AJNR

AMERICAN JOURNAL OF NEURORADIOLOGY

THE JOURNAL OF DIAGNOSTIC AND INTERVENTIONAL NEURORADIOLOGY

DCE perfusion in head and neck schwannomas and paragangliomas
Growth curves of subependymal giant cell tumors
Adrenoleukodystrophy early lesion growth
Contrast bolus interference in CT stroke protocols

Official Journal ASNR • ASFNR • ASHNR • ASPNR • ASSR

PERMEABILITY
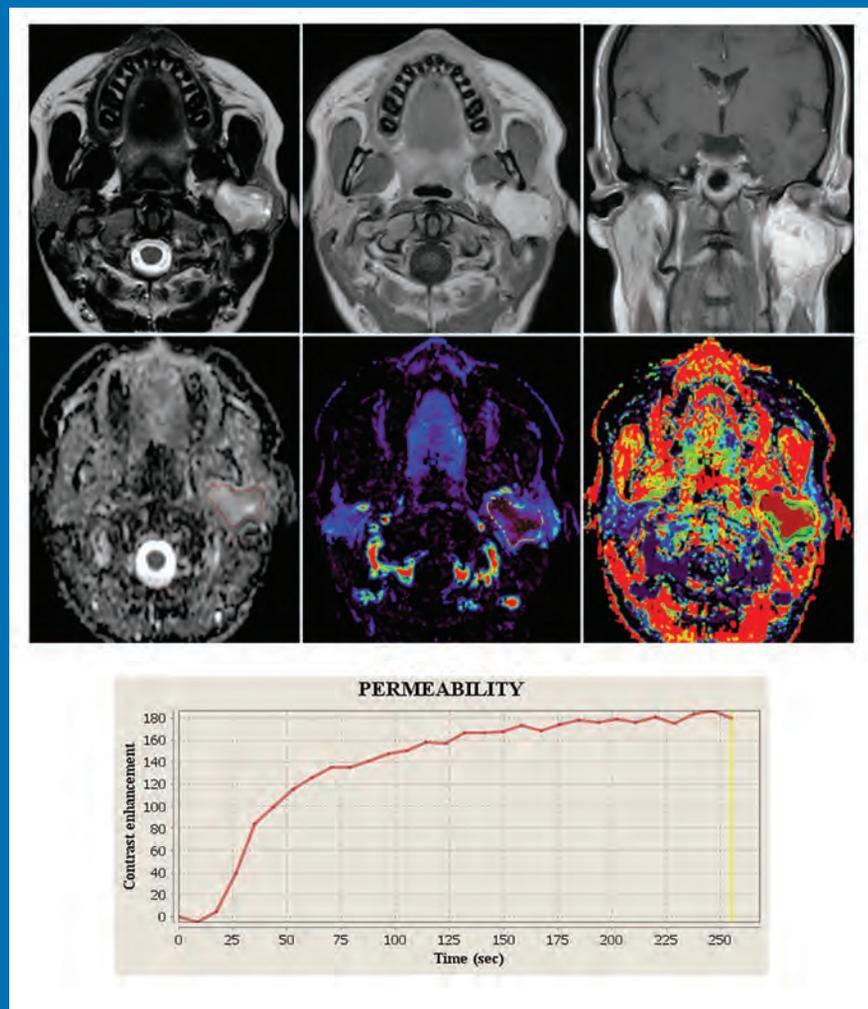
FRED™

**Flow Re-Direction Endoluminal Device**

FLOW DIVERSION.
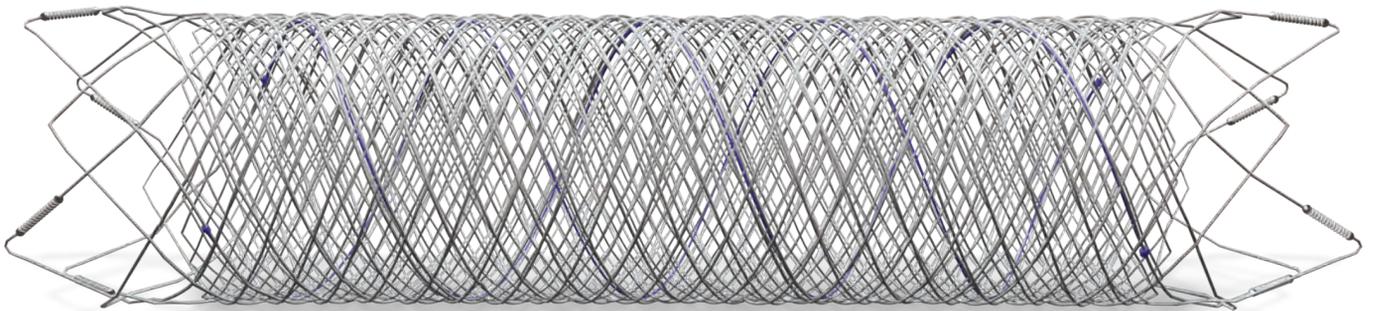
SIMPLIFIED.

The New Standard of
Ease and Simplicity
in Flow Diversion Treatment.
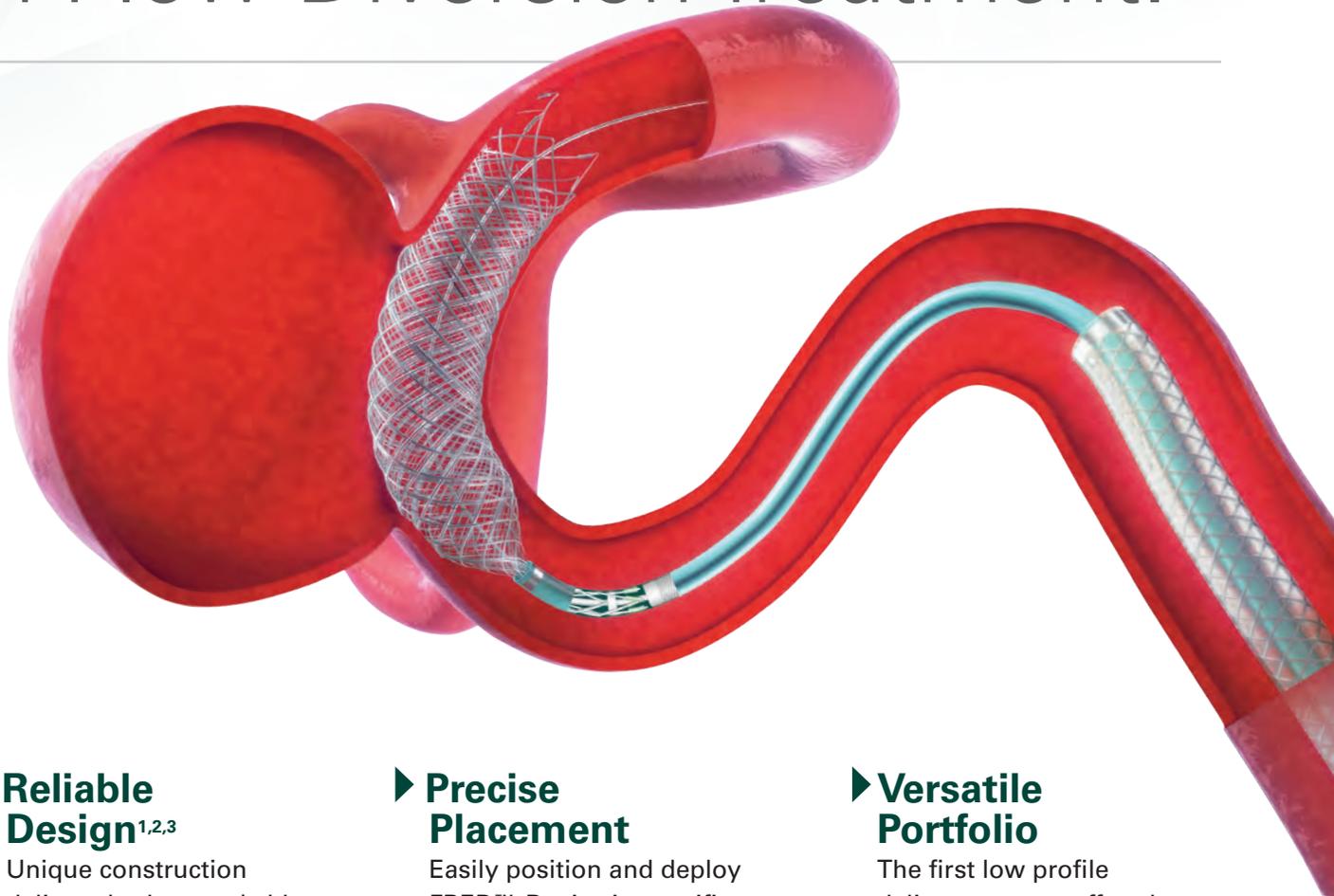
▶ **Reliable Design**[1,2,3]

Unique construction delivers both remarkable ease of use and excellent flow diversion[1,2,3]

▶ **Precise Placement**

Easily position and deploy FRED™ Device in specific, targeted locations[4,5,6]

▶ **Versatile Portfolio**

The first low profile delivery system offered in combination with large diameter and long length options

References:
1. TR11-211 2. TR13-171 3. TR15-055 4. TR13-192 5. TR15-072 6. TR19-145

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.

AJNR *go green*

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.

MIDRC
MEDICAL IMAGING AND DATA RESOURCE CENTER.

Medical Imaging in the Fight against COVID-19
Call for Contributions

The Medical Imaging and Data Resource Center (MIDRC) is a multi-organizational initiative in the fight against the ongoing worldwide pandemic co-led by the ACR®, the RSNA, and the AAPM and is hosted at the University of Chicago.

The primary goal is to accelerate machine learning research by creating a high quality COVID-19 image commons linked to relevant clinical data made available as a public resource.
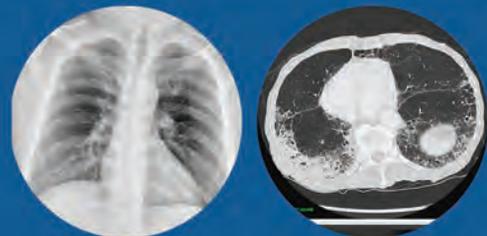
ASNR has teamed up with MIDRC to advance the understanding of the neuroimaging manifestations of COVID-19 and to enhance the MIDRC image commons.
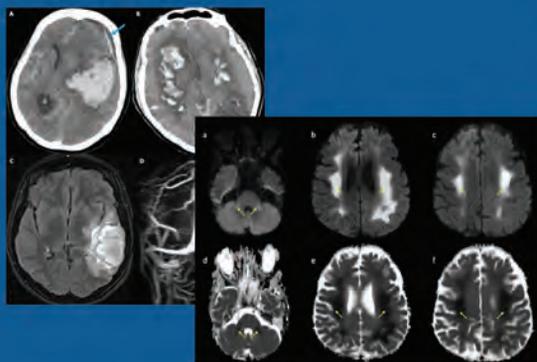
- To ensure success, MIDRC encourages all institutions and hospitals, both small and large, to donate their COVID-19 medical images and data to the cause.
- MIDRC provides a HIPAA-compliant mechanism to securely submit medical images and clinical metadata, and is committed to helping institutions throughout the submission process.

Visit midrc.org for more information

National Institute of Biomedical Imaging and Bioengineering
Creating Biomedical Technologies to Improve Health

Explore the New ASNR Career Center

Your Premier Resource for Professional Development

- Access to jobs in the National Healthcare Career Network
- Confidential resume posting
- One-time free resume review
- Professional online profile
- Saved jobs capability

Start here: careers.asnr.org

ASNR
American Society of Neuroradiology

ASNR 60th Annual Meeting and Symposium Neuroradiologicum XXII

CALL FOR ABSTRACTS

Join us May 13-18, 2022, to present the best scientific research in neuroradiology.

The Symposium Neuroradiologicum XXII will be held on May 13-16, followed immediately by the 60th Annual Meeting of the ASNR on May 16-18. Both meetings will be held in person and virtually.

ABSTRACT SUBMISSION DEADLINE
Monday, November 1, 2021 (11:59 PM ET)
Submit online at https://www.aievolution.com/asn2201/
Acceptance notifications will be sent on or before January 17, 2022, upon conclusion of peer review.

ABSTRACT SUBMISSION CATEGORIES
Scientific Oral Presentation: *will be included during ASNR ONLY*
Electronic Scientific Poster: *will be shown during both meetings*
Electronic Educational Exhibit: *will be shown during both meetings*
Excerpta (Oral Presentations): *will be included during SNR ONLY*

ABSTRACT SUBMISSION INFORMATION AND CRITERIA

1. By submitting an **Oral or Excerpta** abstract, you agree that the presenting author will register, attend and present in person. By submitting an **Electronic Poster or Educational Exhibit** abstract, you agree that the presenting author will register, attend and present virtually OR in person.

2. Presenters of accepted abstracts must register at their own expense for ASNR22 and/or SNR XXII.

3. All authors must have disclosure on file in the submission system. Remove authors with missing disclosures prior to submitting abstracts. Email **education@asnr.org** to add authors after the submission deadline.

4. The use of charts, tables and graphics are encouraged if relevant to the research submitted.

5. DO NOT SUBMIT DUPLICATE ABSTRACTS IN MULTIPLE CATEGORIES OR SUBMISSION TYPES.

6. Changes can be made to submitted abstracts until the deadline.

7. Available ASNR awards include the Cornelius Dyke Memorial Award, Outstanding Oral Presentation Awards and Educational Exhibit Awards.

