Follow-Up Imaging.”

Most patients were female (n = 9, 69%). The median age was 3 years (interquartile range, 0.7–7.5 years). All patients presented clinically with seizures (n = 13, 100%). Other common associated symptomatology included regression, developmental delay, hypotonia, vomiting, and signs of liver damage. All patients had an EEG available around the time of their first brain MR imaging (n = 13, 100%); all the scans had abnormal findings. Clinical or EEG evidence of EPC was detected in most patients (n = 8, 61%). The mean elapsed time between the most contemporaneous EEG and brain MR imaging at diagnosis was 6 days. The most common pathogenic variant found from each parent with another variant was c.1399G>A:p.A467T (n = 7, 54%), which is consistent with findings in previous literature.

At diagnosis, brain MR imaging findings were abnormal in most patients (n = 12, 92%). The most common brain MR imaging findings were unilateral or bilateral perirolandic signal changes (n = 7, 54%) (Fig 1) and unilateral or bilateral thalamic signal changes (n = 10, 77%) (Fig 2). Perirolandic signal abnormalities were unilateral in most cases (n = 5, 71%) and were more frequently seen only affecting the precentral or both the pre- and postcentral gyri. Half of the thalamic changes were unilateral (n = 5, 50%). Simultaneous perirolandic and thalamic signal changes occurred in 6 patients (n = 6, 46%). An association of EPC (clinically/EEG) and perirolandic signal changes was present in 6 patients (n = 6, 75%), and an association of EPC (clinically/EEG) and thalamic signal changes, in 7 patients (n = 7, 87.5%). Two patients with EPC (clinically/EEG) did not present with perirolandic signal abnormalities. The findings positive for lesions overall were on the DWI of 10 patients (n = 10, 83%), on the T2WI of 9 patients (n = 9, 75%), and on FLAIR of 7 patients (n = 7, 58%). In 4 patients, DWI was the only sequence with abnormal findings. The occipital lobe was affected in the early brain MR imaging in 2 patients (n = 2, 15%). Signal changes in other brain regions were found in 5 patients (n = 5, 42%), involving multiple regions, namely the cerebral white matter, insula, putamen, caudate nucleus, fornix, cerebellar vermis, and also the frontal and occipital lobes. One patient did not present with either perirolandic or thalamic changes but instead presented with a diffuse pattern similar to that of leukoencephalopathy, with restricted diffusion in the white matter and white matter tracts (Fig 3). Three patients had an abnormally high lactate peak on MR spectroscopy. None of our patients had ASL or other perfusion-weighted imaging at the time.
of diagnosis. T1 and T2* imaging and SWI had negative findings at the time of diagnosis.

All patients had at least 1 follow-up brain MR imaging. Most patients had imaging findings suggesting disease progression (n = 12, 92%) or a fatal outcome. On follow-up imaging, most demonstrated volume loss (n = 10, 77%), which was typically very rapid, in a matter of a few weeks in most patients. On the follow-up brain MR imaging, the perirlandic and thalamic signal changes had various evolutions, with resolution, stability, contralateral involvement (unilateral → bilateral), or progression with time. The most affected brain regions at follow-up were the occipital lobe (n = 8, 61%) and thalamus (n = 8, 61%). The perirlandic sign was detected in 3 patients who did not have any evidence of perirlandic involvement at diagnosis. Other affected brain regions included the hippocampus, brain stem, dentate nucleus, cerebellar vermis, cerebellar hemispheres, other regions of the frontal and parietal lobes, and the occipital (Fig 4) and temporal lobes. ASL perfusion was obtained in only some of the patients at follow-up imaging. Regions more commonly found with increased ASL perfusion included the thalamus, the perirlandic region, and the occipital and parietal lobes. The areas with increased ASL perfusion also had signal changes visible on T2WI, FLAIR, or DWI. An increased lac-

DISCUSSION
Pol γ is the only human DNA polymerase active during mtDNA replication. POLG is a nuclear gene that encodes for Pol γ, and pathogenic variants lead to errors in mtDNA replication, resulting in multiple small mtDNA deletions, an overall reduction in the number of mtDNA genome copies (mtDNA depletion), and decreased adenosine triphosphate production leading to chronic loss of cellular energy.

Patients with POLG-RD have been grouped into multiple distinct clinically-defined syndromes, which were later attributed to different subtypes of pathogenic variants. The prototypical clinical phenotypes associated with POLG pathogenic variants are the following: 1) Alpers-Huttenlocher syndrome; 2) myocerebrohepatopathy spectrum; 3) myoclonic epilepsy myopathy sensory ataxia; 4) ataxia neuropathy spectrum, which includes mitochondrial recessive ataxia syndrome, spinocerebellar ataxia with epilepsy, and sensory ataxia neuropathy dystrophy and ophthalmoplegia; 5) autosomal recessive progressive external ophthalmoplegia; and 6) autosomal dominant progressive external ophthalmoplegia. Nevertheless, it is increasingly recognized that there is a broad clinical spectrum across these clinical syndromes, in which a patient with POLG-RD may present with clinical features that may not fit within a single classically described syndrome.

POLG pathogenic variants are the most common cause of mitochondrial epilepsy. Seizures are reported as the first clinical presentation in 50% of all patients with POLG-RDs. The most critical factor that triggers epileptic activity is the loss of mtDNA (mtDNA depletion). mtDNA depletion causes loss of the respiratory chain components, which restricts normal energy metabolism, causing a continued neurodegenerative process that interferes with neuronal function. This causes a vicious cycle that ultimately leads to neuronal death and parenchymal necrosis.

EPC is a distinct type of focal motor seizure, first described in 1894. EPC characteristically involves repetitive, sometimes rhythmic, unilateral focal motor twitching of the limbs and/or face, with preservation of consciousness, suggesting an underlying brain lesion. EPC presentation is variable, occurring as single or multiple episodes, and it may be chronic, progressive, or nonprogressive. Several neurologic entities are associated with EPC, such as tuberous sclerosis, Sturge-Weber syndrome, and cortical dysplasia. Metabolic abnormalities such as hyponatremia; hypoglycemia; hyperglycemia; hyperuricemia; uremia; brain tumors such as oligodendroglioma, meningioma, and high-grade glioma; autoimmune processes; and infections. Rasmussen encephalitis also has a well-known association with EPC.

A comprehensive literature review of 136 patients with epilepsy with POLG-RDs has shown that stroke-like changes were the most common imaging findings. The lesions were more commonly located in the occipital lobes. Other structures involved were the parietal, temporal, and frontal lobes; thalamus; basal ganglia; and cerebellum. The age of the patients in this study varied from younger than 30 days to 64 years, with a median of 2 years (first quartile = 0.75 and third quartile = 13.50). Before our study, reports of perirlandic MR imaging signal changes in POLG-RDs have only been mentioned in a few isolated case reports, including a 10-month-old child with Alpers-Huttenlocher syndrome who showed multifocal diffusion restriction areas in the left insula, deep gray matter, and...
bilateral perirolandic area, which were initially related to a nonspecific metabolic process. Right pre- and post-central gyri T2 and DWI signal abnormalities were also detected in a 7-month-old girl with POLG-RD who developed focal clonic status epilepticus, mainly of her left arm and occasionally also involving the left leg. Additional case reports described subtle signal changes on FLAIR in a 15-year-old girl with EPC and a 13-month-old child with bilateral perirolandic restricted diffusion with a history of hypotonia, mild motor delay, and EPC.

In this current study of children with POLG-RDs, all patients presented initially with seizures. Most initial brain MR imaging findings were abnormal. A temporal association between seizures and brain MR imaging abnormalities was found, suggesting lesion-related seizures. The primary brain MR imaging abnormalities included unilateral or bilateral signal changes in the perirolandic region, unilateral or bilateral signal changes in the thalamus, or a combination of the 2 in most of our patients. At diagnosis, the perirolandic region or the thalamus was spared in only 2 patients in our cohort, one with unremarkable brain MR imaging findings and the other in whom diffuse white matter with restricted diffusion was noted (Fig 3).

The perirolandic sign was common and present in most patients at diagnosis. Moreover, the sign was seen in most patients during the course of the disease because half of the patients who did not have the sign initially presented with the sign later. The presence of perirolandic signal changes would reflect neuronal death following acute energy failure in a metabolically demanding region (primary motor and sensory cortex) within the dominant hemisphere. Further studies are needed to confirm this hypothesis.

The MR imaging appearance of the perirolandic sign was varied. Signal changes involved both the pre- and postcentral gyri, more commonly in the precentral gyri. The conspicuity of the perirolandic sign was also variable (Fig 1). In the more notable cases, signal changes were ribbon-like following the course of the gyri, which were readily detectable as T2 hyperintensities and restricted diffusion. DWI was the most sensitive MR imaging sequence to detect signal changes and, therefore, should always be included in the protocol and carefully evaluated when a case of POLG-RD is suspected.

Thalamic signal changes were also frequent at the time of diagnosis and on follow-up imaging. Unilateral or bilateral thalamic involvement was identified in most patients during the onset of their disease. In our cohort, there was no new thalamic involvement at follow-up. On follow-up imaging, thalamic changes had different outcomes: complete resolution, progression from unilateral to bilateral involvement, progression accompanied by volume loss, and fluctuation with periods of an almost-complete resolution and frank progression. Thalamic signal changes were variably detected on DWI, T2WI, FLAIR, or ASL.

Volume loss was detected almost always on follow-up imaging. Volume loss varied from mild to severe, showing rapid evolution in most cases. Most important, volume loss showed no lobar predominance in most patients. However, in severe cases, the occipital lobes were affected asymmetrically and with encephalomalacic changes. Signal changes were detected virtually in any part of the brain, except the medulla. The spinal cord was not part of the scope of this study.

The occipital lobe was rarely affected at diagnosis, though frequently involved on follow-up imaging. Advanced MR imaging sequences such as MR spectroscopy and ASL, albeit not

FIG 3. ADC map images of a 7-month-year old male with POLG-related disorder, demonstrating a diffuse pattern of leukoencephalopathy with restricted diffusion of the periventricular white matter of the bilateral temporal lobes in A, the occipital lobes in B, and also of the bilateral fornix in C and corpus callosum in D.

FIG 4. Bilateral occipital volume loss, gliosis, and encephalomalacia in a 16-year-old female patient with POLG-related disorder (open arrows).
being acquired in all the patients, were useful to demonstrate increased lactate peaks and areas of localized hyperperfusion, respectively. An increased lactate peak is an expected finding in primary mitochondrial disorders and has already been described in patients with POLG-RDs, reflecting mitochondrial dysfunction.9 ASL perfusion has mostly been described in patients with mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS), but rarely in POLG-RDs. Increased ASL perfusion in MELAS preceded signal changes on conventional MR imaging in 3 patients, indicating that ASL imaging has the potential for predicting the emergence of stroke-like lesions.17 In our cohort, increased ASL perfusion was coincident with either T2WI, FLAIR, or DWI signal changes and may have been related to seizure activity. ASL, however, was useful to indicate areas that would undergo further volume loss in some of our patients (Fig 5). The complete role of ASL perfusion in patients with POLG-RDs needs further investigation. 

Valproic acid is an effective antiepileptic drug to manage epilepsy. However, it is strictly contraindicated to treat seizures in patients with POLG-RD due to its potential to cause fulminant hepatotoxicity.5 The best diagnostic test to prevent valproic acid-induced liver toxicity is POLG gene testing before the use of this medication.18 The presence of a unilateral or bilateral perihemispheric, unilateral or bilateral thalamic MR imaging signal changes, or a combination of both in a child with seizures/EPC may be a POLG-RD biomarker, which can also be useful for early suspicion of this diagnosis and to prevent the use of valproic acid in these children. 