8. The *American Journal of Neuroradiology (AJNR)* and *Neurographics* encourage presenters to submit manuscripts based on their work to *AJNR* or *Neurographics* before considering other journals.

Questions? Contact the ASNR Education Department at **education@asnr.org**.

ASNR22 | SNRXXII

May 13–18, NYC • Reconnecting The Global Neuroradiology Community

ONLINE FEATURES

LETTERS

BOOK REVIEWS R.M. Quencer, Section Editor

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

From Ota Y, et al. in this issue. A 41-year-old with a schwannoma in the left parotid space.

Indicates Editor's Choices selection
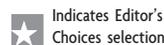
Indicates Fellows' Journal Club selection

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

Indicates article with supplemental online data

Indicates article with supplemental online video

Evidence-Based Medicine Level 1

Evidence-Based Medicine Level 2

Title: Brachial Plexus. Whether it is the crucial attention in tasks or the everlasting need for a keen eye, radiology will continue to be a field that requires precision and a focus on fine detail. As a medical student also striving to embrace my artistic abilities, particularly in human anatomy, it is essential that I am detail-oriented with my work in order to illustrate the various organ systems as realistically as possible. I will always seek to bring life to 2-dimensional images on paper, much like radiologists must interpret 2-dimensional images on a screen, correlating them to real human life. This piece uses pencil on paper to depict the various nerves of the brachial plexus and vertebrae C5 to T1.

Gabriella Mamo, Philadelphia College of Osteopathic Medicine

Imaging of Neuromodulation and Surgical Interventions for Epilepsy

M.E. Adin, D.D. Spencer, E. Damisah, A. Herlopian, J.L. Gerrard, and R.A. Bronen

ABSTRACT

SUMMARY: About one-third of epilepsy cases are refractory to medical therapy. During the past decades, the availability of surgical epilepsy interventions has substantially increased as therapeutic options for this group of patients. A wide range of surgical interventions and electrophysiologic neuromodulation techniques are available, including lesional resection, lobar resection, thermoablation, disconnection, multiple subpial transections, vagus nerve stimulation, responsive neurostimulation, and deep brain stimulation. The indications and imaging features of potential complications of the newer surgical interventions may not be widely appreciated, particularly if practitioners are not associated with comprehensive epilepsy centers. In this article, we review a wide range of invasive epilepsy treatment modalities with a particular focus on their postoperative imaging findings and complications. A state-of-the-art treatment algorithm provides context for imaging findings by helping the reader understand how a particular invasive treatment decision is made.

ABBREVIATIONS: ANT = anterior thalamic nucleus; ATL = anterior temporal lobectomy; DBS = deep brain stimulation; EEG = electroencephalogram; FGATIR = fast gray matter acquisition T1 inversion recovery; LITT = laser interstitial thermal therapy; MST = multiple subpial transection; MTS = mesial temporal sclerosis; RNS = responsive neurostimulation; VNS = vagus nerve stimulation

Epilepsy, a recurrent spontaneous seizure disorder, is the third leading cause of neurologic symptoms, with a prevalence of approximately 1%. In about one-third of cases, optimal seizure control cannot be achieved with medications alone, leading to significant morbidity and mortality. Surgical treatment for epilepsy was pioneered by Dr Victor Horsley in 1886, by resecting a focus of traumatic scar tissue and the surrounding tissue in a 22-year-old patient.[1] Currently, resective surgery is the technique that is most likely to result in seizure freedom for patients with pharmacoresistant epilepsy.[2] Complete resection of a focal MR imaging abnormality causing medically intractable epilepsy may result in >1 year of seizure freedom in a substantial proportion of patients.

Surgical interventions for refractory epilepsy include thermoablation, functional hemispherectomy, corpus callosotomy, and cortical transection as well as resection. During the past decade, the availability of epilepsy interventions has increased substantially beyond the traditional resective surgical treatments. These modalities include 3 FDA-approved neuromodulation devices. The specific type of surgical intervention is determined for each patient after extensive investigation, which sometimes includes localization via intracranial electrodes (Table and Fig 1).

The indications, imaging findings, and imaging features of potential complications of the newer interventions may not be

Types of surgical and neuromodulation epilepsy interventions

Interventions
Resective surgery
Temporal lobe
Temporal lobectomy (characteristically for MTS)
Lesionectomy
Extratemporal lobe resections
Hemispherectomy
Functional hemispherectomy
Hemispherotomy, disconnection
Nonresective surgery
Callosotomy, partial vs complete
Multiple subpial transection
Laser ablation surgery
Radiosurgery
Neuromodulation
Vagus nerve stimulation
Responsive neurostimulation
Deep brain stimulation

FIG 1. Intractable epilepsy decision tree for surgical intervention. Noninvasive investigations indicate whether a patient is a candidate for surgical resection of unilobar lesions or MTS, or whether an intracranial EEG study is needed for further localization (ie, MCD, diffuse or multifocal lesions, normal MR imaging findings, or discordant investigations). For those with unilobar lesions, surgical treatment is based on whether the functional cortex (defined as verbal memory, sensorimotor cortex, and language regions) is involved, determined by functional electrical stimulation mapping. RNS and possibly LITT are used to treat MTS in those with preserved hippocampal function, as opposed to temporal lobectomy for those with poor hippocampal memory. Nonfocal multifocal or network-onset seizures are treated by RNS, DBS, or VNS. Fxn cortex indicates functional cortex; HC fxn, hippocampal function; MCD, malformation of cortical development; Wada, intracarotid amobarbital testing.

widely appreciated, particularly if practitioners are not associated with comprehensive epilepsy centers. In this article, we review a wide range of invasive epilepsy treatment modalities including electrophysiologic neuromodulation, with a particular focus on their postoperative imaging findings and complications. We will initially explore preoperative imaging aspects of intracranial electrodes, then discuss resective and nonresective surgeries, and conclude with neuromodulation. The focus will be on imaging findings and complications specifically associated with epilepsy surgery.

Overview of Surgical Intervention

The goal of resective surgery is to render the patient seizure-free by optimizing the resection of epileptogenic brain and avoiding

neurologic and cognitive sequelae. Identification of the epileptogenic foci is required for all types of interventions for pharmacoresistant epilepsy and is determined initially through noninvasive means, including history and physical examination, audiovisual electroencephalogram (EEG) monitoring, neuropsychologic testing, MR imaging, PET, and sometimes SPECT and magnetoencephalography. Task-based fMRI is widely used for identification of eloquent cortices (sensory-motor and language) before resective surgery and has largely replaced Wada, an invasive test used for determination of the dominant hemisphere for language activity (see the algorithm in Fig 1). In a variety of cases, an intracranial EEG investigation may be necessary. During approximately the past 2 decades, evidence of a new concept of epilepsy has evolved,
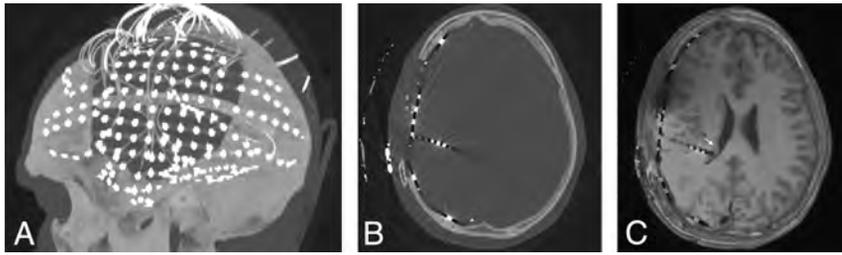
FIG 2. Postelectrode placement with sagittal MIP (*A*) from axial CT images (*B*) demonstrating a subdural electrode grid over an eloquent region (eg, visual-spatial region) and subdural strip and depth electrodes. To determine the exact localization of right temporoparietal seizure onset, we coregistered CT scans to MR images (*C*) because electrode positions are better identified on CT. A depth electrode is placed into the subependymal heterotopia to assess whether seizures are originating there (*arrow*).
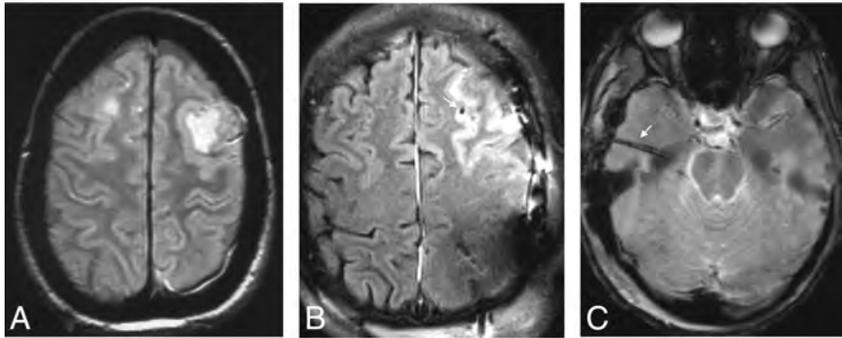
FIG 3. Electrode complications in 2 patients include edema (*A*) and subarachnoid hemorrhage (*B*), possibly related to vascular injury on FLAIR axial images (*arrow* points to adjacent electrode). Imaging findings of the electrode tracts (*arrow*) include hemosiderin on gradient recalled-echo imaging (*C*) and contrast enhancement on axial T1WI (not shown here).

with major implications for treatment; this theory suggests that epilepsy is a disorder of the neural networks in both generalized and focal epilepsies.[3] As a network disorder, destruction (through resection or other surgical techniques) or disruption (through neurostimulation of an EEG focus or the thalamus) of critical nodes or hubs in the network can lead to improvement in seizure outcomes. Convergence of localization of brain dysfunction through seizure semiology, EEG, and neuroimaging helps to determine the epileptogenic foci or critical hub in the network.

A variety of short- and long-term complications is associated with specific types of interventions for epilepsy. Infection and hemorrhage can be seen in all types of interventional epilepsy treatments, though they are rarely encountered. Fever, headache, unexpected seizures, and lethargy as well as other clinical findings of meningoencephalitis may raise suspicion of infectious complications. On imaging, the surgical route and instrumentations must be scrutinized for signs of infection, including, but not limited to, parenchymal enhancement, fluid collection, and restricted diffusion. Interval changes on follow-up imaging may help identify infection, though it may not be possible to reach a conclusion based on imaging alone, particularly in the early postsurgical period due to misleading expected findings, as discussed later in the respective sections.

Intracranial Electrodes

Intracranial electrodes are used when noninvasive investigations fail to localize the ictal onset or when the workup reveals findings discordant with an MR imaging lesion, occurring in 25% of patients with pharmacoresistant epilepsy. The need for intracranial electrodes is often necessary in cases without an MR imaging–identifiable focal lesion (focal lesion such as cavernoma, temporal lobe encephalocele, or tumor). At our institution, we investigate diffuse pathologic changes (such as large ischemic or inflammatory changes) and malformations of cortical development with intracranial electrodes because the seizure foci may not be directly linked to the abnormality on MR imaging or may be more widespread. Because focal cortical dysplasia, which now accounts for a substantial proportion of pharmacoresistant epilepsy, is often difficult to detect by MR imaging and the seizure focus may be divergent, intracranial electrodes are required. Intracranial electrodes may also be used to determine the relationship between nondispensable cortex (e.g., language or motor areas) and an epileptogenic focus. Electrodes are made of MR imaging–compatible materials such as platinum-iridium and can be imaged safely on a 1.5T scanner.[4] Despite the lack of American Society for Testing and Materials (ASTM) MR imaging labeling for some commercially available brands of platinum-iridium–based intracranial electrodes (because they were on the market long before 2005, when the FDA manual was released), there have been decades of excellent MR imaging safety records using transmit-receive head coils.

Electrodes can be placed using frame-based or frameless stereotactic guidance, or with robotic assistance. There are 2 main types of intracranial electrodes: 1) subdural grid or strip electrodes for neocortical coverage, and 2) depth electrodes that extend into the parenchyma (Fig 2, and Online Supplemental Data).

Precise identification of the intracranial electrodes is crucial because they play a central role in diagnosis and treatment. CT and MR images obtained after electrode placement allow coregistration of the exact location of electrodes defined by CT to MR imaging anatomy and abnormal findings. CT is frequently used early after electrode placement to exclude major intracranial complications, such as hemorrhage, edema, or mass effect. Electrode complications are rare (Fig 3). Punctate hyperintense signal changes on long TR images representing gliosis have been found in 41% of depth electrode tracks after removal, while hemosiderin, hematoma, enhancement, or calcification from the tracks are rarely seen.[5,6]
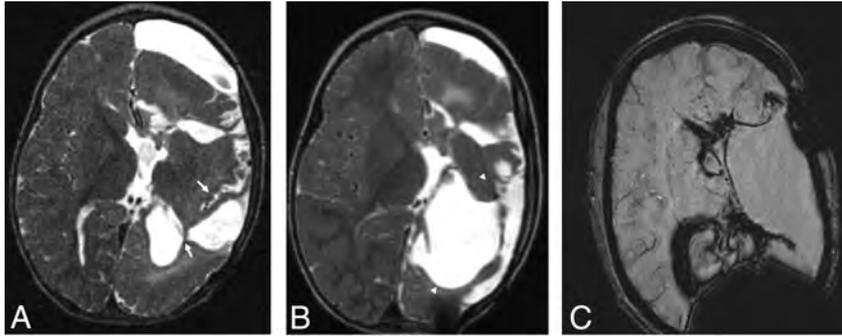
FIG 4. Staged functional hemispherectomy. Initial functional hemispherectomy for hemimegalencephaly consisted of left frontal, parietal, and temporal lobe resections and disconnection of most of the rest of the left hemisphere from the contralateral hemisphere and deep gray matter (*A*). Persistent seizures from the occipital lobe and insula indicated incomplete disconnection (*arrows* in *A*, axial T2WI) necessitating further resection of these structures (*arrowheads* in *B*, axial T2WI). Hemorrhage and hemosiderosis are complications of larger resections, as seen on axial SWI (*C*).

Resective Surgery

The efficacy and safety of surgery for pharmacoresistant epilepsy have been established by clinical trials in adults[7] and children.[8] Surgical decision-making may be relatively straightforward when there is a focal lesion or evidence of mesial temporal sclerosis (MTS) that is proved to cause seizure onset. The most common form of surgical resection is anterior temporal lobectomy (ATL), which is the primary surgical intervention method in refractory MTS. Other temporal lobe resective surgeries include lesionectomy or cortical resection. For extratemporal resections, achieving an excellent outcome is less likely in cases with negative findings on MR imaging,[9] underscoring the importance of presurgical lesion identification with meticulous imaging.

Temporal Lobectomy

The standard surgical resection in ATL involves removing the amygdala, hippocampus, and uncus of the parahippocampus, often en bloc, leaving the pia intact.[10] A number of studies have suggested a relationship between the amount of medial temporal lobe tissue resected and good postoperative seizure outcome. Specifically, if imaging reveals residual tissue involved with postoperative seizures, further surgery may result in an additional cure in 50% of patients.[10] There are other reasons for surgical failures, including extratemporal seizures mimicking temporal lobe epilepsy, such as from the insula.

In ATL surgery, the cerebral peduncle, posterior cerebral and communicating arteries, the anterior choroidal artery, thalamus, oculomotor nerve, and optic radiation/Meyer loop are at risk of injury. With the advent of microscopic neurosurgery, the incidence of surgical complications has notably decreased, though it is still reported to reach 7%.[11] Psychiatric and cognitive disorders, visual field deficits, hemiparesis, language disorders, hemorrhages, infections, and third and fourth nerve dysfunctions are among the clinical complications of ATL surgery. Hemorrhage, usually without mass effect, after ATL may occasionally be seen remote from the surgical site in the cerebellar vermis and folia, thought to be related to CSF loss.[10]

It is important to be aware of imaging findings after ATL epilepsy surgery. MR imaging findings may be misleading when performed in the immediate postoperative period, and MR imaging is only indicated when there is high clinical suspicion for potential complications or a necessity for baseline postoperative imaging (such as in tumors). On MR imaging, enhancement of the surgical margin is thin linear within the first 1–5 days but becomes thick linear or nodular (mimicking a tumor appearance) during the first week to month and then disappears (Online Supplemental Data).[12] Restricted diffusion in the parenchymal rim of ATL resections may occur in half of those imaged within 24 hours postoperatively, while more extensive involvement may indicate vascular injury.[13] Extraparenchymal findings include dural enhancement, which may last years, extra-axial fluid lasting 1–2 months, and pneumocephalus lasting 5 days.[12] Enlargement, sagging, and increased enhancement in the ipsilateral choroid plexus after ATL should not be mistaken for tumor (Online Supplemental Data).[14] Ex vacuo enlargement of the ipsilateral temporal horn is almost unequivocally present.

Extratemporal Lobe Resections. Frontal lobe resections are the most common extratemporal surgeries (48%).[15] In addition to frontal, parietal, and occipital locations (posterior quadrant surgery), the insular, cingulate, and hypothalamic regions are not infrequent extratemporal sites for epilepsy surgery. Extratemporal resective surgery is associated with a higher degree of complications compared with temporal lobe resection, despite a lower chance of disease cure.

Functional Hemisperectomy/Hemispherotomy

The hemispherectomy technique is used for patients with multilobar intractable epilepsy, single hemispheric dysfunction such as dysplastic hemimegalencephaly, Rasmussen encephalitis, Sturge-Weber syndrome, or ischemic infantile hemiplegic syndrome. However, anatomic hemispherectomy has been linked to late-onset hydrocephalus and superficial hemosiderosis, lethal complications that often occur a decade later. These complications have led to several modifications known as functional hemispherectomy and hemispherotomy (including peri-insular, parasagittal, and endoscopic hemispherotomies), with more limited resection of tissue and greater emphasis on disconnection of the epileptogenic hemisphere and an ensuing decrease in complications (Fig 4). Although intermittent intracranial hemorrhage with resultant hydrocephalus or subsequent siderosis has decreased, these complications still occur, particularly with large-volume reductions associated with functional hemispherectomy, which is more commonly performed with hemimegaloencephaly. In addition to assessing postoperative imaging studies for these complications (which may occur years later), it is critical to confirm that the
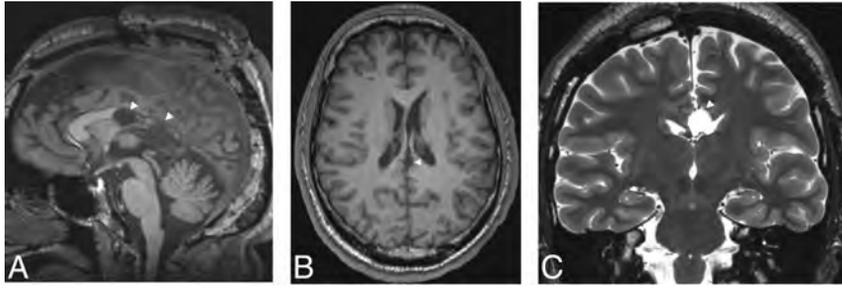
FIG 5. Corpus callosotomy. Disconnection of cerebral hemispheres posteriorly by posterior corpus callosotomy (*arrowheads*) is shown on sagittal (*A*) and axial (*B*) T1WI and coronal T2WI (*C*). Sagittal images may be inadequate for confirming complete disconnection if the plane of the resection undulates.

hemisphere is disconnected on initial postsurgery MR images, particularly for hemimegalencephaly (Fig 4). The most common cause for failure and reoperation is incomplete disconnection.

Nonresective Surgical Techniques

Corpus Callosotomy Disconnection Surgery. Corpus callosotomy is a palliative procedure, particularly for intractable drop seizures. By disconnecting the hemispheres, callosotomy disrupts the propagation of seizures rather than cure them. Total or anterior corpus callosotomy are the most common procedure types. The former has a higher likelihood of seizure reduction, while the latter is less likely to result in disconnection syndromes.[16] Postoperative imaging is often helpful in documenting the extent of callosotomy, particularly when incomplete callosotomy fails to control seizures effectively (Fig 5). In addition to traditional transectional surgery, corpus callosotomy can be accomplished through the use of laser ablative techniques. DTI may be used to confirm whether the intended disconnection has succeeded or to help in planning further interventions against residual connectivity.[17]

Multiple Subpial Transection. Multiple subpial transection (MST) is a surgical intervention that is reserved for patients with medically refractory seizures with epileptogenic foci that overlap the functional (language or motor/sensory) cortex and are therefore not amenable to resection. However, since the approval of neuromodulatory devices such as responsive neurostimulation (RNS), MST is rarely indicated. MST involves severing horizontally oriented cortical fibers of >5 mm to disrupt the putative epileptic discharge, while preserving vertically oriented fibers, perpendicular blood vessels, as well as functional activity.[18] For example, MST for a precentral gyrus epileptogenic site would ideally diminish cortical seizure spread while preserving corticospinal fiber function. Taking advantage of the columnar architecture of the cerebral cortex, the surgical technique is performed as a cluster of small radial cuts, applied through a small pial puncture at the entry point. MST is usually performed as an adjunct to adjacent resective surgery outside the confines of indispensable cortex, though in one-quarter of cases MST has been performed as a stand-alone treatment.[10] By applying MST to functional brain surrounding the resection of most of an epileptogenic region, the surgeon can minimize the amount of excised cortex and avoid neurologic loss.

MR imaging findings depend on the timing of imaging from the intervention, with initial findings of acute hemorrhagic changes, early vasogenic edema, and fluid pocket formation at the transected cortex and varying degrees of gyral atrophy at later stages (Online Supplemental Data). In their retrospective study of 10 cases, Finet et al,[19] reported that T2-weighted images are the most sensitive sequence in the imaging of MST findings, with no gyral atrophy or gliosis, while Smith[20] reported cystic changes and gyral atrophy. Cortical atrophy is presumably the result of vascular injury during the procedure, which depends on the extent and depth of the procedure. Signal changes follow the typical course of traumatic brain injury: an initial hemorrhagic area (hyperintense on T1WI), followed by early vasogenic edema (hyperintense on T2WI) and small pockets of fluid signals mid- to long-term (hyperintense on T2WI) and, eventually, varying degrees of gyral atrophy and gliosis. Blood-sensitive MR imaging sequences like SWI and gradient recalled-echo imaging can precisely delineate transection lines due to hemorrhage and hemosiderin, which can persist for years. On imaging, patients with findings of MST are most likely to have other concomitant findings as a result of presurgical intracranial electrode studies and/or adjacent resective surgery.

Laser Ablative Surgery. Laser interstitial thermal therapy (LITT) is an emerging minimally invasive surgical intervention for medically intractable epilepsy, demonstrating satisfactory efficacy for both medial temporal lobe structures and lesions. LITT effectiveness for seizures from MTS is comparable with that of conventional ATL for the short term.[21] Cavernous malformation, heterotopia, hypothalamic hamartoma, and corpus callosotomy are the more recent applications of LITT in epilepsy treatment. LITT is particularly important for treatment of epileptogenic lesions or abnormalities in the dominant hemisphere, to avoid resective surgery and preserve language. There may be critical clinical implications associated with the specific structures that are targeted; therefore, it is imperative that the location and volume of ablation be detailed in the operative and postoperative imaging reports.[21]

Characteristic time-dependent imaging findings are associated with LITT on MR imaging.[22] During the early periods (first 2 weeks to 3 months) after the ablative procedure, LITT has a typical target-like appearance with 4 concentric zones, including a peripheral rim of enhancement (Fig 6).

Catheter: In the bull's eye of the ablation region, the laser catheter is imaged as a focal signal void on all sequences. With removal of the catheter, CSF fills the catheter track.

Central zone: T1 hyperintensity and T2 hypointensity from coagulative necrosis are classically associated with methemoglobin, high protein, and engorged blood vessels. Heat likely accelerates conversion of hemoglobin to methemoglobin, which leaks from damaged erythrocytes.[22]
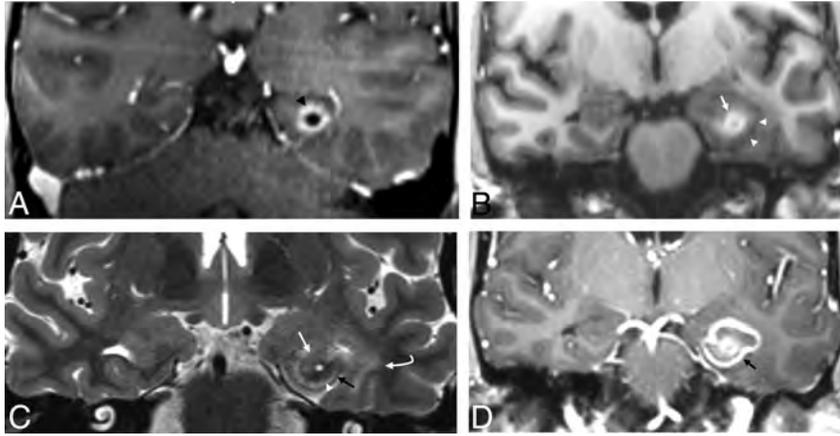
FIG 6. LITT, coronal imaging. Left hippocampal LITT performed in a patient with intractable epilepsy arising from left MTS because of left language dominance. Immediate post-LITT contrast MR imaging (*A*) shows a catheter (as a *central black dot*) with surrounding enhancement (*black arrowhead*) on postcontrast T1W1. The ablated region has enlarged 3 weeks later (*B–D*). The central zone of coagulative necrosis (*white arrow*) is hyperintense on precontrast T1WI (*B*) and hypointense on T2WI (*C*) signal. These are surrounded by a peripheral zone of necrotizing edema (*white arrowhead*), which is hypointense on T1WI and hyperintense on T2WI. The peripheral zone is delineated by a rim (*black arrow*) of signal void on T2WI that enhances on postcontrast imaging (*D*), defining the ablated area. Vasogenic edema (*curved white arrow*) can be seen around the ablated region (hyperintense on T2WI and hypointense on precontrast T1WI). Note the focal central CSF intensity signal on *B–D* from fluid in the laser catheter track.

Peripheral zone: A peripheral T1-hypointense/T2-hyperintense zone surrounding the central zone represents necrotizing edema. This zone is irreversibly damaged and is not viable. Histopathologically, there is an inflammatory reaction, edema, an empty appearance of nerve cell processes and astrocytic foot processes, macrophages, and lymphocytes; this zone liquefies within few months.[23]

Outer rim of the peripheral zone: T1-hypointense/hyperintense and T2-hypointense ring bordering the peripheral zone is considered to represent damaged blood vessels with blood-brain barrier disruption and granulation tissue. The ring persistently enhances with gadolinium on postcontrast MR images and delineates the border of nonviable tissue. It is considered to represent deoxyhemoglobin initially, which evolves into hemosiderin deposition.[22]

Vasogenic edema: T1-hypointense and T2-hyperintense regions of reversible vasogenic edema surround the hemosiderin/enhancing ring. This area appears days to weeks after the procedure, initially expands, and gradually shrinks after 2–9 weeks, with no overt residual tissue damage.