MR imaging findings in our cohort are not pathognomonic for POLG-RD, and the differential diagnosis includes postictal changes and stroke-like episodes similar to MELAS. At this time, we can only speculate about their nature, and further laboratory, genetic, and imaging correlation are necessary to establish their pathogenesis. Bilateral perihemispheric cortical involvement can also be seen in patients with hypoxic-ischemic injury,19 urea-cycle disorders,20 and Wernicke encephalopathy,21 though these patients often have a different clinical context. Perihemispheric white matter involvement can be seen in cases of hypoxic-ischemic injury19 and maple syrup urine disease.22 Unilateral cortical perihemispheric involvement can be seen in vascular infarcts, focal cortical dysplasia, and primary and metastatic tumors.23 Unilateral or bilateral thalamic signal changes can be seen in multiple entities such as arterial infarction, hypoxic-ischemic injuries, several metabolic diseases, infection, inflammatory processes, neoplasms, and veno-occlusive syndromes.24,25

Limitations of our study include the small number of patients with this rare disorder and the fact that this was a retrospective analysis. However, this study is about the first and largest cohort, describing the very early brain MR imaging findings in pediatric patients with confirmed POLG-RD. EPC, in association with perihemispheric signal abnormalities, may serve as an early sign and a potential biomarker of POLG-RD. This clinical presentation should increase awareness and prompt testing for POLG pathogenic variants because the discovery of POLG-RD has relevant implications for acute clinical management and long-term prognosis.

Disclosures: Zarazuela Zollipip-Cunningham—UNRELATED: Employment: Children’s Hospital of Philadelphia; Miami J. Falk—UNRELATED: Board Membership: United Mitochondrial Disease Foundation, Comments: travel reimbursement for serving on scientific and medical advisory boards and board of trustees of the United Mitochondrial Disease Foundation; Consultancy: several pharma companies, Comments: paid consultant for Mitobridge/Astellas, NeuroVive, Cycleronic Therapeutics, Reneo Pharmaceuticals, Mission Therapeutics, Stealth BioTherapeutics, and Inel Therapeutics; Grants/Grants Pending: several pharma companies and foundations, Comments: grants from Mitobridge/Astellas, NeuroVive, Cycleronic Therapeutics, Mission Therapeutics (pending), Stealth BioTherapeutics, Minovia Therapeutics (pending); Stock Options: MitoCureia, Comments: co-founder and scientific advisory board co-Chair of MitoCureia; Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed: education travel, Comments: reimbursed for travel to speak at meetings on mitochondrial diagnostics and therapeutics including at CardioInfantil, Bogota, Columbia; Pediatric Neurology Society, Mexico; American College of Medical Genetics; Emory/Atlanta Children’s Hospital; New York University; Cornell University; Washington University in St Louis; Seattle Children’s Hospital; scientific

FIG 5. Increased ASL perfusion in 3 different patients with POLG-related disorders (A and B), (C and D), and (E and F). Increased perfusion regions are seen in the right perihemispheric region in A (open arrow) and in the left occipital lobe in B (open arrow). Bilateral parietal and occipital increased perfusion is seen in C (open arrows) and in the bilateral cerebellar hemispheres (open arrows) in D. Increased ASL perfusion is seen in the left thalamus in E (solid arrow) and in the left parietal and occipital lobes in E and F (open arrows).
and family meetings hosted by United Mitochondrial Disease Foundation; European Society of Human Genetics; Society for the Study of Inborn Errors of Metabolism; Society of Inherited Metabolic Disease; Hebrew University; Asian Society Of Inborn Errors of Metabolism; Wellcome Genome Campus; Cold Spring Harbor Laboratory. Amy Goldstein—UNRELATED: Board Membership: United Mitochondrial Disease Foundation, Comments: Scientific and Medical Advisory Board. *Money paid to the institution.

REFERENCES

6. Nurminen A, Farnum GA, Kaguni LS. Pathogenicity in POLG syn-


dromes: DNA polymerase gamma pathogenicity prediction server and database. BBA Clin 2017;7:147–56 CrossRef Medline
13. Park S, Kang HC, Lee JS, et al. Alpers-Huttenlocher syndrome first presented with hepatic failure: can liver transplantation be con-


considered as treatment option? Pediatr Gastroenterol Hepatol Nutr 2017;20:259–62 CrossRef Medline


related MRI findings of non-perinatal hypoxic-ischaemic ence-


ABSTRACT

BACKGROUND AND PURPOSE: Focal signal abnormalities at the depth of the cerebellar fissures in children have recently been reported to represent a novel pattern of bottom-of-fissure dysplasia. We describe a series of patients with a similar distribution and appearance of cerebellar signal abnormality attributable to watershed injury.

MATERIALS AND METHODS: Twenty-three children with MR imaging findings of focal T2 prolongation in the cerebellar gray matter and immediate subjacent white matter at the depth of the fissures were included. MR imaging examinations were qualitatively analyzed for the characteristics and distribution of signal abnormality within posterior fossa structures, the presence and distribution of volume loss, the presence of abnormal contrast enhancement, and the presence and pattern of supratentorial injury.

RESULTS: T2 prolongation was observed at the depths of the cerebellar fissures bilaterally in all 23 patients, centered at the expected location of the deep cerebellar vascular borderzone. Diffusion restriction was associated with MR imaging performed during acute injury in 13/16 patients. Five of 23 patients had prior imaging, all demonstrating a normal cerebellum. The etiology of injury was hypoxic-ischemic injury in 17/23 patients, posterior reversible encephalopathy syndrome in 3/23 patients, and indeterminate in 3/23 patients. Twenty of 23 patients demonstrated an associated classic parasagittal watershed pattern of supratentorial cortical injury. Injury in the chronic phase was associated with relatively preserved gray matter volume in 8/15 patients, closely matching the published appearance of bottom-of-fissure dysplasia.

CONCLUSIONS: In a series of patients with findings similar in appearance to the recently described bottom-of-fissure dysplasia, we have demonstrated a stereotyped pattern of injury attributable to cerebellar watershed injury.

ABBREVIATIONS: HII = hypoxic-ischemic injury; PRES = posterior reversible encephalopathy syndrome; SCA = superior cerebellar artery

Focal signal abnormalities at the depth of the cerebellar fissures in children have recently been reported and have been hypothesized to represent a novel pattern of bottom-of-fissure dysplasia.1 On the basis of our observation of a similar imaging pattern in patients—including some with initially normal scan findings—who displayed a depth of cerebellar fissure abnormality attributable to ischemic injury, we sought to review our institutional experience. We report a series of 23 patients with a similar distribution and appearance of cerebellar signal abnormality attributable to watershed injury.

MATERIALS AND METHODS

The study was approved by our institutional review board, and on the basis of minimal risk, the need for written informed consent was waived. Twenty-three patients with MR imaging findings of focal T2 prolongation in the cerebellar gray matter at the depth of the fissures were identified by querying our local radiology data base for the past 15 years and are included in this review. Retrospective analysis of the MR imaging studies was performed by 2 subspecialty board-certified pediatric neuroradiologists with 8 (J.N.W.) and 7 (F.P.) years of experience. Imaging datasets were qualitatively analyzed for the characteristics and distribution of signal abnormality within posterior fossa structures, the presence and distribution of volume loss, the presence of abnormal contrast enhancement, and the presence and pattern of supratentorial injury. Patient demographics, clinical histories, and ultimate clinical-diagnostic diagnoses were determined by chart review.

RESULTS

Patients

Twenty-three patients with characteristic abnormal cerebellar signal were included in our retrospective analysis, of which 16 were males and 7 females (1 proportion Z-test; $z = 1.877; P = .06$). Age
The first MR imaging examination demonstrating cerebellar signal abnormality ranged from 2 months to 18 years, with a mean of 4.4 years and a median of 1.2 years. The mean age for patients with acute-phase imaging (n = 16) was 3.2 years; and for acute-phase imaging excluding posterior reversible encephalopathy syndrome (PRES) (n = 13), it was 0.7 years. Final clinicoradiologic determination of the etiology of brain injury was hypoxic-ischemic injury (HII) in 17 patients, PRES in 3 patients, and indeterminate in 3 patients. Among patients with HII, 5 cases were secondary to abusive head trauma; 4, to cardiopulmonary arrest; 3, to perinatal birth injury; 2, to profound septic shock; 2, to perioperative hypotension; and 1, to accidental asphyxiation. Ten patients had follow-up imaging at least 1 year after the initial MR imaging demonstrating cerebellar abnormality, with a median follow-up of 64 months (range, 15–142 months). Five patients had past MR imaging studies available, obtained before the first MR imaging that demonstrated the cerebellar abnormality.

**Infratentorial Imaging Findings**

The key imaging feature was T2 and FLAIR hyperintensity and T1 hypointensity confined to the gray matter and immediate subjacent white matter at the depths of the cerebellar fissures (Fig 1). The findings were best seen on coronal and parasagittal fluid-sensitive sequences. In all patients, bilateral hemispheric involvement was present, with injury typically involving multiple fissures centered at the expected location of the borderzone between the superior cerebellar artery (SCA), AICA, and PICA (Fig 2). The cerebellar tonsils were injured in 39% (9/23) of cases. The cerebellar vermis was spared in 91% (21/23) of patients. Five of the 23 patients had prior imaging studies, all of which demonstrated a normal cerebellar cortex (Fig 3).

Imaging was available for 16/23 patients during the acute phase of injury, defined as <7 days from the suspected clinical event. All patients demonstrated normal thickness or a mildly thickened cerebellar cortex (Figs 3 and 4). Eighty-one percent (13/16) of patients demonstrated acute diffusion restriction (12 with an eventual clinical diagnosis of HII, 1 of unknown etiology) (Fig 4A–C), while 19% (3/16) of patients demonstrated no diffusion abnormality (all with an eventual clinical diagnosis of PRES) (Fig 4D–F). The 3 patients with PRES were imaged in the acute phase only, with no imaging follow-up available.

Imaging was available for 7/23 patients in the subacute phase of injury, defined as 7–30 days from the suspected clinical event. Three of 7 patients (patients 4, 18, and 22) demonstrated intrinsic T1 signal hyperintensity at the depths of the cerebellar fissures (Fig 5A). In 1 case, this was demonstrated to have resolved on a follow-up study performed 1 month later (patient 22) (Fig 5B). One patient who underwent contrast-enhanced MR imaging in the subacute phase demonstrated focal contrast enhancement in the areas of cerebellar injury (patient 18).

Imaging was available for 15/23 patients in the chronic phase of injury, defined as >30 days from the suspected clinical event. All had variable thicknesses of T2 prolongation without diffusion restriction in the cerebellar cortex and the immediate subcortical...
white matter (Fig 6). Seven of 15 patients had marked cortical volume loss with minimal residual T2 hyperintensity (Fig 6A, B), while the remaining 8 demonstrated a relatively thick layer of T2 hypersignal involving the foliar cortex without substantial volume loss (Fig 6C, D). Subtle fissural prominence and mild foliar volume loss were identified in 14/15 patients. Ten patients had follow-up for >1 year after the initial MR imaging demonstrating cerebellar signal abnormality, with a median follow-up interval of 5.3 years (range, 1.3–11.8 years). Cerebellar signal abnormality persisted in all cases and for up to 11.8 years in the case with the longest follow-up (patient 1). There were no discernable group differences in terms of the presence of supratentorial injury, overall injury extent or severity, the presence of acute-phase diffusion restriction, underlying etiology, or mean length of follow-up between patients with thick or thin T2 foliar hyperintensity in the chronic phase.