The ablation region (central and peripheral zones) initially increases during the first 2 weeks and then shrinks during the next 6 months. Enhancement may extend along the track of the laser catheter as well as in the peripheral rim. In the subacute period, the central zone T1 hyperintensity decreases, while the peripheral T1 hypointensity increases, leading to a more uniform appearance. The diameter of the enhancing rim decreases exponentially after mild initial increase, with a half-life of 93 days, but a focal area of enhancement can be seen for years.[22] As stated above, vasogenic edema generally dissipates within 1–2 months. However, this idealized pattern may vary, particularly in the center of the lesion, as

the lesion can be heterogeneous. However, a T1-hyperintense central zone, an enhancing T2 hypointense peripheral ring, and perilesional white matter edema will always be present.

Radiosurgery. The role, efficacy, and safety of radiosurgery in epilepsy treatment is not well-understood, and its application in temporal lobe epilepsy remains controversial. The antiepileptic effects of radiosurgery were initially observed with treatments of cavernomas and AVMs, in which the primary treatment end point was lesion size reduction to prevent hemorrhage, rather than antiepileptic efficacy. For MTS, there is no consensus on the optimal dosing of radiation, and long-term procedural outcomes remain largely unknown. One prospective randomized study found superior efficacy of a higher radiation dose (24 versus 20 Gy), though the antiepileptic effects still remained modest,[24] while another demonstrated ineffectiveness due to resultant radiation necrosis necessitating resective surgery.[25] Imaging characteristics of the radiation-related parenchymal changes parallel typical brain radiation exposure with initial white matter signal alteration followed by edema and then radiation necrosis, typified by lace-like, irregular peripheral enhancement and central nonenhancing necrosis. Solid nodular enhancement may be seen with lesions of <2 cm.[26] Cyst formation may occur in the mid- to long-term. Eventually, parenchymal involution, gliosis, and encephalomalacia occur months to years after treatment. Optimal imaging follow-up for irradiated MTS depends on the evolution of clinical symptoms; worsening focal seizures may precede edema and mass effect.[24]

Neuromodulation
As opposed to pharmacotherapy, which targets the entire brain, or surgical treatment, which consists of resection, disconnection, or ablation of certain anatomic regions, neuromodulation uses electrical stimulation to target specific regions of the brain to disrupt and modulate epileptic activity or networks. Currently, there are 3 main types of microinvasive neuromodulation methods that have FDA approval for the treatment of medically refractory epilepsy: vagus nerve stimulation (VNS), deep brain stimulation (DBS), and RNS. Other neuromodulation techniques such as transcranial magnetization, transcranial electrical stimulation, and focused sonography are under investigation.[27] Familiarity with neuromodulation techniques, associated hardware, expected procedural outcomes, and their temporal evolution may prevent imaging pitfalls. While the effectiveness of seizure control by current neuromodulation devices is lower than resective surgery, these devices have a specified role in the treatment of intractable epilepsy (Fig 1). Implanted intracranial and extracranial
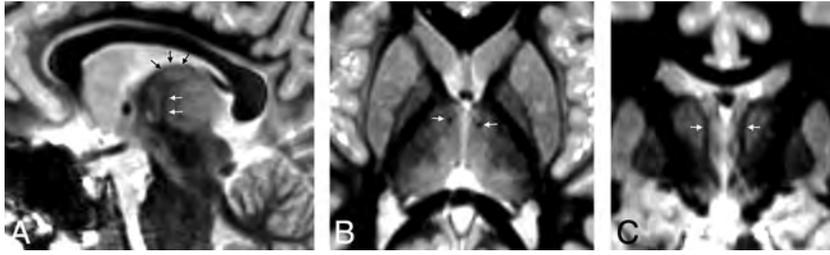
FIG 7. FGATIR, ANT localization. An 11-year-old child needing DBS because of bilateral poorly localized focal epilepsy despite left hippocampal sclerosis on MR imaging. ANT (*black arrows*) can be localized superior to the mammillothalamic track (*white arrows*), delineated on FGATIR parasagittal (*A*), axial (*B*), and coronal (*C*) images. The FGATIR sequence nulls white matter in this track.

components should be compared with baseline postprocedural scans, particularly for location verification, and device migration or failure. The course of leads and extensions should be scrutinized for any structural signal change or fluid collection, particularly in case of suspected infections.

Vagus Nerve Stimulation. VNS was the first neuromodulation device used for intractable epilepsy, with FDA approval in 1997 as an adjunctive treatment for pharmacoresistant focal epilepsy. This technique is often used in patients with unidentifiable seizure foci or with multiple potential epileptogenic foci not amenable to interventional treatment. VNS is an extracranial procedure, with the infraclavicular subcutaneous implantation of a commercially available programmable pulse-generator device. Lead wires are attached to (and surround) the midcervical vagus nerve through a second neck incision via a subcutaneous tunnel (Online Supplemental Data). The generator runs continuously, but its activity can be controlled by the patient or programmer. VNS is placed on the left side to prevent damage to the sinoatrial node, which is innervated by the right vagus nerve. Patients with an implanted VNS device can undergo MR imaging using a transmit-receive head coil without a limit of active scanning time.[28] For receive-only coils, specific conditions are applied for newer models, such as specific absorption rate restriction or a time limit for "safe active scanning," as little as 15 minutes, depending on the type and brand of the specific devices (VNS and/or DBS). On imaging, the position of the device, including wires, should be investigated. A potential complication of VNS is temporary or permanent left vocal cord paralysis. Other local or device-related complications are akin to those of neuromodulation systems in general, which are discussed in the DBS section below.

Responsive Neurostimulation. RNS has been an approved treatment for focal intractable epilepsy since 2013. It is indicated for a localized epileptogenic focus, in which there is a significant risk of a neurologic deficit associated with resection or ablation or for patients unwilling to undergo invasive therapies. As opposed to VNS and DBS, RNS has both recording and stimulation capability, with feedback similar to that of a cardiac pacemaker or defibrillator. Once abnormal electrocorticographic activities are detected, the RNS device (NeuroPace Inc., Mountain-view, CA, USA) will deliver electrical current pulses to disrupt and abort the ictal activity.[29] RNS is a closed-loop system, which means that it delivers electrical impulses only after the targeted epileptiform activity is detected. A programmable neurostimulator device is implanted in the skull and connected to 2 subdural strips and/or depth electrodes (with 4 electrodes each) that are placed at or near a previously determined epileptogenic zone (Online Supplemental Data). Skull x-rays are helpful in demonstrating the position of the device. CT is typically used for determination of intracranial device position as well as for potential surgical complications, but severe beam and streak artifacts limit evaluation. Angle cuts or dual-energy CT may provide further information in challenging cases. In March 2020, the FDA approved a commercial RNS device for MR imaging compatibility under appropriate conditions.

Deep Brain Stimulation. The discovery that damage to the anterior thalamic nucleus (ANT) in Rhesus monkeys decreased epileptogenic activity, likely by increasing the threshold of cortical epileptic discharges, led to the possibility of a new treatment technique, DBS.[30] The ANT is part of the circuit of Papez, which relays information among the hippocampus, mamillary bodies, cingulum, and fornices.[31] The Stimulation of the Anterior Nucleus of the Thalamus for Epilepsy (SANTE) trial and others demonstrated the safety and efficacy of ANT stimulation in pharmacoresistant focal epilepsy with reduction in seizures with variable results.[32-34] The FDA recently approved commercially available MR imaging–conditional DBS systems for epilepsy treatment, which allow preprogramed intermittent stimulation of bilateral ANTs. Apart from the ANT, neurostimulation studies were also performed in the centromedian thalamus, subthalamic nucleus, caudate nucleus, hippocampus, hypothalamus, cortex, and cerebellum, which resulted in variable effects on seizures. For example, DBS targeting the centromedian thalamic nuclei has been shown to be effective in generalized epilepsy[35] due to circuitry connections to the medial frontal and cingulate gyri.[36] Further understanding of epilepsy as a network disorder and new large-scale studies will likely broaden the clinical landscape of neuromodulation.

Although variations may occur in individual approaches to ANT stimulation across different institutions, the placement of electrodes is commonly performed in a transventricular fashion, which is connected to a subcutaneous generator (Online Supplemental Data). Electrode placement uses preoperative coregistered high-resolution MR imaging and CT scans for frame-based stereotactic neuronavigation, with or without robotic assistance, and electrodes directed toward the anterior superior portion of the ANT. Before the report of Grewal et al,[37] in 2018, describing direct targeting of the ANT, most surgeons used indirect targeting methods (eg, based on an atlas coregistered to the patient's anterior/posterior commissure line), which introduced some variability in electrode placement (including 10% being placed outside the ANT). Grewal et al used
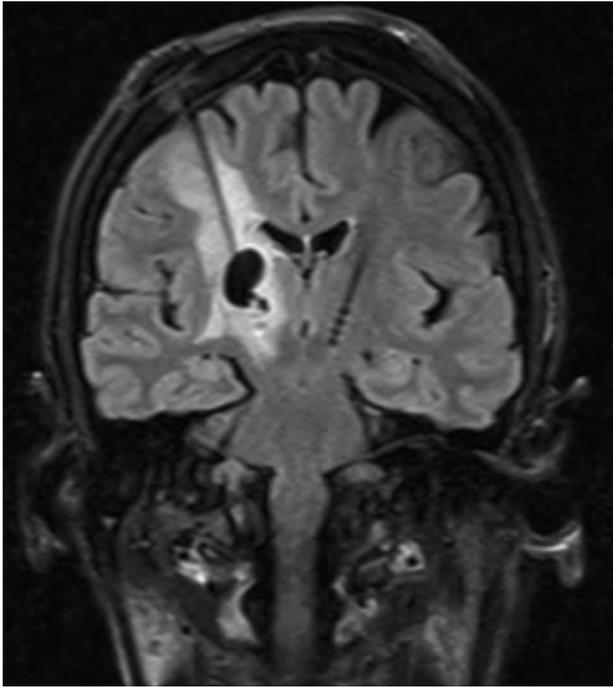
FIG 8. DBS complications in a patient without epilepsy. Left foot weakness and dysarthria occurred 3 months after bilateral DBS for refractory essential tremor. Coronal T2 FLAIR image demonstrates a nonhemorrhagic cyst with vasogenic edema surrounding the right electrode and no restricted diffusion or enhancement to suggest infection. This rare inflammatory DBS electrode complication was treated with steroids, resulting in regression of the cyst and vasogenic edema 7 weeks later (not shown here).

an 8-minute 0.8-mm isotropic inversion recovery sequence (known as fast gray matter acquisition T1 inversion recovery or fast gray matter acquisition [FGATIR]), inorder to null the white matter to delineate the mammillothalamic track, in turn, allowing direct localization of the anterior nucleus of the thalamus situated superiorly (Fig 7). The mammillothalamic track is easily delineated on parasagittal FGATIR images, as opposed to 3D T1 gradient-echo imaging, which does not have sufficient contrast. MR or CT angiography can aid in the determination of vascular anatomy, for a safe trajectory.

The clinical effectiveness of DBS is primarily determined by the precision of the electrode implantation within the ANT and the burden of adverse effects. A long-term follow-up report from the SANTE trial demonstrated a procedural infectious complication rate of 12.7%; however, none were intracranial.[32] The neurostimulator migration rate was 5.5%. Most other complications were neuropsychiatric disturbances. Delayed complications of DBS systems also include skin granuloma, bowstringing, extension lead failure, hardware migration, and perilead cyst formation and brain edema (Fig 8).

CONCLUSIONS

Pharmacoresistant epilepsy has a wide range of surgical interventional treatment options from microinvasive neuromodulation techniques to the functional hemispherectomy, which drastically alters cranial anatomy. As treatment options constantly evolve,

practitioners need to be cognizant of unique imaging findings and potential short- and long-term complications associated with these interventions.