**Supratentorial Imaging Findings**

In the supratentorial brain, 87% (20/23) of patients demonstrated a classic parasagittal watershed pattern of cortical injury (Fig 7), with 11 with additional evidence of associated basal ganglia injury. Seventeen patients had watershed infarctions and a concordant clinical history of HII. The remaining 3 patients demonstrated a watershed pattern of injury typical of PRES. In all cases, the apparent timing of the supratentorial injury was consistent with the estimated timing relative to the suspected clinical event.

Three of 23 patients demonstrated no evidence of supratentorial watershed cortical injury. One (patient 11) demonstrated diffuse progressive cortical atrophy and a deep pattern of watershed injury in the supratentorial white matter, in the context of underlying drug-resistant chronic epileptic encephalopathy. The remaining 2 patients had no demonstrable supratentorial cortical injury. The first of these had a history of remote aortic stenosis and coarctation repair, and the second had end-stage renal disease.
complicated by chronic hypotension and respiratory failure, both of which may have predisposed to recurrent hypoperfusion events that were clinically unsuspected. Of the 3 patients without supratentorial cortical watershed injury, 2 had prior imaging demonstrating normal cerebella (patients 10 and 11).

DISCUSSION
We hypothesize that the pathologic basis for the stereotyped depth-of-cerebellar-fissure MR imaging findings in our case series is cerebellar watershed injury for the following reasons: 1) location at the deep borderzone between the cerebellar artery vascular territories, 2) association with a classic watershed pattern of supratentorial cortical injury, 3) frequent acute-phase diffusion restriction with expected evolution across time, and 4) frequent late foliar volume loss with fissural prominence. Furthermore, among patients with MR imaging performed before the suspected clinical insult, the cerebellar abnormalities were not previously present, indicating an acquired insult rather than a congenital malformation. Diaschisis secondary to the supratentorial injury was considered an unlikely contributor to the observed infratentorial injury pattern, due to the simultaneous evolution of brain injury in both locations and the multifocal rather than diffuse nature of the cerebellar findings.

Two distinct etiologies of watershed injury were observed in our case series, including PRES and watershed infarctions in the setting of HII. The cortical injury in PRES is poorly characterized at the cellular level but is presumed secondary to abnormal perfusion in the setting of disturbed autoregulation. As is typical of this entity, acute-phase imaging demonstrated T2 prolongation associated with facilitated diffusion. PRES has a strong and well-described predilection for vascular borderzones in the supratentorial brain. While cerebellar involvement has been reported, its distribution has not been well-characterized.

A classic parasagittal distribution of supratentorial watershed infarctions is well-recognized, with preferential involvement of the end-arterial borderzones between the anterior, middle, and posterior cerebral artery vascular territories. Cerebellar watershed infarcts remain less well-described, with scattered case reports and case series describing predominantly embolic watershed infarcts in adults. A few authors have emphasized the distinction between superficial cerebellar borderzone embolic infarcts and deep cerebellar watershed infarcts, the latter related to hypoperfusion and centered at the interface between the penetrating arteries of the cerebellar arteries.

Although some authors have asserted that deep watershed infarcts predominantly affect the deep cerebellar white matter, several lines of evidence suggest that focal injury to the gray matter at the depth of the cerebellar fissures is, in fact, the expected pattern. In a primate model of induced cerebral hypoperfusion, cerebellar injury in the watershed between the SCA and PICA was noted in 7/15 adult rhesus monkeys, in conjunction with classic supratentorial watershed injury. The injury predominantly affected cortical gray matter at the depths of the cerebellar fissures. Depth of fissure gray matter injury was also noted on gross and histopathologic evaluations of human brains after hypoxic injury, in a pattern exactly corresponding to the distribution of injury in our case series. This distribution may reflect the unique angioarchitecture of the penetrating arteries in the cerebellum, analogous to the propensity for depth-of-sulcus cortical injury in neonates with HII reported by Takashima et al.

Sparing of the cerebellar vermis was a notable feature in our patients, possibly because the arterial supply to the vermis is almost entirely derived from the SCA, with relatively minor contributions from the inferior cerebellar arteries. In contrast, cerebellar tonsillar injury was observed in more than one-third of patients in our series, likely reflecting its location at the borderzone between the AICA and PICA territories.
Our patient series comprised a disproportionate number of male patients, though this finding did not reach statistical significance regarding variance from an expected population proportion of 50% male. This may, in part, be related to the number of cases of HII secondary to abusive head trauma in our series, which is known to disproportionately affect male patients. Haas-Lude et al. recently reported a strikingly similar pattern of cerebellar injury in 2 patients with abusive head trauma, both with a supratentorial watershed pattern of HII. It would be interesting to investigate the prevalence of this finding in a larger cohort of patients with suspected abusive head trauma to better understand its predictive value above and beyond the presence of HII more generally.

Our case series is among the first to describe the characteristic appearance of cerebellar watershed injury in children. Recognition of this pattern of injury is important for identifying patients with HII or PRES and for avoiding misdiagnosis. For example, 1 of our patients was originally misdiagnosed with cerebellitis based on abnormal cerebellar T2 hyperintensity (patient 10), delaying definitive diagnosis. Similarly, Mills et al. reported a case of "severe cerebellitis" in a child with supratentorial watershed infarcts following methadone poisoning. The imaging findings in that case are strikingly similar to those in our case series, raising the possibility of unrecognized watershed injury rather than cerebellitis.

A second patient in our case series was initially diagnosed with bottom-of-fissure dysplasia; however, a prior MR imaging subsequently became available demonstrating a normal cerebellum, indicating an acquired rather-than-congenital lesion. The similarity between our reported imaging findings and those of Poretti et al suggests that unrecognized watershed injury may have been present in some or all of their cases, or that cerebellar watershed injury and bottom-of-fissure dysplasia share a common imaging appearance. Although they considered the possibility of cerebellar watershed injury as the etiology of the imaging findings in Poretti et al, they ultimately concluded that lack of substantial white matter involvement, variability of distribution of involved fissures, and stability with time all argued against this possibility. It is clear from our patient series—corroborated by the available experimental neurophysiology and neuropathology literature—that the foliar gray matter at the depths of the fissures is an expected location of cerebellar watershed injury, with minimal or absent white matter involvement.

Variable fissural involvement is expected because it reflects the relatively variable vascular territories and shifting borderzones of the cerebellar arteries. Furthermore, stability with time in the chronic phase, often with unexpectedly subtle volume loss, was the norm in our series.

There are several limitations to our analysis, primarily reflecting the small sample size and retrospective design. We were unable to assess the clinical implications of cerebellar watershed injury in our case series because of the small and heterogeneous sample. Our study does not permit an analysis of the prevalence of cerebellar imaging findings among children with supratentorial watershed injury. It is also possible that a subset of our cases may represent a dysplastic process as suggested by Poretti et al, particularly those without corroborative supratentorial findings. However, 5 of our patients had a prior MR imaging study demonstrating a normal cerebellum, including 2 of the 3 patients without clear supratentorial cortical watershed injury, arguing against the likelihood of a developmental abnormality.

CONCLUSIONS
We have demonstrated a stereotyped pattern of injury that is most plausibly attributable to cerebellar watershed injury. This report extends the reported imaging findings in watershed cortical injury related to HII and PRES, which are an important diagnostic consideration when cerebellar dysplasia and inflammatory pathologies such as cerebellitis are being entertained.

REFERENCES


Sensitivity of the Inhomogeneous Magnetization Transfer Imaging Technique to Spinal Cord Damage in Multiple Sclerosis

H. Rasoanandrianina, S. Demortière, A. Trabelsi, J.P. Ranjeva, O. Girard, G. Duhamel, M. Guye, J. Pelletier, B. Audoin, and V. Callot

ABSTRACT

BACKGROUND AND PURPOSE: The inhomogeneous magnetization transfer technique has demonstrated high specificity for myelin, and has shown sensitivity to multiple sclerosis-related impairment in brain tissue. Our aim was to investigate its sensitivity to spinal cord impairment in MS relative to more established MR imaging techniques (volumetry, magnetization transfer, DTI).

MATERIALS AND METHODS: Anatomic images covering the cervical spinal cord from the C1 to C6 levels and DTI, magnetization transfer/inhomogeneous magnetization transfer images at the C2/C5 levels were acquired in 19 patients with MS and 19 paired healthy controls. Anatomic images were segmented in spinal cord GM and WM, both manually and using the AMU40 atlases. MS lesions were manually delineated. MR metrics were analyzed within normal-appearing and lesion regions in anterolateral and posterolateral WM and compared using Wilcoxon rank tests and z scores. Correlations between MR metrics and clinical scores in patients with MS were evaluated using the Spearman rank correlation.

RESULTS: AMU40-based C1-to-C6 GM/WM automatic segmentations in patients with MS were evaluated relative to manual delineation. Mean Dice coefficients were 0.75/0.89, respectively. All MR metrics (WM/GM cross-sectional areas, normal-appearing and lesion diffusivities, and magnetization transfer/inhomogeneous magnetization transfer ratio) were observed altered in patients compared with controls (P < .05). Additionally, the absolute inhomogeneous magnetization transfer ratio z scores were significantly higher than those of the other MR metric z scores, suggesting a higher inhomogeneous magnetization transfer sensitivity toward spinal cord impairment in MS. Significant correlations with the Expanded Disability Status Scale (ρ = −0.73/P = .02, ρ = −0.81/P = .004) and the total Medical Research Council scale (ρ = 0.80/P = .009, ρ = −0.74/P = .02) were observed for inhomogeneous magnetization transfer and magnetization transfer ratio z scores, respectively, in normal-appearing WM regions, while weaker and nonsignificant correlations were obtained for DTI metrics.

CONCLUSIONS: With inhomogeneous magnetization transfer being highly sensitive to spinal cord damage in MS compared with conventional magnetization transfer and DTI, it could generate great clinical interest for longitudinal follow-up and potential remyelinating clinical trials. In line with other advanced myelin techniques with which it could be compared, it opens perspectives for multicentric investigations.

ABBREVIATIONS: AMU = Aix-Marseille University; CSA = cross-sectional areas; EDSS = Expanded Disability Status Scale; FA = fractional anisotropy; HC = healthy controls; RMT = inhomogeneous magnetization transfer; iMTR = inhomogeneous magnetization transfer ratio; Axy = axial diffusivity; Arad = radial diffusivity; MRC = Medical Research Council; MT = magnetization transfer; MTR = magnetization transfer ratio; NA = normal-appearing; SC = spinal cord; TWT = Timed 25-Foot Walk Test; SCT = Spinal Cord Toolbox

S is a chronic autoimmune disease of the CNS affecting the brain and the spinal cord (SC), characterized by multifocal demyelination and inflammation, edema, axonal loss, and gliosis, whose diagnosis is based on anatomic MR imaging. Although this work was performed by a laboratory member of France Life Imaging network (grant ANR-11-INBS-0006) and supported by the French Centre National de la Recherche Scientifique, the French Ministère de l’Enseignement Supérieur et de la Recherche, and the Fondation pour l’Aide à la Recherche sur la Sclérose En Plaques.

Please address correspondence to Virginie Callot, PhD, CRMBM-CEMEREN, UMR 7339, CNRS, Aix-Marseille Université, Faculté de Médecine, 27, bd Jean Moulin, 13385 Marseille Cedex 5, France; e-mail: virginie.callot@univ-amu.fr

Indicates article with supplemental on-line tables.

Indicates article with supplemental on-line photo.

http://dx.doi.org/10.3174/ajnr.A6554


Original Research

Spine
extensive investigations including demonstration of multicentric capabilities and scan time reduction are still necessary to bridge the gap to clinical practice, quantitative MR imaging techniques now have the potential to help improve disease diagnosis and monitoring, as well as treatment evaluation.

Given the clinical relevance and the risk of long-term disability in spinal MS, a better characterization of pathologic mechanisms underlying SC tissue impairment, including myelin damage, is therefore essential.

Unfortunately, SC MR imaging lags behind brain MR imaging, mainly due to technical challenges related to the small size, CSF pulsatility, respiratory motion, and susceptibility artifacts of the spinal cord due to surrounding bony structures. Nonetheless, due to recent methodologic improvements, a growing number of SC MR imaging studies in MS have been reported, mostly using DTI and conventional magnetization transfer (MT) imaging, highlighting SC microstructural changes with significant correlations to physical disability.

Although highly sensitive to MS-related tissue changes, both DTI and MT metrics are not specific to myelin damage because they could also be affected by inflammation and/or axonal loss. Tremendous effort has thus recently been made toward the development of more myelin-specific MR imaging techniques. These particularly include myelin-water imaging, which gives access to the myelin water fraction or water protons associated with myelin density, or the recently described inhomogeneous MT (ihMT) technique. This latter allows selectively isolating the myelin sheaths, as well as quantitative MT, which gives a robust estimate of the macromolecular-to-free pool size ratio, which has been shown to strongly correlate with white matter myelin density, or the recently described inhomogeneous MT (ihMT) technique. This latter allows selectively isolating the contribution of dipolar order within the broad MT signal and hence appears particularly sensitive to myelinated structures due to their relatively long dipolar relaxation times. Applied in a preliminary clinical MS study, ihMT was found to be highly sensitive to cerebral microstructural changes with strong correlation to the Expanded Disability Status Scale (EDSS). Furthermore, embedded within a protocol optimized for the SC, it has also already demonstrated high sensitivity to SC tissue changes in both aging healthy subjects and patients with amyotrophic lateral sclerosis.

In line with these studies, the main purpose of this work was to assess the sensitivity of the ihMT technique to SC impairment in MS relative to more conventional MR imaging markers, including atrophy, MT, and DTI measures and to demonstrate its potential in clinical monitoring. To do so, we applied a multi-parametric MR imaging protocol, including high-resolution anatomic and quantitative DTI and MT/ihMT imaging techniques, together with a SC template–based approach to quantify data collected in patients with MS and age-matched healthy controls (HC) in different ROIs.

**MATERIALS AND METHODS**

**Subjects**

This retrospective study was approved by our institutional ethics committee (CCP Sud Méditerranée I), and written informed consent was obtained from all subjects. Nineteen patients with MS (13 relapsing-remitting and 6 secondary-progressive forms) and 19 sex- and age-matched HC were included during 2 years. Inclusion criteria for patients were a diagnosis of MS according to the revised McDonald criteria and clinical symptoms suggesting SC involvement. Clinical symptoms were evaluated >3 months after steroid infusion. Exclusion criteria were any other neurologic disease other than MS and chronic psychiatric diseases. Neurologic examinations were performed right after the MR imaging acquisitions. They included a specific MS global disability rating scale with EDSS scoring. A muscle strength grading with the Medical Research Council scale (MRC) for the 4 limbs was also performed. Specific assessment of the upper limbs was performed with a handgrip strength test and the Nine-Hole Peg Test, currently used in MS to assess manual dexterity. For lower limb evaluation, a Timed 25-Foot Walk Test (TWT) was performed. Additional details are found in Table 1.

**Image Acquisition**

MR imaging was performed on a 3T scanner (Magnetom Verio; Siemens) using the body coil for transmission and standard 12-channel head, 4-channel neck, and 24-channel spine matrix array coils for reception. The subjects were positioned with the neck in a neutral position using a cervical collar to reduce subject motion.

The MR imaging protocol (summarized in On-line Table 1) was optimized to last approximately 50 minutes. It included a 3D-T2WI sequence for vertebral level identification and cervical lesion depiction (see postprocessing in Fig 1), a 2D multislice multangle T2WI multiecho gradient-echo sequence (7 slices, 1 per cervical level from C1 to C7), providing a good contrast to delineate GM and WM, as well as 2-level DTI and ihMT acquisitions (single-section) centered at mid-C2 and mid-C5 cervical levels, placed strictly perpendicular to the cord axis. Patients also underwent, during the same session, conventional brain and SC acquisitions at the cervical level (T1WI, T2WI, FLAIR, STIR, and postgadolinium T1) for lesion burden evaluation.

| Table 1: Demographic and clinical features of the study population including HC and patients with MS* |
|---|---|---|
| Demographics | HC | MS |
| Count (female/male) | 19 (11/8) | 19 (11/8) |
| Age (years) | 52 (22–74) | 46 (26–70) |
| Disease duration (months) | – | 117 (48–372) |
| EDSS | – | 2.5 (1–7) |
| Total MRC (1/120) | 118 (103–120) | |
| NHPT (sec) | 24.1 (17.8–73.4) | |
| TWT (sec) | 5.5 (3.4–73) | |
| Handgrip (kPa) | 82.5 (23.5–122.5) | |
| No. of SC lesions | 4 (2–7) | |
| Volume of brain lesions (mm3) | – | 3261 (726–38,013) |

Note: NHPT indicates Nine-Hole Peg Test; –, not applicable measurements.

a Reported values are medians (minimum–maximum). Lesion burden in the cervical SC and brain of patients with MS is also reported.
b Medical Research Council scale normal value per muscle; 5, range, 0–5, for upper (7 muscles on each side) and lower limbs (5 muscles on each side), with a total maximum score of 120.

*The Nine-Hole Peg Test for each hand and TWIs were performed twice for each patient, and the best performance was preserved.

†Handgrip measurements for each patient were repeated thrice for each hand and then averaged to get 1 value per limb.
Data Analysis

The entire postprocessing pipeline, detailed in Fig 1, was performed off-line using a custom Matlab script (Matlab R2014a; MathWorks), relying on the Spinal Cord Toolbox (SCT; https://sourceforge.net/projects/spinalcordtoolbox/)27 functions and dedicated SC templates and atlases (MNI-Poly-AMU; https://sourceforge.net/p/spinalcordtoolbox/wiki/MNI-Poly-AMU/28 and AMU4029), allowing automatic delineation of the GM and WM.

To evaluate the efficacy of the AMU40-based automated GM/WM segmentation applied for the first time in MS, 2 independent experts (S.D., an in-training neurologist, and H.R., a 3-year MR imaging expert) manually outlined the SC, GM/WM, and lesional/NA tissue segmentations on T2*WI; and T2WI by the SCT PropSeg function; T2WI registration to the MNI-Poly-AMU template28 to obtain cervical levels; T2*WI registration to the AMU40 template29 to obtain GM/WM segmentations, warped back into the subject’s space followed by CSA (square millimeters) extraction and intrasubject multiparametric data normalization, taking into account the presence of lesions, allowing metrics extraction from lesioned and NA anterolateral WM and posterolateral WM, derived from the WM tracts atlas and defined as delimited by the blue horizontal line. Masks were eroded (2D surface) before quantification, y.o. indicates years of age; Gd, gadolinium; anterolat, anterolateral; posterolat, posterolateral; FSL view refers to the FSL viewer software; ANTs, Advanced Normalization Tools (https://stnava.github.io/ANTs/); FLIRT, FMRIB’s Linear Image Registration Tool (http://www.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT); WM tracts atlas (https://sourceforge.net/p/spinalcordtoolbox/wiki/White%20Matter%20atlas/).32

FIG 1. Design of the study. Acquisition and postprocessing pipeline for MR imaging data shown for the C2 cervical level in a patient with MS, including the following; manual SC, GM/WM, and lesional/NA tissue segmentations on T2*WI and registration of the 3D-T2WI to the MNI-Poly-AMU template to automatically identify the cervical levels from C1 to C7 (warped back into the subject’s space; this step allows cord volume per level measurements, not reported in this study); SC segmentation on T2WI and T2WI by the SCT PropSeg function; T2WI registration to the MNI-Poly-AMU template28 to obtain cervical levels; T2*WI registration to the AMU40 template29 to obtain GM/WM segmentations, warped back into the subject’s space followed by CSA (square millimeters) extraction and intrasubject multiparametric data normalization, taking into account the presence of lesions, allowing metrics extraction from lesioned and NA anterolateral WM and posterolateral WM, derived from the WM tracts atlas and defined as delimited by the blue horizontal line. Masks were eroded (2D surface) before quantification, y.o. indicates years of age; Gd, gadolinium; anterolat, anterolateral; posterolat, posterolateral; FSL view refers to the FSL viewer software; ANTs, Advanced Normalization Tools (https://stnava.github.io/ANTs/); FLIRT, FMRIB’s Linear Image Registration Tool (http://www.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT); WM tracts atlas (https://sourceforge.net/p/spinalcordtoolbox/wiki/White%20Matter%20atlas/).32

The Dice similarity coefficient (range, 0–1, with 1 indicating perfect similarity), the Hausdorff distance (in millimeters), and the maximum median distance between skeletonized segmentations (in millimeters), using the SCT.27 For each subject, a comparison of the 3 WM/GM metrics between the 2 raters was performed; then, the average value per structure was calculated. All automatic segmentations presenting manual/automatic Dice coefficients of <0.8 (considered imperfect) were manually corrected under strict consensus between the 2 experts and subsequently used for multiparametric metrics quantification.

Morphologic measurements were then performed on all subjects by extracting the cross-sectional areas (CSA, in square millimeters) from automatic GM/WM segmentations (manually corrected if necessary) at each cervical level from C1 to C6 (C7 was discarded due to subject motion in 7/19 subjects [37%]).

Cervical SC lesions (characterized by signal hyperintensities) were finally depicted and counted from both STIR and T2WI. Then, for quantitative analyses, manual segmentation of lesions was performed on the T2*WI by expert 1 (S.D.) using both STIR and T2WI series as additional references to improve the accuracy.
of the lesion location. All the segmented lesions were visible on all 3 images. Normal-appearing (NA) regions were subsequently differentiated from lesion regions for metrics quantification.

For quantitative analysis, 4 DTI metrics (fractional anisotropy [FA], ADC, axial diffusivity [λ∥], and radial diffusivity [λ⊥], along with MT and ihMT ratios [MTR/ihMTR]) were derived afterward and quantified within the anterolateral WM and posterolateral WM at C2 and C5. The GM was not considered. These regions were chosen to comply with current ihMT HASTE limitations in terms of partial volume effect and the loss of spatial resolution (see Discussion). To match ihMT single-section quantification, only the middle section was considered for the DTI metrics. At the end of the postprocessing pipeline, the quantitative dataset for each subject included SC, GM, and WM CSA at each level from C1 to C6, as well as mean ± SD values of FA, ADC, λ⊥, λ∥, ihMTR, and MTR from anterolateral and posterolateral WM regions free from or having MS lesions (NA/lesional regions). For each metric, values from C2 and C5 were averaged to obtain a single value per ROI per subject for statistical purposes.