REFERENCES

1. Zentner J. **Surgical Treatment of Epilepsies.** In: *Advances in Epilepsy Surgery and Radiosurgery.* Springer, Vienna; 2002:27–35
2. Spencer DD, Gerrard JL, Zaveri HP. **The roles of surgery and technology in understanding focal epilepsy and its comorbidities.** *Lancet Neurol* 2018;17:373–82 CrossRef Medline
3. Spencer SS. **Neural networks in human epilepsy: evidence of and implications for treatment.** *Epilepsia* 2002;43:219–27 CrossRef Medline
4. Davis LM, Spencer DD, Spencer SS, et al. **MR imaging of implanted depth and subdural electrodes: is it safe?** *Epilepsy Res* 1999;35:95–98 CrossRef Medline
5. Merriam MA, Bronen RA, Spencer DD, et al. **MR findings after depth electrode implantation for medically refractory epilepsy.** *AJNR Am J Neuroradiol* 1993;14:1343–46 Medline
6. Nguyen HS, Doan N, Gelsomino M, et al. **Dilemmas surrounding the diagnosis of deep brain stimulation electrode infection without associated wound complications: a series of two cases.** *Surg Neurol Int* 2016;7:121–24 CrossRef Medline
7. Wiebe S, Blume WT, Girvin JP, et al. **A randomized, controlled trial of surgery for temporal-lobe epilepsy.** *N Engl J Med* 2001;345:311–18 CrossRef Medline
8. Mansouri A. **Surgery for drug-resistant epilepsy in children.** *N Engl J Med* 2018;378:399 CrossRef Medline
9. Noe K, Sulc V, Wong-Kisiel L, et al. **Long-term outcomes after nonlesional extratemporal lobe epilepsy surgery.** *JAMA Neurol* 2013;70:1003–08 CrossRef Medline
10. Schramm J, Clusmann H. **The surgery of epilepsy.** *Neurosurgery* 2008;62(Suppl 2):463–81; discussion 481 CrossRef Medline
11. Brotis AG, Giannis T, Kapsalaki E, et al. **Complications after anterior temporal lobectomy for medically intractable epilepsy: a systematic review and meta-analysis.** *Stereotact Funct Neurosurg* 2019;97:69–82 CrossRef Medline
12. Sato N, Bronen RA, Sze G, et al. **Postoperative changes in the brain: MR imaging findings in patients without neoplasms.** *Radiology* 1997;204:839–46 CrossRef Medline
13. Oztürk A, Oğuz KK, Akalan N, et al. **Evaluation of parenchymal changes at the operation site with early postoperative brain diffusion-weighted magnetic resonance imaging.** *Diagn Interv Radiol* 2006;12:115–20 Medline
14. Saluja S, Sato N, Kawamura Y, et al. **Choroid plexus changes after temporal lobectomy.** *AJNR Am J Neuroradiol* 2000;21:1650–53 Medline
15. Delev D, Oehl B, Steinhoff BJ, et al. **Surgical treatment of extratemporal epilepsy: results and prognostic factors.** *Neurosurgery* 2019;84:242–52 CrossRef Medline
16. Graham D, Tisdall MM, Gill D. **Corpus callosotomy outcomes in pediatric patients: a systematic review.** *Epilepsia* 2016;57:1053–68 CrossRef Medline

17. Choudhri AF, Whitehead MT, McGregor AL, et al. **Diffusion tensor imaging to evaluate commissural disconnection after corpus callosotomy.** *Neuroradiology* 2013;55:1397–1403 CrossRef Medline

18. Morrell F, Whisler WW, Bleck TP. **Multiple subpial transection: a new approach to the surgical treatment of focal epilepsy.** *J Neurosurg* 1989;70:231–39 CrossRef Medline

19. Finet P, Grandin C, Vaz G, et al. **Multiple subpial transections and magnetic resonance imaging.** *Neurochirurgie* 2017;63:449–52 CrossRef Medline

20. Smith MC. **Multiple subpial transection in patients with extratemporal epilepsy.** *Epilepsia* 1998;39:S81–89 CrossRef Medline

21. Wu C, Jermakowicz WJ, Chakravorti S, et al. **Effects of surgical targeting in laser interstitial thermal therapy for mesial temporal lobe epilepsy: a multicenter study of 234 patients.** *Epilepsia* 2019;60:1171–83 CrossRef Medline

22. Schwabe B, Kahn T, Harth T, et al. **Laser-induced thermal lesions in the human brain: short- and long-term appearance on MRI.** *J Comput Assist Tomogr* 1997;21:818–25 CrossRef Medline

23. Schober R, Bettag M, Sabel M, et al. **Fine structure of zonal changes in experimental Nd:YAG laser-induced interstitial hyperthermia.** *Lasers Surg Med* 1993;13:234–41 CrossRef Medline

24. Barbaro NM, Quigg M, Broshek DK, et al. **A multicenter, prospective pilot study of gamma knife radiosurgery for mesial temporal lobe epilepsy: seizure response, adverse events, and verbal memory.** *Ann Neurol* 2009;65:167–75 CrossRef Medline

25. Chen N, Du SQ, Yan N, et al. **Delayed complications after Gamma Knife surgery for intractable epilepsy.** *J Clin Neurosci* 2014;21:1525–28 CrossRef Medline

26. Wang YX, King AD, Zhou H, et al. **Evolution of radiation-induced brain injury: MR imaging-based study.** *Radiology* 2010;254:210–18 CrossRef Medline

27. Stern JM, Spivak NM, Becerra SA, et al. **Safety of focused ultrasound neuromodulation in humans with temporal lobe epilepsy.** *Brain Stimulation* 2021;14:1022–31 CrossRef

28. de Jonge JC, Melis GI, Gebbink TA, et al. **Safety of a dedicated brain MRI protocol in patients with a vagus nerve stimulator.** *Epilepsia* 2014;55:e112–15 CrossRef Medline

29. Bergey GK, Morrell MJ, Mizrahi EM, et al. **Long-term treatment with responsive brain stimulation in adults with refractory partial seizures.** *Neurology* 2015;84:810–17 CrossRef Medline

30. Bouwens van der Vlis TA, Schijns O, Schaper F, et al. **Deep brain stimulation of the anterior nucleus of the thalamus for drug-resistant epilepsy.** *Neurosurg Rev* 2019;42:287–96 CrossRef Medline

31. Dalkilic EB. **Neurostimulation devices used in treatment of epilepsy.** *Curr Treat Options Neurol* 2017;19:7 CrossRef Medline

32. Salanova V, Witt T, Worth R, et al. SANTE Study Group. **Long-term efficacy and safety of thalamic stimulation for drug-resistant partial epilepsy.** *Neurology* 2015;84:1017–25 CrossRef Medline

33. Lee KJ, Shon YM, Cho CB. **Long-term outcome of anterior thalamic nucleus stimulation for intractable epilepsy.** *Stereotact Funct Neurosurg* 2012;90:379–85 CrossRef Medline

34. Krishna V, King NK, Sammartino F, et al. **Anterior nucleus deep brain stimulation for refractory epilepsy: insights into patterns of seizure control and efficacious target.** *Neurosurgery* 2016;78:802–11 CrossRef Medline

35. Valentín A, García Navarrete E, Chelvarajah R, et al. **Deep brain stimulation of the centromedian thalamic nucleus for the treatment of generalized and frontal epilepsies.** *Epilepsia* 2013;54:1823–33 CrossRef Medline

36. Kim SH, Lim SC, Yang DW, et al. **Thalamo-cortical network underlying deep brain stimulation of centromedian thalamic nuclei in intractable epilepsy: a multimodal imaging analysis.** *Neuropsychiatr Dis Treat* 2017;13:2607–19 CrossRef Medline

37. Grewal SS, Middlebrooks EH, Kaufmann TJ, et al. **Fast gray matter acquisition T1 inversion recovery MRI to delineate the mammillothalamic tract for preoperative direct targeting of the anterior nucleus of the thalamus for deep brain stimulation in epilepsy.** *Neurosurg Focus* 2018;45:E6 CrossRef Medline

Calcified Pseudoneoplasm of the Neuraxis

J.C. Benson, J. Trejo-Lopez, J. Boland-Froemming, B. Pollock, C.H. Hunt, and J.T. Wald

ABSTRACT

SUMMARY: Calcified pseudoneoplasms of the neuraxis are extremely rare non-neoplastic lesions that can exist anywhere in the CNS. Although benign, the lesions can cause substantial neurologic symptoms, typically related to mass effect on adjacent structures. Calcified pseudoneoplasms of the neuraxis can also mimic other entities such as calcified oligodendrogliomas and meningiomas. Nevertheless, the lesions can usually be strongly suggested at the time of imaging due to a number of fairly unique imaging characteristics. Here, the clinical presentation of a patient with a posterior fossa calcified pseudoneoplasm of the neuraxis is described, along with its imaging and pathologic features.

ABBREVIATION: CAPNON = calcified pseudoneoplasm of the neuraxis

A 58-year-old man presented with a long-standing history of headaches. As a child, his symptoms were attributed to temporomandibular joint dysfunction; migraines were common during adulthood. Twenty-five years before presentation, he underwent CT of his head after a traumatic injury, which revealed a posterior fossa abnormality described as being "scar tissue," per the patient (Fig 1). An MR imaging performed approximately 2 decades later showed a highly calcified lesion that appeared stable compared with the prior CT. The patient ultimately presented to our institution with progressively worsening headaches that were not successfully managed with over-the-counter medications.

Imaging

Imaging revealed a densely calcified extra-axial mass in the posterior fossa, which appeared to originate from the tentorium. The mass had mildly increased in size since the initial CT. MR imaging demonstrated markedly hypointense intralesional signal on T2 and FLAIR. Modest internal enhancement was observed, which was superimposed on heterogeneous intrinsic T1 hyperintensity. More substantial enhancement was also observed around the margins of the lesion. Mild associated mass effect and vasogenic edema were present, with partial effacement of the fourth

ventricle. There was also a suggestion of a dural tail, leading to the favored diagnosis of calcified meningioma (Fig 2).

The decision was made to proceed to surgical resection on the basis of the presumption that the mass was contributing to the patient's ongoing headaches.

Operative Report and Hospital Course

The mass was resected via a right suboccipital craniotomy. When the surgeon entered the dura, substantial mass effect was observed in the posterior fossa. The tumor was visualized along the superior surface of the cerebellum and was adherent to the tentorium. Because of its solid, chalky consistency, the tumor was removed in a piecemeal fashion. No gross tumor was visible at the completion of the surgery. Postoperatively, the patient developed obstructive hydrocephalus and required placement of an external ventricular drain but did not require permanent CSF diversion. He was discharged to inpatient rehabilitation and made a progressive recovery back to his neurologic baseline.

Pathology

Histologic examination revealed a coarsely calcified mass, in which a chondromyxoid fibrillary matrix was identifiable (Fig 3). While the periphery of the tumor occasionally demonstrated palisading reactive spindle cells and a mild chronic inflammatory infiltrate, the lesion itself was composed of acellular mineralization devoid of a viable cytologic component. The lesion was, therefore, incompatible with a neoplastic entity such as a heavily calcified glioma, which is characterized by a discernible proliferation of cells.

FIG 1. Initial CT scan of the mass obtained 15 years before presentation at our institution. The images demonstrate a densely calcified mass in the right posterior fossa (*asterisk*), without substantial local mass effect.
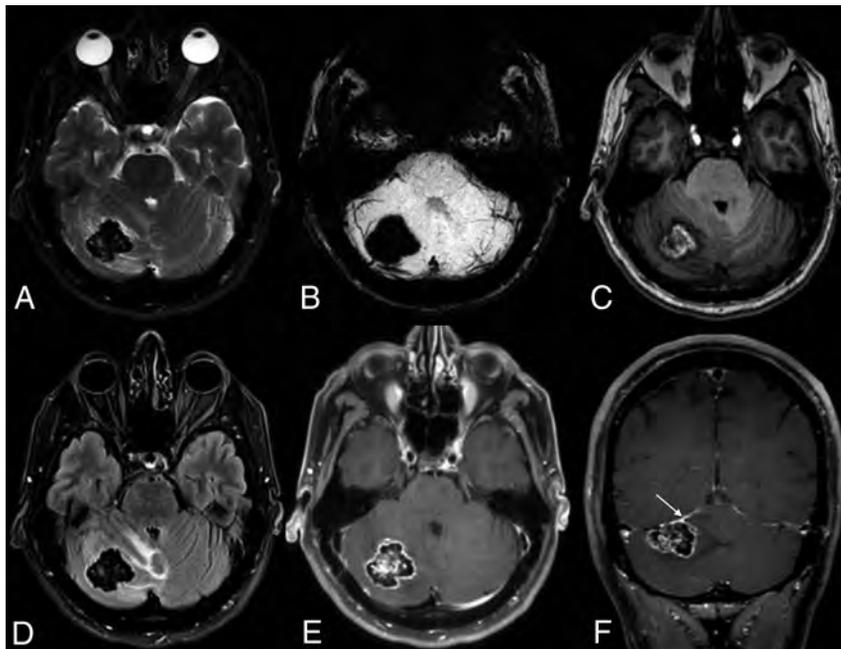
FIG 2. MR imaging at the time of presentation demonstrates a slight interval increase in the size of the mass. Dense intralesional calcifications are again observed, as evidenced by substantial signal drop-out on T2 (*A*) and susceptibility-weighted (*B*) images. Axial T1 image (*C*) shows areas of internal hyperintense signal. FLAIR image (*D*) shows adjacent vasogenic edema, with a mild associated mass effect on the fourth ventricle. Axial (*E*) and coronal (*F*) postcontrast images demonstrate intense peripheral enhancement around the perimeter of the lesion, as well as a possible dural tail (*arrow*).

Multiple features distinguished this entity from other intracranial masses with a chondromyxoid matrix. Metaplastic meningiomas demonstrate varying degrees of mesenchymal elements that may manifest as myxoid changes, and chordoid meningiomas feature chords and trabeculae of eosinophilic, variably vacuolated cells in a mucoid background reminiscent of chordomas.[1,2] Nevertheless, both meningioma variants invariably demonstrate an overt neoplastic cell population, readily differentiating them from the entity observed in this case. In addition, masses exhibiting chondromyxoid elements should be evaluated for the possibility of chondroma and chondrosarcoma. Conventional chondrosarcomas exhibit variable degrees of cellularity, cytologic atypia, and mitotic activity and, similarly, show a distinctive hyaline-to-myxoid cartilage matrix.[3] However, the presence of a malignant cytologic component differentiates chondrosarcomas from the current process. The mature hyaline cartilage present in chondromas appears morphologically distinct from the irregular pattern of a chondromyxoid matrix and calcification seen in this lesion.[4]

In addition, some processes exhibit robust mineralization in the absence of an overt neoplastic cell population, including calcium pyrophosphate disease and myositis ossificans. While calcium pyrophosphate disease may exhibit robust mineralization, it is characteristically composed of rhomboid pyrophosphate crystals that exhibit weak, positive birefringence following examination under polarized light, a quality lacking in the current lesion. Moreover, although myositis ossificans often exhibits prominent calcification, it is, nevertheless, histologically distinctive for having a so-called "zonation" phenomenon, in which progressive maturation of bone is deposited around a central fibroblastic core—features lacking in this process.[5]

Ultimately, the robust calcification in patchy alternation with the chondromyxoid fibrillary matrix seen in this lesion was most consistent with calcified pseudoneoplasm of the neuraxis (CAPNON). The diagnosis was CAPNON.