Last, brain lesions were manually segmented on T2 and FLAIR images to estimate their volume.

**Statistical Analysis**

Statistical analyses were performed using JMP9 (SAS Institute), considering P < .05 for statistical significance.

**Comparison of MR Imaging Data between Patients with MS and Healthy Controls**

For MS-related morphologic change evaluation, GM, WM, and SC CSA at each level were compared between patients with MS and HC, using nonparametric Wilcoxon rank tests corrected for multiple comparisons (n = 6 cervical levels) with the Holm-Bonferroni method.

Considering multiparametric data, we used the nonparametric Steel-Dwass all-pairs tests corrected for multiple comparisons to assess DTI/MT/ihMT metrics differences between patients and HC.

**Sensitivity of ihMT to MS-Related Changes**

DTI, MT, and ihMT metrics z scores were computed for NA and lesional anterolateral WM, posterolateral WM, and total WM for each patient with MS relative to mean values of HC. CSA z scores, all cervical levels considered, were calculated for global GM, WM, and SC.

Z scores of ihMTR were then compared with z scores of all metrics using nonparametric Wilcoxon rank tests to assess the sensitivity of ihMT to MS-related changes compared with the more conventional atrophy (CSA), DTI, and MT measures. Henceforth, the term “metric variation” will be used to refer to a change in the absolute z score value of the metrics.

**Correlations with Clinical Scores.** Correlations among brain lesion volume, SC lesion number, SC/GM/WM CSA, and DTI/MT/ihMT metrics variations (ie, z scores) in lesion and NA WM regions and each clinical score (disease duration, total MRC, handgrip, TWT, Nine-Hole Peg Test, and EDSS) were evaluated using the statistical nonparametric Spearman rank test. The resulting significant correlations were then corrected to control for multiple comparisons using the false discovery rate procedure.

**RESULTS**

**Patient Characteristics**

Demographic and clinical features of patients with MS are summarized in Table 1. The median EDSS score and disease duration of the patients at the time of MR imaging were 2.5 and 117 months, respectively.

**Reliability of the Automated AMU40 GM/WM Segmentation**

On-line Table 2 summarizes the similarity indices (Dice coefficient, Hausdorff, and skeletonized mean distances) with regard to manual GM/WM segmentations. The reader is referred to the On-line Figure for illustrations of representative AMU40-based GM segmentations from patients with MS with/without the presence of lesions. In the HC group, AMU40-based segmentations were found to be highly reliable with mean Dice coefficients of 0.85 ± 0.05 and 0.92 ± 0.02 for GM and WM, respectively. In patients with MS, 8/114 slices were discarded due to visually observed poor contrast in the images, mostly due to breathing motion of the patients. When applied to the remaining 106 slices, indices from WM (Dice, 0.89 ± 0.04) were close to those found in HC, whereas the values from GM segmentations (0.75 ± 0.14) were found to be lower, with higher SDs. Worse GM indices were found in slices with lesions (0.73 ± 0.12) compared with slices without lesions (0.79 ± 0.09), whereas WM indices were similar whether or not lesions were present in the slices. Afterward, 29% (31/106 slices) of the AMU40-based segmentations (with Dice, <0.8) were manually corrected for measurements of CSA and multiparametric analysis.

**Comparison of MR Imaging Data between Patients with MS and Healthy Controls**

Boxplots of GM and WM CSA from HC and patients with MS along with the corresponding z scores and percentages of variation are presented in Fig 2. A significant decrease in CSA for the SC (P < .001, data not shown) and WM (P < .003) across all cervical levels was found in patients with MS compared with HC. Decrease trends were additionally found for GM areas, significant at C6 (P = .003).

Figure 3 shows representative multiparametric datasets of 2 patients with MS (MS_1: 64 years of age, EDSS score: 2; MS_2: 60 years of age, EDSS score: 6) and a healthy control (65 years of age). Overall, in patients with MS, the DTI, ihMTR, and MTR values were all found altered in both NA and lesional WM regions (Table 2 and Fig 4) compared with the healthy control. λ⊥ and ADC showed trends toward an increase, albeit not significant, whereas ihMTR and MTR significantly decreased in both regions, with higher metric variations found in regions with lesions compared with NA tissue.

**Sensitivity of ihMT to MS-Related Changes**

Among all DTI, MT, and ihMT metrics, ihMTR showed the most important metrics variations (ie, higher absolute z score values) as illustrated in Fig 4. Results of Wilcoxon rank tests showed significant differences between these ihMTR variations and z scores.
of DTI and MT metrics ($P < .0001$), in both NA and lesional regions.

**Correlations with Clinical Scores**

Significant correlations were found between ihMTR $z$ scores in NA whole-WM and total MRC scores ($\rho = 0.80, P = .009$) as well as EDSS scores ($\rho = -0.73, P = .02$). MTR $z$ scores were also found correlated with total MRC scores ($\rho = 0.74, P = .02$) and EDSS scores ($\rho = -0.81, P = .004$). On the other hand, weaker and nonsignificant correlation trends were found for DTI metrics variations, such as between FA $z$ scores and MRC scores ($\rho = 0.64, P = .06$), for example.

**FIG 2.** CSAs (square millimeters) in HC and patients with MS. Mean values were extracted from AMU40-based GM and WM segmentations for each cervical level (C1 to C6). Boxplots of subjects’ mean CSA values across all cervical levels are also presented in the first column of each graph. The boxplots represent the minimum, first quartile (25%), the median, the third quartile (75%) and the maximum from upper to lower, respectively. Significant decreases in CSA in patients with MS relative to HC, assessed using Wilcoxon rank tests are reported as *$P < .05$, **$P < .001$, ***$P < .0001$. The section symbol indicates statistical differences that survive the Holm-Bonferroni correction for multiple comparisons. The percentage difference of CSA between MS and HC groups (in red) and the average $z$ scores (no unit, in blue) in patients with MS relative to HC at each level are also reported; a negative value indicates a decrease in CSA in patients.

**FIG 3.** Representative multiparametric maps extracted from a patient with relapsing-remitting MS (MS_1, 64 years of age and EDSS score, 2) and a patient with secondary-progressive MS (MS_2, 60 years of age and EDSS score, 6) along with maps from an age-matched HC (65 years of age) for illustration. yo, years of age.
Brain lesion volume was also significantly correlated with the TWT ($\rho = 0.66$, $P = .008$). Because correlations with the TWT were mostly related to data from a patient with important clinical disabilities (TWT = 73 seconds), they were not considered afterward. No other significant correlations were found.

DISCUSSION

This study provides 3 main findings: 1) AMU$_{40}$-based SC WM/GM segmentation is feasible in patients with MS; 2) ihMTR shows higher sensitivity to MS-related SC pathophysiologic changes compared with more conventional approaches; and 3) ihMTR might be useful in a clinical context.

MR templates and atlases are very useful tools for automated postprocessing pipelines. Across the years, effort has been directed toward construction and application of such templates in the SC. To our knowledge, only 2 studies applied and evaluated automated SC GM segmentation in MS: the work of Prados et al based on WM/GM segmentation using the Similarity and Truth Estimate for Propagated Segmentation algorithm and the work of Datta et al based on active contour segmentation, with both studies mostly focusing on the C2/C3 cervical levels. Our work is the first to apply the AMU$_{40}$ template in SC MS and investigate WM/GM segmentation capability from C1 to C6 cervical levels in such cases. The AMU$_{40}$-based segmentation was previously validated in healthy subjects and lesions extending on

| Table 2: Multiparametric values for anterolateral/posterolateral/total WM in HC and normal-appearing and lesional regions in patients with MS* |
|---------------------------------|-------------|---------------|-----------|
| ROI/MRI Metrics | HC (n = 19) | NA (n = 19) | Lesions (n = 19) |
| Anterolateral WM | | | |
| ADC (10$^{-3}$ mm$^2$/s) | 1.04 ± 0.08 | 1.08 ± 0.14 | 1.11 ± 0.17 |
| FA (n.u) | 0.65 ± 0.04 | 0.62 ± 0.07 | 0.61 ± 0.06 |
| $\lambda_\perp$ (10$^{-3}$ mm$^2$/s) | 1.90 ± 0.08 | 1.92 ± 0.15 | 1.96 ± 0.19 |
| ihMTR (%) | 5.88 ± 0.40 | 5.62 ± 0.43 | 4.30 ± 0.43 |
| MTR (%) | 32.04 ± 1.87 | 30.65 ± 2.25 | 29.69 ± 2.89 |
| Posterolateral WM | | | |
| ADC (10$^{-3}$ mm$^2$/s) | 1.00 ± 0.08 | 1.01 ± 0.10 | 1.09 ± 0.18 |
| FA (n.u) | 0.69 ± 0.03 | 0.67 ± 0.08 | 0.65 ± 0.06 |
| $\lambda_\perp$ (10$^{-3}$ mm$^2$/s) | 1.95 ± 0.07 | 1.90 ± 0.11 | 2.05 ± 0.17 |
| ihMTR (%) | 5.82 ± 0.39 | 5.49 ± 0.43 | 4.18 ± 0.70 |
| MTR (%) | 31.42 ± 1.23 | 30.12 ± 1.45 | 28.99 ± 2.57 |
| Total WM | | | |
| ADC (10$^{-3}$ mm$^2$/s) | 1.01 ± 0.08 | 1.04 ± 0.10 | 1.09 ± 0.16 |
| FA (n.u) | 0.68 ± 0.03 | 0.66 ± 0.07 | 0.64 ± 0.05 |
| $\lambda_\perp$ (10$^{-3}$ mm$^2$/s) | 1.94 ± 0.07 | 1.93 ± 0.10 | 2.02 ± 0.16 |
| ihMTR (%) | 5.84 ± 0.40 | 4.95 ± 0.55 | 4.27 ± 0.65 |
| MTR (%) | 31.58 ± 1.53 | 29.85 ± 1.56 | 29.32 ± 2.48 |

Note: (n.u) indicates no unit.

*Values are reported as group mean ± SD. Significant differences between metrics from control subjects and NA tissue in patients with MS as well as between HC and lesional tissue were assessed with Steel-Dwass tests and are indicated with the $P$ values below.

$^b$ $P < .001$.

$^c$ Differences that survived the Holm-Bonferroni correction for multiple comparisons.

$^d$ The only significant metric difference between NA and lesions in patients with MS ($P = .003$).

$^e$ $P < .0001$.

$^f$ Absolute (Abs) $z$ scores are used for visualization purposes and statistical comparison.

**FIG 4.** Spider graphs showing mean absolute $z$ scores of each quantitative MR metric in patients with MS relative to HC computed for NA and lesional anterolateral WM (left graph), posterolateral WM (middle graph), and total WM (right graph). Wilcoxon tests for paired values performed between ihMTR absolute $z$ scores and each of the other metrics to assess the sensitivity of the ihMT technique to microstructural changes in MS compared with DTI and MT show significant differences of all tested pairs, with $P < .0001$. Absolute (Abs) $z$ scores are used for visualization purposes and statistical comparison.
both GM and WM, mitigating the difference of contrasts between these structures. The template, nonetheless, demonstrated fairly good compliance with manual segmentations, with Dice coefficients (0.76 ± 0.09 at C2/C3) in line with the 2 previously mentioned MS studies, including patients with visible lesions (GM Dice coefficients: 0.80 ± 0.07 and 0.88 ± 0.03, respectively, at C2-3). The AMU40 template is not available for whole-spinal cord coverage or for 3D data; however, a former version of it (AMU13) could be found in the recently proposed PAM50 template (formerly the MNI-Poly-AMU template; https://sourceforge.net/p/spinalcordtoolbox/wiki/MNI-Poly-AMU), which presents the advantages of being compatible with larger spatial coverage, 3D acquisitions, and acquisition with different contrasts, while also proposing segmentation approaches based on machine learning. Comparison with this template version was beyond the scope of the study but would be interesting.

MS-related morphologic changes have been largely reported with evidence of cervical SC and GM atrophy in patients with MS relative to HC, with a strong correlation with clinical disability. In our study, SC, WM, and GM semiautomated measurements of CSA were performed along the cervical cord from C1 to C6 and results suggested atrophy of both SC GM and WM in patients with MS (5%/14% decrease on average, respectively), consistent with the literature, with more important and more significant atrophy in WM.

DTI and MT imaging have already been largely used to evaluate MS-related SC microstructural changes, with reports of higher ADC and lower FA and MTR in patients with MS relative to HC. In this study, DTI and MT metrics in patients with MS demonstrated overall similar variations, with higher variations in lesions compared with NA tissue, suggesting MS-related microstructural changes in the cervical SC as described in previous reports. However, both DTI and conventional MT techniques have been proved to lack specificity toward any pathophysiologic mechanism, due to confounding factors linked to inflammation, demyelination, edema, or axonal degeneration, making the use of more myelin-specific or sensitive methods interesting.

This work focused on ihMT. Previously published studies reported a supposedly higher ihMT specificity to myelinated structures compared with conventional MT as well as higher sensitivity to associated microstructural changes, in both healthy and pathologic SC (amyotrophic lateral sclerosis), and in the pathologic (MS) brain. In the present SC study, among all the considered MR metrics (CSA, DTI, ihMTR, and MTR), ihMTR had the most important and significant variations (absolute value decreases compared with HC and highest z scores), particularly in lesions. Although less important than in lesions, an ihMTR decrease in NA tissue of patients with MS was also found highly significant, suggesting tissue changes occurring away from a lesion or before lesion occurrence, currently not visible on conventional MR imaging. Better detection of such subtle changes could improve disease monitoring, particularly for future treatment based on neuroprotection. Of interest and though not assessed in the current MS work, the ihMT technique was previously demonstrated to be highly reproducible, with a small scan-rescan ihMTR variability (2.9%).

Finally, despite a readout module inducing severe blurring and precluding quantification within specific WM tracts or GM areas, ihMTR z scores measured in total NA WM regions were significantly associated with MRC and EDSS scores, whereas correlations between MTR and DTI with most clinical scores were not significant. These results could advocate for clinical and trial applications of the ihMT technique. They also support the higher sensitivity of the ihMT to MS-related deficits compared with DTI or MT. Further investigations related to the ihMT myelin specificity should now be performed, in conjunction with other myelin imaging techniques that have been successfully applied to SC imaging in recent years, such as quantitative MT imaging or myelin water fraction, which have already demonstrated great interest for SC MS investigations. Future studies should also focus on the ability to detect subtle and early changes.

Indeed, from the clinical point of view, in addition to the small number of patients, the main limitations were the cohort heterogeneity (especially regarding phenotypes, combining relapsing-remitting and secondary-progressive) and the disease duration. It would be relevant to include early forms of MS for longitudinal monitoring of ihMTR variation at different times to judge its utility to better characterize myelin-related processes. From the methodologic point of view, the readout module used for the ihMT measurements induced severe resolution loss and tissue blurring (~6 voxels for a T2 of 75 ms) in the phase-encoding direction (right-left), hence only allowing investigating anterolateral and posterolateral regions free from lesions or presenting lesions. Further acquisition strategies, such as recently presented, should help in achieving refined tissue separation. These latter strategies should also permit investigating multiple SC levels and hence characterize the diffuse pattern of cervical SC tissue alteration in MS (DTI/MT/ihMT acquisitions were limited here to C2 and C5 levels to keep the protocol duration to <1 hour). Optimized excitation pulse schemes and translation to higher fields (eg, 7T) should also be considered to increase spatial resolution, SNR, and sensitivity.

Finally, lesions were manually delineated; further studies should evaluate the use of the recently proposed automated MS lesion segmentation algorithms using convolutional neural networks or lesion-prediction algorithms.

Nevertheless, despite these limitations, the MR imaging acquisition and template-based postprocessing strategies used in the current study led to promising results supporting higher ihMT sensitivity toward cervical SC pathologic changes in MS, thereby inspiring future work to use this technique for MR imaging monitoring and as a tool in remyelination therapeutic strategies. Meanwhile, further investigations should be directed toward confirming this higher sensitivity in different cohorts and further evaluation of the specificity of ihMT to these myelin-related pathologic changes.

CONCLUSIONS

The use of a multiparametric MR imaging protocol combined with an automatic template-based GM/WM segmentation approach in the current study outlined a higher sensitivity of the ihMT technique toward SC pathophysiologic changes in MS compared with atrophy measurements, DTI, and conventional MT, presumably
due to its supposed higher specificity to myelin-content variations. The clinical correlations between ihmTR and functional impairment observed in patients with MS also argue for its potential clinical relevance, paving the way for future longitudinal multicentric clinical trials in MS.

ACKNOWLEDGMENTS

The authors thank V. Gimenez and L. Pini for study logistics, A. Rico for patient recruitment, and S. Gherib for brain lesion segmentation.


REFERENCES

MR Myelography for the Detection of CSF-Venous Fistulas

J.L. Chazen, M.S. Robbins, S.B. Strauss, A.D. Schweitzer, and J.P. Greenfield

ABSTRACT

SUMMARY: CSF-venous fistula is an important treatable cause of spontaneous intracranial hypotension that is often difficult to detect using traditional imaging techniques. Herein, we describe the technical aspects and diagnostic performance of MR myelography when used for identifying CSF-venous fistulas. We report 3 cases in which the CSF-venous fistula was occult on CT myelography but readily detected using MR myelography.

ABBREVIATIONS: CVF = CSF-venous fistula; SIH = spontaneous intracranial hypotension; VIBE = volumetric interpolated brain examination

CSF-venous fistula (CVF) is an important cause of spontaneous intracranial hypotension (SIH) and a critical diagnosis because treatment can result in a clinical cure.1-5 The fistulous connection is typically associated with a nerve root sleeve diverticulum; therefore, a detectable epidural fluid collection is often absent. Some authors advocate decubitus CT myelography to improve CVF detection,1,2,6 while others suggest that digital subtraction myelography may increase rates of fistula identification.7,8 We report surgically confirmed cases of CVF that were identified using MR myelography following the intrathecal administration of gadolinium.

MR Myelography Technique

Intrathecal gadolinium was instilled in the subarachnoid space via CT-guided lumbar puncture using a Gertie Marx Whitacre-style spinal needle (IMD). The opening pressure was measured and recorded, and 0.5 mL of gadobutrol (Gadavist; Bayer Schering Pharma) mixed with 4.5 mL of iohexol (Omnipaque 180; GE Healthcare) was intrathecally administered. After a 1-mL test injection, CT imaging confirmed intrathecal contrast, and 5 mL was instilled before the needle was removed with the stylet in place. The patient was positioned in a decubitus Trendelenberg position and rolled multiple times to disperse contrast throughout the subarachnoid space. The patient was then transferred to the MR imaging suite where MR myelography was performed on a 3T Biograph mMR scanner (Siemens) using multiplanar fat-suppressed T1 sequences (On-line Table). Imaging was reviewed in real time, and delayed thin-section axial sequences were acquired as needed. MR imaging was completed approximately 1 hour after intrathecal gadolinium administration.

Cases

Case 1. A 26-year-old woman with a 3-year history of orthostatic headaches after a spontaneous onset had brain MR imaging findings characteristic of SIH, including sagging, pituitary enlargement, venous distention, small subdural collections, and pachymeningeal enhancement. A CT myelogram was unrevealing for CSF leak, and 3 epidural blood patches were unsuccessful. She subsequently developed upper extremity pain and paresthesias, leading to the discovery of a large cervicothoracic syrinx on cervical spine MR imaging (Fig 1). She was treated with a posterior fossa decompression to improve craniocervical CSF flow but obtained neither symptom nor imaging improvement. Eventually, she underwent MR myelography with an opening pressure of ~4 cm H2O and was discovered to have a prominent left T7–T8 spinal meningeal diverticulum fistulizing to a left paraspinal vein draining into the azygous system. She also had early contrast opacification of her renal collecting system on the MR myelography completed approximately 1 hour after intrathecal contrast injection (Fig 2). The patient was offered and accepted surgical CVF ligation at T7–T8. The diverticulum was identified with dilated draining veins, and the entire nerve root complex with the CVF was interrupted with...
an aneurysm clip and silk suture (Online Fig 1). The patient experienced an immediate improvement in symptoms followed by a transient new headache and sixth-nerve palsy 3 days following the operation. This was presumed to be secondary to transient rebound intracranial hypertension, and her symptoms slowly resolved. Eight months postoperatively, she is without headaches, and follow-up MR imaging revealed both a dramatic resolution of brain sagging and near-complete resolution of the spinal cord syrinx (Fig 1).

Case 2. A 64-year-old man presented with a 9-month history of orthostatic headache, distorted hearing, and tinnitus with onset during Valsalva exacerbated by severe constipation. Brain imaging demonstrated a venous distention sign without other intracranial sequelae of SIH. Standard CT myelography could not identify a CSF leak, and multiple epidural blood patches were unsuccessful. An MR myelogram with an opening pressure of 9 cm H$_2$O revealed a left T8–T9 spinal meningeal diverticulum and adjacent opacification of a paraspinous vein (Fig 3). Small-volume contrast was also seen in the renal collecting systems. The patient was offered and accepted surgical CVF ligation at T8–T9. The diverticulum was identified with dilated draining veins, and the entire nerve root complex with the CVF was interrupted with an aneurysm clip and silk suture. The patient reported an immediate improvement in symptoms with return to baseline and resolution of headache, sustained at the most recent 4-month follow-up visit.