DISCUSSION

CAPNONs, also known as fibro-osseous lesions of the CNS, are rare nonneoplastic lesions, first described by Rhodes and Davis,[6] in 1978. They are found essentially everywhere in the central nervous system; lesions can be intra- or extra-axial and occur in both the brain and spine.[7,8] However, CAPNONs may have a slight predilection for intracranial involvement, specifically occurring along the dura of the skull base.[9] The lesions have no predilection for age or sex. Cases have been reported in patients as young as 12 and as old as 90 years of age, with about half of patients presenting between 40
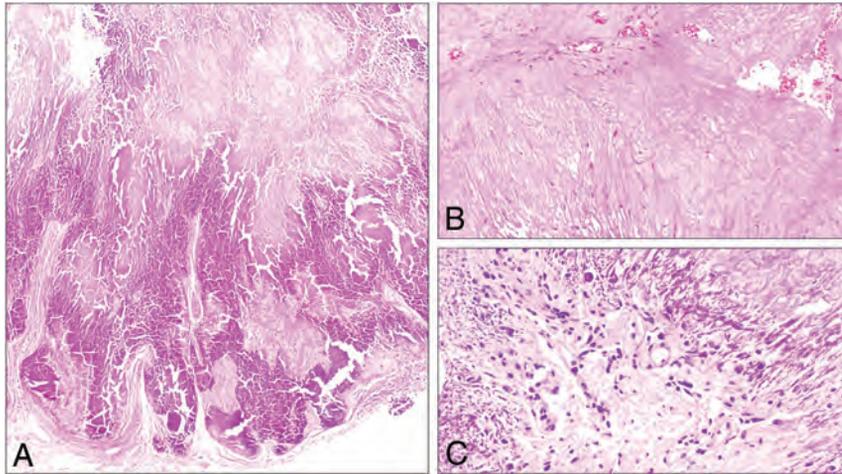
FIG 3. Histologic examination revealed a tumefactive lesion characterized by exuberant calcification (*A*, original magnification ×50) around areas of chondromyxoid fibrillary matrix (*B*, original magnification ×20). Palisading spindle cells and a mild chronic inflammatory infiltrate are present on the periphery (*C*, original magnification ×200). All specimens are stained with hematoxylin-eosin.

and 60 years of age.[9,10] Because symptoms are related to local mass effect, clinical presentations are extremely variable. Headaches, epileptic seizures, cranial nerve dysfunction, and neck/back pain have all been reported; some CAPNONs are discovered incidentally.[11-13]

The diagnosis of a CAPNON can be strongly suggested by its imaging appearance, especially when intracranial. True to their name, CAPNONs are highly calcific, a feature that is readily apparent on CT. Although lesions are classically described as hypointense on both T1WI and T2WI, T1 hypointensity is not a reliable characteristic, likely due to interactions between calcification and free water molecules, as was seen in the case presented here.[14] CAPNONs are usually extra-axial lesions, and dural attachments are observed in most cases. Modest enhancement may or may not be visualized, with either intralesional serpentine or peripheral/rim enhancement patterns.[15] Perilesional T2 prolongation, too, is sometimes observed, either representing vasogenic edema or gliosis related to chronic mass effect on the nearby parenchyma.[16,17]

The imaging differential of CAPNONs mirrors that of other cerebral calculi ("brain stones"). If extra-axial, osteomas and highly calcified meningiomas are the greatest mimickers of CAPNONs.[18] Osteomas often develop from the calvaria; thus, intracranial osteomas are often contiguous with the inner table or even the diploic space of the adjacent bone. They also typically lack intracranial enhancement.[19] Nevertheless, osteomas can arise from dural attachments, similar to CAPNONs; rarely, some osteomas enhance.[20] Large meningiomas are often associated with hyperostosis of the underlying bone, a feature not described in CAPNONs. Meningiomas also usually demonstrate enhancement, though this may be absent in densely calcified lesions. Intra-axial CAPNONs can appear similar to calcified tumors and vascular malformation—particularly oligodendrogliomas and cavernous malformations.[21,22] Dystrophic calcifications, too, can share imaging features of CAPNONs, though these typically arise in post-traumatic or postischemic regions of the brain.[23]

Spinal lesions tend to be smaller and less characterizable than their intracranial counterparts. The large majority of spinal CAPNONs (>80%) are in the epidural space.[10] Garcia Duque et al[10] found that the cervical region was the most frequent location of spinal CAPNONs. Other studies, however, have suggested otherwise; a large single-institution study by Ho et al[14] noted that most spinal CAPNONs were in the lumbar spine. Calcified psammomatous meningiomas, calcium pyrophosphate dihydrate deposition disease, exophytic osseous lesions, synovial cysts, calcified hematomas, and herniated discs would be the most likely alternative diagnostic considerations.[24]

Histologically, CAPNON is typically characterized as a process exhibiting robust calcification in a patchy, nonzonal alternation with chondromyxoid fibrillary matrix. While a chondromyxoid fibrillary matrix has been classically described as a defining feature of this entity, a recent case series has noted that calcification and cartilaginous elements are more consistent features across patients.[14] While the lesion may incite reactive meningothelial cells and macrophages around the periphery of the lesion, these cellular elements should not be an underlying substrate of the lesion itself.

The etiology of CAPNONs remains uncertain. However, both their histologic characteristics and behavior have led to the presumption that CAPNONs are reactive entities.[18] The source of the reactive process can vary, ranging from inflammation to trauma. Ho et al[14] found that nearly one-third of CAPNONs were so-called "collision" lesions, in which CAPNON tissue coexisted with a separate, distinct pathologic entity. The coexistent primary tumors found in such collision lesions included meningiomas, pleomorphic xanthoastrocytomas, dysembryoplastic neuroepithelial tumors, and vascular malformations. Dual lesions have similarly been reported elsewhere.[8,25] Histologic analyses of collision lesions have demonstrated inflammatory tissue outside the CAPNON boundaries, suggesting that the primary lesions initiated a reactive process, ultimately resulting in the formation of a CAPNON.[14] In the spine, too, associated pathologies are also observed, including synovial cysts and degenerated discs.

The management and prognosis of CAPNONs are highly dependent on lesion location. Total resection tends to be curative, though exceptionally rare cases have been reported in which recurrence or progression of lesions or both occur after incomplete resection.[11] Still, surgical removal can be particularly challenging when the lesion is located within eloquent regions because piecemeal resection is often required due to the dense intralesional calcifications.[26]

Case Summary

CAPNONs are non-neoplastic lesions, found anywhere within the CNS and can appear as intra- or extra-axial masses. Although

extremely rare, the diagnosis of a CAPNON can be highly suggested on the basis of its imaging appearance, particularly when large and intracranial. Histologically, a CAPNON is typically characterized as a process exhibiting robust calcification with cartilaginous elements and may exhibit a characteristic chondromyxoid fibrillary matrix. Prognosis tends to be excellent, and surgical resection is curative.

REFERENCES

1. Krisht KM, Altay T, Couldwell WT. **Myxoid meningioma: a rare metaplastic meningioma variant in a patient presenting with intratumoral hemorrhage.** *J Neurosurg* 2012;116:861–65 CrossRef Medline

2. Couce ME, Aker FV, Scheithauer BW. **Chordoid meningioma: a clinicopathologic study of 42 cases.** *Am J Surg Pathol* 2000;24:899–905 CrossRef Medline

3. Rosenberg AE, Nielsen GP, Keel SB, et al. **Chondrosarcoma of the base of the skull: a clinicopathologic study of 200 cases with emphasis on its distinction from chordoma.** *Am J Surg Pathol* 1999;23:1370–78 CrossRef Medline

4. Erdogan S, Zorludemir S, Erman T, et al. **Chondromas of the falx cerebri and dural convexity: report of two cases and review of the literature.** *J Neurooncol* 2006;80:21–25 CrossRef Medline

5. Lacout A, Jarraya M, Marcy P-Y, et al. **Myositis ossificans imaging: keys to successful diagnosis.** *Indian J Radiol Imaging* 2012;22:35–39 CrossRef Medline

6. Rhodes RH, Davis RL. **An unusual fibro-osseous component in intracranial lesions.** *Hum Pathol* 1978;9:309–19 CrossRef Medline

7. Grabowski M, Recinos P, Chen T, et al. **Calcifying pseudoneoplasm of the neuraxis overlying the corpus callosum: a case report and review of the literature.** *Clin Neuropathol* 2013;32:515–21 CrossRef Medline

8. Rodriguez FJ, Scheithauer BW, Fourney DR, et al. **Ependymoma and intraparenchymal calcifying pseudoneoplasm of the neural axis: incidental collision or unique reactive phenomenon?** *Acta Neuropathol* 2008;115:363–66 CrossRef Medline

9. Nussbaum ES, Hilton C, Defillo A, et al. **Extradural petromastoid calcifying pseudoneoplasm of the neuraxis (CAPNON): case report and literature review.** *Clin Neurol Neurosurg* 2018;166:99–106 CrossRef Medline

10. García Duque S, Medina Lopez D, Ortiz de Méndivil A, et al. **Calcifying pseudoneoplasms of the neuraxis: report on four cases and review of the literature.** *Clin Neurol Neurosurg* 2016;143:116–20 CrossRef Medline

11. Stienen MN, Abdulazim A, Gautschi OP, et al. **Calcifying pseudoneoplasms of the neuraxis (CAPNON): clinical features and therapeutic options.** *Acta Neurochir (Wien)* 2013;155:9–17 CrossRef Medline

12. Conway KS, Jentzen J, Pratt D, et al. **Sudden death due to calcifying pseudoneoplasm of the neuraxis: a case report and a review of sudden death due to undiagnosed central nervous system mass lesions.** *Am J Forensic Med Pathol* 2020;41:70–74 CrossRef Medline

13. Lopes AJ, Brock RS, Martins TG, et al. **Intradural calcifying pseudoneoplasm of the neuraxis presenting as a cauda equina syndrome.** *Surg Neurol Int* 2016;7:S1102–05 CrossRef Medline

14. Ho ML, Eschbacher KL, Paolini MA, et al. **New insights into calcifying pseudoneoplasm of the neuraxis (CAPNON): a 20-year radiological-pathological study of 37 cases.** *Histopathology* 2020;76:1055–69 CrossRef Medline

15. Aiken AH, Akgun H, Tihan T, et al. **Calcifying pseudoneoplasms of the neuraxis: CT, MR imaging, and histologic features.** *AJNR Am J Neuroradiol* 2009;30:1256–60 CrossRef Medline

16. Shrier DA, Melville D, Millet D, et al. **Fibro-osseous lesions involving the brain: MRI.** *Neuroradiology* 1999;41:18–21 CrossRef Medline

17. Merola J, Jain A, Ziad F, et al. **Calcified pseudoneoplasm of the neuraxis (CAPNON): a lesson learnt from a rare entity.** *J Neurol Neurosci* 2016;7:2–4 CrossRef

18. Bertoni F, Unni KK, Dahlin DC, et al. **Calcifying pseudoneoplasms of the neural axis.** *J Neurosurg* 1990;72:42–48 CrossRef Medline

19. Colas L, Caron S, Cotten A. **Skull vault lesions: a review.** *AJR Am J Roentgenol* 2015;205:840–47 CrossRef Medline

20. Chen SM, Chuang CC, Toh CH, et al. **Solitary intracranial osteoma with attachment to the falx: a case report.** *World J Surg Oncol* 2013;11:221 CrossRef Medline

21. Jun C, Burdick B. **An unusual fibro-osseous lesion of the brain: case report.** *J Neurosurg* 1984;60:1308–11 CrossRef Medline

22. Saade C, Najem E, Asmar K, et al. **Intracranial calcifications on CT: an updated review.** *J Radiology Case Rep* 2019;13:1–18 CrossRef Medline

23. Celzo FG, Venstermans C, De Belder F, et al. **Brain stones revisited: between a rock and a hard place.** *Insights Imaging* 2013;4:625–35 CrossRef Medline

24. Liccardo G, Lunardi P, Menniti A, et al. **Calcifying pseudo-tumor of the spine: description of a case and review of the literature.** *Eur Spine J* 2003;12:548–51 CrossRef Medline

25. Paolini MA, Ho ML, Monahan HR, et al. **Supratentorial CAPNON associated with WHO grade II meningioma: a case report.** *Neuropathology* 2018;38:535–38 CrossRef Medline

26. Peker HO, Aydin I, Baskaya MK. **Calcifying pseudoneoplasm of the neuraxis (CAPNON) in lateral cerebellomedullary junction: clinical image with surgical video.** *World Neurosurg* 2018;115:206–07 CrossRef Medline

Qualifying Certainty in Radiology Reports through Deep Learning–Based Natural Language Processing

F. Liu, P. Zhou, S.J. Baccei, M.J. Masciocchi, N. Amornsiripanitch, C.I. Kiefe, and M.P. Rosen

ABSTRACT

BACKGROUND AND PURPOSE: Communication gaps exist between radiologists and referring physicians in conveying diagnostic certainty. We aimed to explore deep learning–based bidirectional contextual language models for automatically assessing diagnostic certainty expressed in the radiology reports to facilitate the precision of communication.