Case 3. A 53-year-old woman presented with a 2-year history of spontaneous-onset orthostatic headache and neck pressure. Brain imaging findings were suspicious for SIH with sagging, pituitary enlargement, venous distention, small subdural collections, and pachymeningeal enhancement. A CT myelogram revealed multiple spinal meningeal diverticula and Tarlov cysts without evidence of spinal fluid leak, and the patient received multiple epidural blood patches with only transient improvement of symptoms. An MR myelogram with an opening pressure of <4 cm H$_2$O revealed a left T9–T10 spinal meningeal diverticulum and an opacified left paraspinous vein (On-line Fig 2). The patient was offered and accepted surgical CVF ligation at T9–T10. The diverticulum was identified with dilated draining veins, and the entire nerve root complex with the CVF was interrupted with an aneurysm clip and silk suture. The patient reported an immediate improvement in symptoms with return to baseline and resolution of headache, sustained for 7 months at the most recent follow-up visit.
CSF-venous fistulas are an under-recognized cause of SIH. Improved detection using advanced imaging techniques may increase our sensitivity in detecting CSF leaks. CT myelography may detect CVF, and the technique can be augmented with decubitus positioning and provocative maneuvers to increase CSF pressure.\(^5\)\(^6\) Digital subtraction myelography has also been used to identify CSF-venous fistulas. However, both of these techniques require substantial ionizing radiation exposure to the patient, and the digital subtraction myelography technique is operator-dependent with different techniques described in the literature.\(^7\)\(^8\) Intrathecal administration of gadolinium is an off-label use in the United States but is well-tolerated and frequently used for clinical and research applications.\(^10\)\(^11\)

However, while no long-term harmful effects from intrathecal gadolinium are known, this has not been extensively studied. Only a macrocyclic gadolinium agent (gadobutrol) was used in this study, which has a weaker association with deposition compared with linear gadolinium agents.\(^12\) MR myelography eliminates the ionizing radiation exposure aside from any used for lumbar puncture. Additionally, multiple imaging sequences can be repeated as necessary to confirm a suspected fistula. Furthermore, while the attenuation of intrathecal CT contrast fades rapidly within an hour of intrathecal administration, intrathecal gadolinium maintains robust T1-shortening MR imaging signal for hours after administration. We found the axial T1 fat-suppressed volumetric interpolated brain examination (VIBE) imaging sequences to be particularly helpful. These may help highlight a fistula on MR myelography that may be less apparent on other imaging modalities. The technique is further bolstered by the improved soft-tissue contrast of MR imaging and the prominent signal intensity of gadolinium on fat-suppressed T1-weighted images. MR myelography has been shown previously to detect leaking gadolinium from perineural cysts.\(^13\)

Outstanding questions remain regarding the diagnostic performance of MR myelography and other myelographic techniques in the detection of CVF because no known criterion standard is available. We report 3 patients who had unrevealing CT myelograms in whom CVF was detected on MR myelography, all of whom experienced marked clinical improvement following ligation of the suspected fistula. This report raises important questions for further study regarding the optimal technique for imaging patients with suspected CSF leak.

In conclusion, MR myelography with intrathecal gadolinium may identify CSF-venous fistulas in patients without readily apparent fistulas on CT myelography. Further study is needed to confirm these promising findings.

**REFERENCES**


**FIG 3.** CT and MR myelogram images from case 2. A, CT myelogram reveals perineural cysts at T8–T9 without opacification of the left-sided cyst. B, MR myelogram axial T1 VIBE demonstrates opacification of the left-sided spinal meningeal diverticulum and left paraspinal vein draining into the azygous system (arrows), consistent with a CSF-venous fistula.
Chymopapain Chemonucleolysis: Correlation of Diagnostic Radiographic Factors and Clinical Outcome

The therapeutic response to treatment of lumbar disk herniations with chymopapain chemonucleolysis is significantly influenced by the extent of the radicular symptoms. Although classic radiological techniques may determine the frequency of lumbar radiculopathy, some patients do not achieve satisfactory relief of pain with chemonucleolysis, in an attempt to identify the factors which influence the outcome of chemonucleolysis. The correlation of preoperative radiographs and clinical response were studied in 55 cases. No significant correlation between the severity of disk herniation and the level of improvement was found. Although Disk height and disk herniation were not significantly related to the clinical response, in patients who had radicular radiographic evidence of disk space narrowing and discal material, evidence of coronal and sagittal disk degeneration, evidence of foraminal encroachment, or evidence of adjacent intervertebral disc disease, the clinical response to chemonucleolysis was significantly better than those patients with a complete disk height preserved. It was also observed that, in patients with a single severe root compression, those patients with a single severe root compression were (91%) than those with mild disc space narrowing. Clinical variations were not related to the frequency of successful root relief, length of follow-up, or the amount of radicular compression responsible for chemonucleolysis.

Chymopapain, a pepsin-like enzyme derived from papaya, degrades the major components of the extracellular matrix of the intervertebral disk. It was introduced into clinical practice in 1954. Smith (1) first demonstrated the possibility of using chymopapain to enzymatically disassemble and resorb intervertebral disk material. He reported the use of chymopapain in patients with lumbar radiculopathy. On the other hand, Preminger et al. (2) reported improved results when a more prolonged procedure was followed. By 1970, over 11,000 patients had been treated with chymopapain chemonucleolysis in 75 centers (2, 3). Variations were reported in the frequency of successful root relief, length of follow-up, and the amount of radicular compression responsible for chemonucleolysis.

Computed Tomographic Changes of Hypertensive Encephalopathy

Computed tomographic (CT) scan in a 69-year-old hypertensive hypertensive patient showed focal areas of decreased density, and normal radiologic changes were seen in the remaining cerebral structures. Hypertensive encephalopathy is an entity that is associated with severe stroke, diastolic, and (or) systolic hypertension, and may affect the thalamus, thalami, and periventricular white matter. These abnormalities are seen on CT scans and are considered to represent hypertensive encephalopathy. Hypertensive encephalopathy is an entity that is associated with severe stroke, diastolic, and (or) systolic hypertension, and may affect the thalamus, thalami, and periventricular white matter. These abnormalities are seen on CT scans and are considered to represent hypertensive encephalopathy.

Materials and Methods

Computed tomography (CT) scans of 11 patients with hypertensive encephalopathy were performed at New York University Medical Center and New York University Medical Center. The scans were obtained, in all cases, at 7 or 8 mm intervals during the time of fluid or fluid diversion. The examination was repeated in the same manner at 3 to 10 days. Statistical analysis was performed using Student's t-test. The results were compared to the previous findings.

Results

The results of the study suggest that the changes seen on CT scans in patients with hypertensive encephalopathy are related to the degree of hypertension and the duration of hypertension. The degree of hypertension is determined by the level of diastolic and systolic blood pressure. The duration of hypertension is determined by the length of time the patient has been hypertensive. The results of the study suggest that the changes seen on CT scans in patients with hypertensive encephalopathy are related to the degree of hypertension and the duration of hypertension. The degree of hypertension is determined by the level of diastolic and systolic blood pressure. The duration of hypertension is determined by the length of time the patient has been hypertensive.
Identification of Vortex Cores in Cerebral Aneurysms on 4D Flow MRI

We read with interest the study by Futami et al,1 “Identification of Vortex Cores in Cerebral Aneurysms on 4D Flow MRI.” They described the visualization, complexity, and stability of vortex cores on 4D flow MR imaging in a series of 40 cerebral aneurysms. Discriminating single, stable, vortex cores from complex flow patterns with temporarily unstable vortices, they found an association between a complex vortical flow pattern and complex aneurysm morphology. Thus, they concluded that the identification of vortex core flow patterns on 4D flow MR imaging may help to stratify aneurysm rupture risk. Moreover, they explained the mathematic definition of a vortex core line and referred to other previous research using computational fluid dynamics simulation data models that describe the phenomenon of vortex cores with cerebral aneurysm models.2,3

We agree with Futami et al1 that vortex core identification is a promising new tool for aneurysm assessment, which is in line with our findings published more than 10 years ago.4 Our group first described normal intracranial artery hemodynamics with a helical flow phenomenon in the carotid siphon using in vivo 4D flow MR imaging.5 Moreover, we were the first to describe intra-aneurysmal hemodynamics, including the visualization of intra-aneurysmal vortex cores.4 In these studies, we performed postprocessing and visualization of aneurysm flow using a commercially available 3D visualization software package (EnSight; Computational Engineering International, Apex, North Carolina). Besides 3D streamlines, 2D vector graphs, and time-resolved particle traces, a dedicated integrated algorithm of this software was applied to segment and illustrate vortex core patterns from 3D vector fields. Thus, the core of the vortical is represented by colored bars indicating the location and magnitude of the vortical flow.

We were also able to analyze the temporal and spatial stability of vortex cores inside the aneurysm during the whole cardiac cycle (Fig 3 from our original work4). Likewise in the work by Futami et al,1 a link between the stability and complexity of intra-aneurysmal vortex core patterns and aneurysm geometry, including the aspect ratio, was disclosed in a series of 5 cerebral aneurysms. In 2 of these aneurysms with a smooth surface and low aspect ratio, temporally stable, single vortex cores were identified, whereas in the other aneurysms with a lobulated surface and/or high aspect ratio, vortex cores showed spatio-temporal and magnitude alterations during the cardiac cycle.

We acknowledge the authors’ work studying vortex core patterns in a higher number of aneurysms and support their conclusions drawn from their cohort, though mostly ruptured cases were included. However, we would like to stress that correct reference to the existing body of literature should be given relative to the original description of in vivo visualization of aneurysmal vortex core flow.

REFERENCES

Disclosures: Michael Markl—UNRELATED: Consultancy: Circle Cardiovascular Imaging; Grants/Grants Pending: research support: Siemens; research grant, Circle Cardiovascular Imaging; research grant, Cryolife.4 *Money paid to the institution.

http://dx.doi.org/10.3174/ajnr.A6480

S. Meckel
Department of Neuroradiology
Faculty of Medicine
Medical Center, University of Freiburg
Freiburg, Germany

M. Markl
Departments of Radiology and Biomedical Engineering
Feinberg School of Medicine and McCormick School of Engineering
Northwestern University
Chicago, Illinois

S. Wetzel
Institute of Neuroradiology
Hirslanden Clinic Zurich
Zurich, Switzerland
We applaud the advanced and multidirectional studies performed by Wetzel et al.1 and Meckel et al.2 We agree with Meckel et al., who showed that aneurysms with an irregular shape or a high aspect ratio may have complex and unstable flow.

In our study,3 we differentiated vortex core lines from vortex cores. We identified vortex cores as thin streamline bundles with minimum velocities by progressively decreasing a threshold to visualize flow streamlines with velocities under the threshold in aneurysms. In extremely low velocity thresholds, a vortex core was exhibited as a line. In our study, we did not include such lines into vortex cores, because we could not definitely discriminate the lines from irregular streamlines which were usually visualized in aneurysms via 4D flow MR imaging. In our study, the velocity-encoding (VENC) for 4D flow MR imaging was set to 40–60 cm/s to correctly extract low-velocity vectors in aneurysms. We verified that the thin streamline bundle was a vortex core by showing that the bundle passed through the center of vortical flow vectors in the aneurysmal dome.3

Meckel et al.2 showed a vortex core line using an algorithm to identify the points on the face where reduced velocity is zero and to connect these distinct points.4 However, the locations of vortex core lines can vary between algorithms.3 In 4 of 5 aneurysms examined in their study, the VENC for 4D flow MR imaging was 90 cm/s to detect high-velocity flow components.2 This VENC value can lead to errors in the extraction of flow vectors with low velocities in aneurysms, which may cause mislocations of the zero velocity points. Therefore, we suggest that it should be verified that their vortex core lines are located in the center of vortical flows in aneurysms. This verification would guarantee the usefulness of the tool integrating the algorithm to automatically identify vortex core lines in aneurysms and to avoid the interobserver differences. We hope that the tool helps to clarify the role of the vortical flow pattern in aneurysm behavior in future studies.

REFERENCES

Sundaram et al\textsuperscript{1} reported lucidly on agreements and correlations of ASPECTS for NCCT (manual, automatic) and CTP-CBV (manual). But why use ASPECTS, not volume? Invented as a volume surrogate, ASPECTS takes the form of “content analysis”—a research method to quantify texts and symbols,\textsuperscript{2} wherein ASPECTS unitizes the MCA territory by defining 10 regions and categorizes each unit as having either no or any acute infarction. Measurement is degraded by the unitization’s vague definitions and diverse volumes, the binary categorization of so few units, and the bias of “any.” ASPECTS promotes interobserver disagreement (masked by “auto-”) and scatters markedly away from true volume. It is not excused for individual patient use by mere population correlation. Moreover, ASPECTS tends to advance infarcts by sometimes large margins of error beyond a supposed threshold of risk or futility, denying thrombectomy. A detailed explanation of ASPECTS and its flaws and a quick remediation by manual volumetry using the diameter formula $ABC/2$ (or a PACS-friendly refinement called $2Sh/3$) are available.\textsuperscript{3}

Besides the well-documented $ABC/2$ and $2Sh/3$, the software behind any auto-ASPECTS can, in principle, report infarct volumes. If these and other authors skip ASPECTS and use volume, their agreements and correlations might be stronger and their plots less scattered.