MATERIALS AND METHODS: We randomly sampled 594 head MR imaging reports from an academic medical center. We asked 3 board-certified radiologists to read sentences from the Impression section and assign each sentence 1 of the 4 certainty categories: "Non-Definitive," "Definitive-Mild," "Definitive-Strong," "Other." Using the annotated 2352 sentences, we developed and validated a natural language-processing system based on the start-of-the-art bidirectional encoder representations from transformers (BERT), which can capture contextual uncertainty semantics beyond the lexicon level. Finally, we evaluated 3 BERT variant models and reported standard metrics including sensitivity, specificity, and area under the curve.

RESULTS: A κ score of 0.74 was achieved for interannotator agreement on uncertainty interpretations among 3 radiologists. For the 3 BERT variant models, the biomedical variant (BioBERT) achieved the best macro-average area under the curve of 0.931 (compared with 0.928 for the BERT-base and 0.925 for the clinical variant [ClinicalBERT]) on the validation data. All 3 models yielded high macro-average specificity (93.13%–93.65%), while the BERT-base obtained the highest macro-average sensitivity of 79.46% (compared with 79.08% for BioBERT and 78.52% for ClinicalBERT). The BioBERT model showed great generalizability on the heldout test data with a macro-average sensitivity of 77.29%, specificity of 92.89%, and area under the curve of 0.93.

CONCLUSIONS: A deep transfer learning model can be developed to reliably assess the level of uncertainty communicated in a radiology report.

ABBREVIATIONS: AUC = area under the receiver operating characteristic curve; BERT = bidirectional encoder representations from transformers; BioBERT = biomedical variant of BERT; ClinicalBERT = clinical variant of BERT; NLP = natural language processing; QC-RAD = qualifying certainty in radiology reports

The American College of Radiology has stressed a critical need for "precision communication" in radiologic reports,[1,2] and clarity has also been recognized by referring physicians as a key quality metric of radiologic reports.[3] However, there are communication gaps when the referring physician may interpret the radiologist's textual expressions that convey diagnostic uncertainty as different from what was intended.[4-6]

A standardized lexicon for diagnostic certainty[7] has been proposed to address this challenge. However, using a restricted lexicon can only help with lexion-level "one word at a time" interpretation, while diagnostic uncertainty in the radiology report is typically context-dependent. It means that the same term may have different interpretations based on how differential diagnoses are reported in the context. For instance, "likely to be Arnold Chiari I malformation" indicates a mildly certain diagnosis, while "likely differential considerations include demyelinating/inflammatory processes" indicates uncertainty in the diagnosis. Additionally, standardized lexicons will not mitigate the problem of overusing "hedge" words when the diagnosis could be more certain.[8]

Natural language processing (NLP), an artificial intelligence technology that analyzes free texts to understand underlying semantics, has been widely applied to radiology research for

FIG 1. Overview of the QC-RAD system workflow.

Table 1: Diagnostic certainty of diagnosis in the Impression section of a radiology report—categories for annotation

Certainty Categories	Interpretation	Examples
Non-Definitive	Describing differential diagnoses without indicating any confidence or only findings without any diagnosis	"Less likely differential considerations include demyelinating/inflammatory processes"
Definitive-Strong	Describing discrete diagnostic findings without hedging words	"Stable right sphenoid intraosseous lipoma"
Definitive-Mild	Describing discrete diagnostic findings with hedging words	"Findings suggestive of Arnold Chiari I malformation"
Other	Describing recommendations, imaging techniques, prior studies	"Another follow-up is recommended"

automatic identification and extraction of clinically important information.[9] However, little work has been done on applying NLP to assess the diagnostic certainty in radiology reports beyond the lexicon level (eg, hedging cue terms). NegBio (https://github.com/ncbi-nlp/NegBio)[10] was developed to detect negation and uncertainty in radiology reports, but the rule-based system highly depends on identifying relevant hedging terms only, without considering linguistic contexts. It only performs binary identification regarding whether a finding is uncertain or not. Other research has attempted to analyze the certainty of scientific statements in the biomedical literature through conventional machine learning methods (eg, conditional random fields,[11] support-vector machines[12-14]), but the ground truth annotation used to train these systems is based on which hedging expressions appear in the sentence, and these hedging expressions do not provide a reliable metric in diagnostic radiology reporting.

Recent innovations in deep learning technology provide improved NLP performance in radiology-related research.[15] Exploring state-of-the-art deep learning approaches to develop a reliable NLP system to qualify the context-aware certainty in radiology reporting has great potential to facilitate communications between radiologists and referring physicians. The goal of this study is to investigate the potential of deep learning NLP techniques for qualifying the certainty expressed in the Impression sections of radiology reports (QC-RAD). Specifically, our goal is the following:

1. Establish an NLP system for capturing contextualized certainty semantics at the sentence level in radiology reports and build the first annotated data of its kind
2. Develop a deep transfer learning approach so that knowledge representation learned from a very large universal textual data set can be transferred through fine-tuning the pretrained neural networks
3. Conduct experiments with 3 variants of bidirectional encoder representations from transformers (BERT) models with demonstrated promising results.

MATERIALS AND METHODS

Figure 1 shows the overview of the QC-RAD system, and we will describe each component in the following subsections.

Data Annotation

In this institutional review board–exempt quality improvement project, we initially randomly selected 1500 brain MRIs performed at a single academic medical center (UMass Memorial Medical Center). During the data-cleaning process, we transformed the reports into plain texts and extracted the Impression sections, which contain free-form texts written by the radiologists. We then loaded them into the extensible Human Oracle Suite of Tools (eHOST; https://github.com/chrisleng/ehost)[16] for context-aware certainty annotation. The eHost is versatile for annotation tasks and has been used by several institutions and projects for a variety of tasks, including i2b2: Informatics for Integrating Biology & the Bedside[17] and the Consortium for Health Care Informatics Research[18] projects.

We asked 3 board-certified radiologists (S.J.B., M.J.M., and N.A.) to read sentences from the Impression sections and to assign each sentence 1 of the 4 certainty categories shown in Table 1. We define "diagnostic findings" as a diagnostic opinion regarding a specific disease or other condition.[19] The certainty category is not only dependent on the hedging terms used but is also based on the holistic context expressed in the sentence. In contrast to previous studies, we did not try to distinguish different hedging terms or phrases from their different levels of certainty because they are perceived so differently by physicians, radiologists, and patients.[5,6] In our annotation guideline, we compiled a list of hedging terms (Online Supplemental Data), and using any of them is considered one of the factors contributing to uncertainties (defined in Table 1).

Initially, the 3 annotators each reviewed 30 head MR imaging reports and then went through an iterative process of adjusting the guideline and amending the annotations to reach consensus. Each annotator then independently annotated an additional 24

MR imaging reports according to the finalized annotation guideline (derived from the consensus review of the initial 30 reports). The interrater agreement was calculated (0.74) by mean pair-wise Cohen κ statistics, which showed substantial strength of agreement across annotators.[20] Finally, each annotator annotated 180 reports, resulting in a total of 594 reports.

The annotated data were then analyzed for certainty qualification, as follows: We performed word tokenization and sentence boundary identification on all the sentences from the Impression section of the radiology reports. We then removed sentences that contained fewer than 4 words (76 sentences) because these short sentences are typically noise caused by sentence-splitting errors, resulting in 2352 sentences in total for further analysis; 88.7% of the sentences were <25 words in length, with a mean of 14 (SD, 9.5). We then split the annotated certainty data into training data (80%), validation data (10%), and testing data (10%). The training and validation data were used for fine-tuning, and the test data were used as heldout (unseen) data to evaluate the final performance of the system. The data statistics on 3 datasets are shown in Table 2.

Deep Transfer Learning

We formulated the certainty assessment task into a multiclass sentence-classification problem and exploited NLP techniques to capture fine-grained semantics for classifying each sentence into 1 of the 4 categories defined in Table 1. Recent progress in NLP has been driven by using deep learning approaches,[21] and different deep learning architectures have been applied for text classification, which typically can be grouped into 2 model families: Convolutional neural networks are good at extracting local and position-invariant pattern features,[22] while recurrent neural networks are shown to perform better in modeling long dependencies among texts.[23] All those approaches require large amounts of labeled data to reliably estimate the numerous model parameters; however, compared with general domains, annotated data are more difficult and expensive to obtain in clinical domains because they require subject matter expertise for high-quality annotation.

Deep transfer learning[24] makes it possible to harness the power of deep neural architecture when only limited labeled data are available. In this study, we explored the state-of-the-art NLP transferring learning model BERT,[25] which was developed by Google Artificial Intelligence and demonstrated breakthrough performance improvement in a variety of NLP tasks. BERT is a contextualized word-representation model based on a masked language model and pretrained using bidirectional transformers,[26] and it can take into account sequential dependencies among words in a sentence for a semantically meaningful representation. More information about the BERT can be found in the Online Supplemental Data.

Preprocessing. We extracted data from the annotation tool and performed sentence segmentation using the Natural Language Toolkit (https://www.nltk.org/)[27] so that each sentence was paired with a certainty label described in Table 2. We then used the WordPiece tokenizer (https://www.paperswithcode.com/method/wordpiece)[28] for tokenization. It breaks each word down into its prefix, root, and suffix (subwords) in order to mitigate the out-of-vocabulary issue. For instance, "infarction" will be tokenized as 3 subwords: "in," "##far," and "##ction." To be compatible with BERT input format, we added the "[CLS]" token at the beginning of each sentence, and the "[SEP]" token at the end of each sentence.

Model Training (Fine-tuning). To address our sentence-classification problem, we added a drop-out regularization[29] and a softmax classifier layer on top of the pretrained BERT layer. In our work, we adopted 3 variants of pretrained BERT models: 1) The BERT-base model (https://github.com/google-research/bert) consists of an encoder with n ($n = 12$ in Fig 2) layers of transformer blocks and was pretrained using BookCorpus (https://github.com/soskek/bookcorpus) and Wikipedia;[25] 2) BioBERT

Table 2: Data statistics of the 3 data sets[a]

	Train Data Set	Valid Data Set	Test Data Set
Non-Definitive	585 (30.97%)	73 (30.93%)	73 (30.8%)
Definitive-Mild	329 (17.42%)	41 (17.37%)	42 (17.7%)
Definitive-Strong	503 (26.63%)	63 (26.69%)	63 (26.58%)
Other	472 (24.97%)	59 (25%)	59 (24.89%)
Total	1889 (100%)	236 (100%)	237 (100%)

[a] Data are the number of sentences and corresponding percentage.
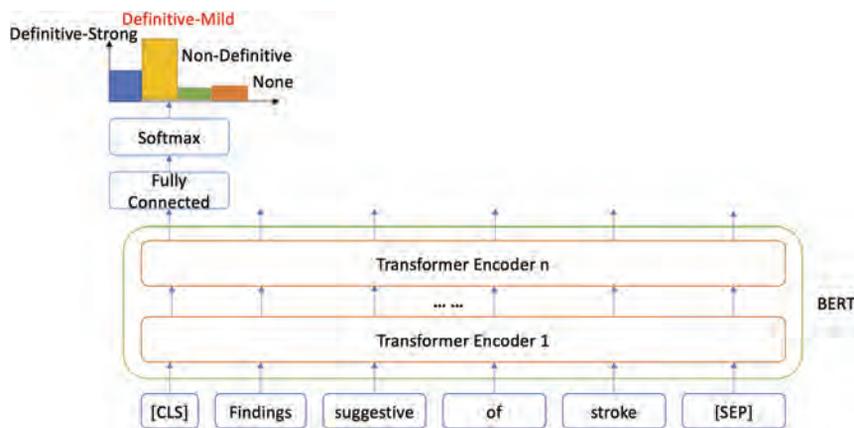
FIG 2. Illustration of using BERT for certainty classification. The input "Findings suggestive of stroke" was classified as "Definitive-Mild."

Table 3: Performance comparison among 3 BERT variants (with their optimal parameters) on the validation data set

Model	No of Epochs	Batch Size	Learning Rate	Macro-Sensitivity (%) (95% CI)	Macro-Specificity (%) (95% CI)	Macro-AUC (95% CI)
BERT-base	4	24	0.00003	79.46 (68.02–87.82)	93.65 (89.26–96.46)	0.928 (0.883–0.973)
BioBERT	6	32	0.00003	79.08 (67.13–87.78)	93.13 (88.58–96.13)	0.931 (0.886–0.975)
ClinicalBERT	5	32	0.00005	78.52 (66.91–87.07)	93.19 (88.57–96.25)	0.925 (0.878–0.971)

Note:—Macro indicates the average on the macro level across different categories.

(https://github.com/dmis-lab/biobert)[30] was initialized using the BERT-base and was pretrained using BookCorpus, Wikipedia, PubMed abstracts, and PubMed Central full text articles; 3) ClinicalBERT (https://github.com/EmilyAlsentzer/clinicalBERT)[31] was initialized using BioBERT and pretrained using around 2 million clinical notes in the MIMIC-III database (https://physionet.org/content/mimiciii/1.4/).[32] All 3 model variants shared the same architecture with the BERT-base model, which consists of 12 layers of transformer blocks with a hidden size of 768, and 12 self-attention heads. We fine-tuned all the layers in the model-training process. The 3 model variants shared the following hyperparameters: drop-out probability of 0.1 and maximum sequence (subwords) length of 80. The optimal batch size, learning rate, and number of epochs were chosen using the validation data (see the Results section).