Disclosures: Marco C. Pinho—UNRELATED: Consultancy: ABC Medical Education. Comments: On-line educational activities related to MRI.

REFERENCES


http://dx.doi.org/10.3174/ajnr.A6485
We thank Drs Suss and Pinho for their interest in reading our article and sharing their opinion on the use of ASPECTS in patients with acute ischemic stroke. Several valid points have been raised about the limitations of ASPECTS, with which we agree.

Regardless of its limitations, ASPECTS is currently being used for treatment decision-making in patients with acute stroke who present <6 hours from symptom onset. This use is supported by a strong level of evidence that has emerged from randomized clinical trials and is included in the latest American Heart Association guidelines.1

Volume (of ischemic core) measurements if accurate and reliable can improve the screening process of patients with acute ischemic stroke and avoid several shortcomings of ASPECTS, including variability and ambiguity associated with the binary nature of ASPECTS.

While we have seen the successful use of ischemic core volume estimation obtained from CT perfusion or MR imaging to treat patients with stroke presenting beyond 6 hours from onset,2 current data on volumetric estimates of ischemic core via NCCT alone are limited. It is plausible that with a future successful large, prospective study, volumetric assessment of ischemic core on NCCT could potentially replace ASPECTS. This scenario may be not too far ahead, in particular considering the emergence of deep learning techniques with the potential for making this process even more reliable and effortless.3

REFERENCES

K. Nael
Department of Radiological Sciences
David Geffen School of Medicine at UCLA
Los Angeles, California

http://dx.doi.org/10.3174/ajnr.A6504
How Far Can We Take Vessel Wall MRI for Intracranial Atherosclerosis? The Tissue is Still the Issue

We read with great interest the recent article by Zwartbol et al, and the accompanying commentary by Dr Chan. While the data from the SMART-MR cohort add to the body of research on intracranial vessel wall MR imaging (vwMRI) studies, we have concerns with the interpretation of the results, specifically that the findings do not support “a different etiology” between intracranial (ICAD) and extracranial (ECAD) atherosclerosis. While vwMRI allows us to visualize disease in the vessel walls themselves, this provides only a snapshot of a dynamic process (atherogenesis) that is years-to-decades in the making. We are far from being able to see underlying processes at a molecular and tissue level that lead to these vwMRI findings.

Since the Virchow era, it has been understood that the distribution of atherosclerosis is not uniform; plaques occur primarily at bifurcations, in curved segments, and in coronary arteries. Intracranial and extracranial arterial systems are derived from different germ cell layers, ectoderm and mesoderm, respectively, and have distinct structural elements and physiologic flow characteristics. Accordingly, the risk factors for the development of ECAD and ICAD differ with respect to systemic processes like hyperlipidemia, hyperglycemia, or hypertension. Furthermore, atherosclerosis causes symptoms through different mechanisms and recurrence rates by site. As providers who are often frustrated by suboptimal treatment options for ICAD, we can confirm what has been known for decades, that atherosclerosis responds to treatment differently depending on location. Indeed, Dr Chan cites a seminal manuscript from 1964. Pathophysiologic mechanisms must be taken into account when considering what is seen on imaging.

The association between ICAD and general vascular risk factors examined in Zwartbol et al has previously been identified. A major limitation of these factors is over-reliance on stenosis. Prior work suggests that high-grade stenosis in ICAD is a feature of late-stage disease, and algorithms relying on stenosis measurement alone can ignore nonstenotic-but-clinically-active disease. A major benefit of vwMRI is the ability to identify plaque features like remodeling pattern, enhancement, fibrous cap thickness, proximity to perforator ostia, intraplaque hemorrhage, or lipid-rich necrotic core volume. Using vwMRI to merely identify the presence of ICAD misses an opportunity to better characterize the disease.

Another limitation of the methodology in Zwartbol et al is identification of ICAD burden regardless of symptomatology. The prevalence of ICAD is rather high, but most lesions prove harmless. The advantage of the granularity of vwMRI is that it could potentially distinguish the multiple ways ICAD can cause ischemic strokes and possibly predict which ICAD plaques are most likely to cause stroke, regardless of level of stenosis. This study also evaluates ICAD and ECAD cross-sectionally, without any clarification of the temporal relationship of disease development in each vascular bed. In addition, there is no indication of ethnicities within the study cohort of 130 patients; certain ethnic populations preferentially develop atherosclerosis in different vascular beds.

This critique is not meant to disparage the work of Zwartbol et al or the fine research conducted by the UCC-SMART study group. However, this must be viewed as an opportunity to refocus on the need for tissue validation of vwMRI for the investigation of ICAD, a feat that will prove difficult and likely rely on animal models. The burden of proof lies in such validation, and the value of vwMRI lies in its potential to identify processes that mediate disease rather than assuming there is a common pathophysiologic driver of atherosclerosis in all vascular beds.

Disclosures: Adam de Havenon—RELATED: Grant: National Institutes of Health—National Institute of Neurological Disorders and Stroke, Comments: supported by the National Institutes of Health—National Institute of Neurological Disorders and Stroke K23NS050924.* Joseph Scott McNally—RELATED: Grant: American Heart Association, National Institutes of Health, Comments: support from the following grants: 1) American Heart Association Scientist Development Grant (Principal Investigator) 17SDG33460420; 2) American Cancer Society, Huntsman Cancer Institute grant (Principal Investigator) 129785RG1619001IRG; 3) National Institutes of Health (Co-Investigator under D. Dennis Parker, Principal Investigator) 1R01 HL127582-01A1.* *Money paid to the Institution.

REFERENCES
1. Zwartbol MH, Geerlings MI, Ghaznawi R, et al. Intracranial atherosclerotic burden on 7T MRI is associated with markers of


M.D. Alexander
Departments of Radiology and Imaging Sciences and Neurosurgery

A. de Havenon
Department of Neurology
University of Utah
Salt Lake City, Utah

M. Mossa-Basha
Department of Radiology
University of Washington
Seattle, Washington

J.S. McNally
Departments of Radiology and Imaging Sciences
University of Utah
Salt Lake City, Utah
We thank Dr Alexander and colleagues for their interest in our work and appreciate the opportunity to respond to their concerns. A concern is raised regarding the interpretation of our results, which led us to conclude that our findings do not support “a different etiology” between extracranial atherosclerosis (ECAS) and intracranial atherosclerosis (ICAS). We agree with the data used to substantiate their concern, including data on histopathology and etiology, but do not share their opinion that our findings are contradictory. As described in our article—and noted by Alexander et al—there are several limitations to our study design, which is why we tried to carefully word our conclusion to avoid overinterpretation of our results. We showed that there is a relationship between ICAS and several clinical markers of ECAS, suggesting a similar etiology; or in other words (used in the article), our (limited) results do not support a different etiology. However, there are indeed differences in etiology and pathophysiology, which we do not claim to contradict. Apparently, we were not able to successfully convey this subtle difference to all readers.

As carefully described by Alexander et al, ICAS is a complex disease with many facets, and intracranial vessel wall MR imaging (vwMRI) facilitates our attempts to elucidate parts of this disease process. We agree that vwMRI can provide us with information not only on the presence or absence of atherosclerotic plaques but also on plaque components and the (from extracranial atherosclerosis) well-known “vulnerable plaque.” However, most vessel wall lesions encountered in our populations, Caucasian subjects, which Alexander et al justly noted were not specifically mentioned in our article, are too small to accurately measure, let alone characterize, apart from the presence or absence of enhancement. It would indeed also have been of interest to correlate vessel wall lesion locations with cerebrovascular symptoms, but this would necessitate a different study population—with more (cerebrovascular) symptomatic patients—to reach any statistically significant results.

Finally, we would like to provide our whole-hearted support for the suggestion of Alexander et al to “refocus the need for tissue validation of vwMRI.” The ease of characterizing extracranial carotid atherosclerotic plaques is in stark contrast with the difficulty of performing the same analyses in intracranial plaques, which is mainly caused by the lack of methods to compare in vivo–acquired vwMRI with ex vivo histopathology. Nonetheless, sequence validation is essential for our conception of how we should interpret vessel wall lesions on vwMRI and how we should incorporate intracranial vwMRI in clinical practice. We would like to thank Alexander and colleagues for their critical review and compliment them on highlighting the possibilities of vwMRI and opportunities for further research.

REFERENCES

M.H.T. Zwartbol
A.G. van der Kolk
Department of Radiology
M.I. Geerlings
Julius Center for Health Sciences and Primary Care
University Medical Center Utrecht and Utrecht University
Utrecht, the Netherlands

http://dx.doi.org/10.3174/ajnr.A6536
We read with interest the editorial of Aftab and Salman about discontinuing the use of the term “hemorrhagic stroke.” The word “stroke” was coined in medicine to define an acute focal deficit event of the central nervous system of vascular origin and has been very useful in clinical practice. The division into ischemic and hemorrhagic has served to make a rapid classification of acute cerebrovascular events.

The term “stroke” is unique in the medical literature in English and is not used in other languages. Older words such as “apoplexy” that have their equivalent in other languages have long since ceased to be used, and it is preferred to use “cerebrovascular attack,” “event,” or “accident.” With the advances in the images of the central nervous system and with the globalization of medical knowledge, it is no longer justified to use the term “stroke” in neuroradiology to define lesions of vascular origin. Terms such as “infarct,” “hypoperfusion,” or “focal ischemia” are more precise and better explain the pathophysiology of the neurologic deficit.

Finally, the editorial says that ischemic events are caused by hypoperfusion of a region of the brain secondary to vascular occlusion or hypoxia. It seems controversial to say that hypoxia causes cerebral hypoperfusion; usually, the opposite is true.

REFERENCES

http://dx.doi.org/10.3174/ajnr.A6479
I agree with Ortega et al that the term "stroke" is vague and more precise terms such as infarct or ischemia should be used in radiology reporting. Unfortunately, the term stroke is a ubiquitous part of the general (nonradiologic) medical lexicon and national guidelines, particularly given the new emphasis on timely stroke therapy. In this context, what is being referred to is ischemic stroke.

What used to be classified as hemorrhagic stroke is now best classified as simply "intracranial hemorrhage," which, as was pointed out in the original editorial, carries a different set of etiologies, symptoms, and therapies than ischemic stroke.

To ease the confusion on this terminology, it is being suggested that the term stroke be reserved for what used to be called ischemic stroke, supplanting the ancient definition of the term stroke.

Finally, examples of hypoxia causing ischemia would include etiologies of hypoxic-ischemic injury such as carbon monoxide poisoning. In these cases, vascular perfusion may be preserved but the perfusing blood may lack adequate oxygen, resulting in ischemia. Perhaps a comma may have helped, "Ischemic events are caused by hypoperfusion of a region of the brain secondary to vascular occlusion, or hypoxia."

http://dx.doi.org/10.3174/ajnr.A6528