Model Evaluation. We evaluated the performance QC-RAD using standard metrics: Sensitivity, specificity, and area under the receiver operating characteristic curve (AUC). For aggregated evaluation across 4 certainty categories, we use the macro-average value, which calculates each metric (sensitivity, specificity, or AUC) independently for each category and then takes the average. We prefer macro-average to micro-average because we value the ability of QC-RAD to perform equally well across different categories. In the context of imbalanced data in which the majority category has many more samples than other categories, micro-average will be biased toward the dominating majority categories, while macro-average is less sensitive and considers each category equally.

Figure 2 shows an example. The input sentence "Findings suggestive of stroke" is composed of a sequence of 4 words. [CLS] and [SEP] are added prior to being fed into the BERT model. We first initialize n transformer encoders ($n = 12$ orange blocks in Fig 2) using the pretrained BERT model, and all the parameters, including the fully connected layer, will be fined-tuned through supervised learning using the labeled data (see the "Data Annotation" section) for certainty classification. Through a multilayer deep neural network architecture (encoder) in BERT, each input token will be transformed to a final output embedding (vector representation). For this sentence-classification task, we use only the final hidden state of the first token [CLS], which is considered aggregated sentence representations, and feed it into a fully connected layer to obtain a probability distribution across 4 certainty categories through the softmax function.

RESULTS

We reported the classification performance against the reference standard categories assigned by radiologists. The aggregated

results are presented with the macro-average across 4 categories defined in Table 1. For 3 pretrained language models, we used grid-search to optimize the batch size (range in 24, 32, and 64) and learning rate (range in 0.000005, 0.00001, 0.00003, and 0.00005) during the fine-tuning process for a fair comparison. The number of epochs for fine-tuning training was selected on the basis of the peak AUC score on the validation data.

Comparative Performance among 3 BERT Models (Validation Data)

We compared the performance on the certainty classification task using the 3 pretrained BERT models, and the results are shown in Table 3. The BioBERT model obtained the best macro-AUC of 0.931, and the BERT-base yielded the best macro-sensitivity of 79.46% and specificity of 93.65%, while ClinicalBERT achieved the relatively lower macro-sensitivity of 78.52% compared with the other 2 models (Table 3).

Performance Curve in the Fine-tuning Process (Validation Data)

Figure 3 shows the performance curve of the BERT-base (left) and BioBERT (right) across the number of epochs during the fine-tuning process. Here, we also show the F1 score, which is the harmonic mean of the positive predictive value and sensitivity. We observed similar trends on both models. With fine-tuning, all the performance metrics increased initially and plateaued after approximately 5 epoch trainings.

Performance on the Test Data

On the basis of the evaluation results (AUC scores) on the validation data, we chose the best system (BioBERT) and applied it to the test data as shown in Table 4. The system performs the best on the "Other" category, with the highest sensitivity of 98.31%, specificity of 97.19%, and AUC of 0.994. Among the other 3 categories, the system obtained the highest sensitivity for Non-Definitive (76.71%), the highest specificity for Definitive-Strong (95.4%), and the highest AUC for Definitive-Strong (0.964). Overall, it obtained the macro-average sensitivity of 77.29%, specificity of 92.89%, and AUC of 0.93 on the heldout unseen data. Although the Non-Definitive class has a lower AUC score than Definitive-Strong, the sensitivity of Non-Definitive was better than that of Definitive-Strong (76.71% versus 74.6%) as shown in Table 4. In Fig 4, receiver operating characteristic curves of Definitive-Strong (class 2) and Other (class 3) are closer to the ideal spot (the closer to the upper left corner, the better).

Error Analysis

We conducted error analysis on the validation data, and the confusion matrix is shown in Table 5. Rows represent the truth label
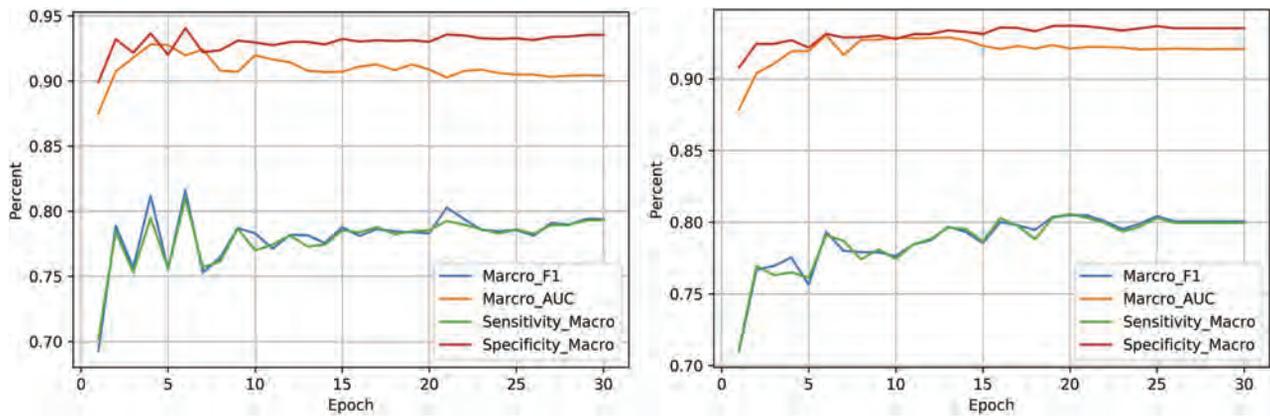
FIG 3. Performance curve across the number of fine-tuning epochs. The left figure is for BERT, and the right one is for BioBERT.

Table 4: System performance of BioBERT on the test data set[a]

Category	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	AUC (95% CI)
Non-Definitive	76.71 (56/73) (65.35–85.81)	90.24 (148/164) (84.64–94.32)	0.919 (0.874–0.964)
Definitive-Mild	59.52 (25/42) (43.28–74.37)	88.72 (173/195) (83.42–92.79)	0.843 (0.76–0.92)
Definitive-Strong	74.6 (47/63) (62.06–84.73)	95.4 (166/174) (91.14–97.99)	0.964 (0.931–0.997)
Other	98.31 (58/59) (90.91–99.96)	97.19 (173/178) (93.57–99.08)	0.994 (0.979–1)
Macro Avg	77.29 (65.4–86.22)	92.89 (88.19–96.05)	0.93 (0.888–0.972)

Note:—Macro Avg indicates average on the macro level across different categories.
[a] Numerators and denominators for sensitivity and specificity are included in parentheses.
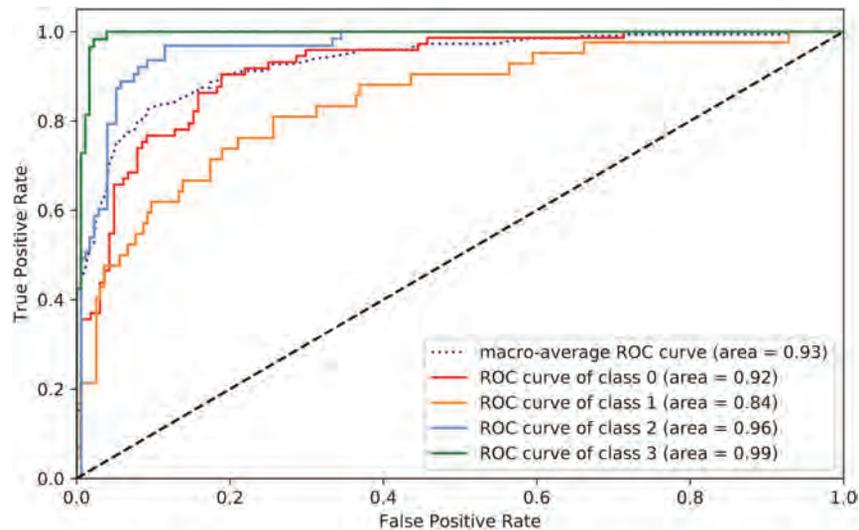
FIG 4. Receiver operating characteristic curves of individual classes on the test data set. Class 0 = Non-Definitive, class 1 = Definitive-Mild, class 2 = Definitive-Strong, and class 3 = Other.

assigned by domain experts, and columns indicates the system predictions. It shows that only 1 (1.7%) sentence in the Other category was wrongly classified as Non-Definitive, which explains the high performance of this category in Table 4. On the basis of the definition of the "Other" category, it covers a narrow scope of semantics, and it is easy for the system to pick up the representative patterns (eg, "follow-up" is a reliable indicator for recommendations). The top 3 error patterns are the following: 1) Non-Definitive → Definitive-Mild (11 of 73, 15%), 2) Definitive-Mild → Non-Definitive (11 of 41, 26.8%), and 3) Definitive-Strong →

Non-Definitive (11 of 63, 17.5%). Those error patterns suggest that the Non-Definitive category is more challenging to distinguish from the other 2 definitive categories.

DISCUSSION

We demonstrated that the interrater agreement of subject matter experts on certainty interpretations when considering both hedging term usage and surrounding linguistic contexts is excellent. Building on this ground truth, we then demonstrated that deep

Table 5: Confusion matrix among different categories

Truth	Prediction			
	Non-Definitive	Definitive-Mild	Definitive-Strong	Other
Non-Definitive	56	11	3	3
Definitive-Mild	11	28	2	0
Definitive-Strong	11	5	46	1
Other	1	0	0	58

transfer learning shows great potential for unlocking contextualized semantics for certainty assessment of radiology reports using limited annotated data. Our novel QC-RAD system holds the potential to facilitate precision communication of imaging findings, as well as to serve as a new quality measure in radiology reporting.

Multiple publications have stressed the importance of the accurate conveying of diagnostic certainty in the radiology report. Most recently, a study developed a certainty scale, specifying recommended and nonrecommended certainty terms. The adoption of such a scale significantly increased the proportion of recommended certainty terms during a voluntary period.[33] Similar to the idea of a standardized certainty lexicon,[7] this approach is limited to term-level certainty, without taking into account contextual semantics. In contrast to prior work that has used NLP to simply identify the instances of these predefined terms and to count the frequency of specific, predefined hedging terms, our work has developed a deep learning–based NLP algorithm that will read and interpret the level of certainty conveyed in the free text, including surrounding contexts.

Our study categorized the certainty of each sentence in the Impression section of radiology reports into different certainty categories. We showed that our automated categorization scheme has strong operating characteristics compared with a ground truth based on the radiologists' consensus. The sentence-level certainty categorization is a first step toward a more general quantification of diagnostic uncertainty conveyed in radiology reports.

Deep learning approaches have shown breakthrough results in many tasks, but it is challenging to train a reliable deep neural network with limited annotation data in a specific domain. Pretrained language models, such as BERT, greatly alleviate this problem by training a deeply bidirectional language representation in an unsupervised manner using only a plain text corpus. By leveraging this universally learned knowledge, we performed the fine-tuning using the task-specific annotated data so that the learned deep encoder can be adapted to better fit our target task. Our experiments show promising results on all 3 variants of the BERT models, leading to the best macro-average AUC score of 0.93 on the unseen test data.

Among the 3 variant models of BERT, though ClinicalBERT was pretrained using clinical notes, it did not show any advantage compared with the other 2 models (BERT and BioBERT) as it did in other benchmark tasks.[31] This finding is possibly because clinical notes often contain ill-formed, nongrammatic sentences, arbitrary abbreviations, and typographic errors, which are less likely to be present in the Impression section of radiology report. Therefore, the expanded knowledge from clinical notes may not necessarily benefit the current task and could potentially introduce noise. BioBERT did show a slight overall performance gain in terms of the AUC score; however, at a certain threshold, BERT-base outperformed it on the basis of sensitivity and specificity.

The Impression of the radiology report reflects the radiologist's interpretation of actionable findings on the imaging study. Sometimes the diagnoses are inherently uncertain by imaging alone. However, when there is near-certainty about the diagnosis, radiologists should convey confidence.[8] Our ultimate goal is to provide the radiologist with a real-time, automatic measurement of the level of diagnostic certainty in their report before signing. Thus, the radiologists will have objective information about the level of certainty that they are conveying. The current system classifies 1 sentence into a discrete certainty category. In future work, we will expand the system by combining the discrete certainty category with the probability that the certainty level is correct. Once this expanded system has been developed and appropriately validated, the quantification of the certainty level conveyed in the radiology report may be used as a quality metric to evaluate radiologists' performance.

There are limitations in this study. First, the data size used for this study is relatively small, yet the high-quality annotation and transfer-learning strategy enable the system to learn efficiently, yielding promising results on the withheld testing data. Leveraging a larger amount of unlabeled data (eg, unsupervised representation learning[34]) would potentially further improve the performance of the system, which we will explore for future work. Second, we developed and evaluated the QC-RAD system using head MR imaging radiology reports only, and its generalizability to other radiologic specialties needs further investigation. However, because our system is fully trainable and does not depend on any heuristic rules, we speculate that it can be easily generalized to other radiology subspecialties through active-learning[35] and domain-adaptation learning[36] techniques. Third, cross-institutional external validation on the performance of the system and ground truth consensus is needed to verify the overall generalizability of the study, which we will pursue in the near future. We did conduct an external validation within our institution. Specifically, we randomly sampled a new set of 40 MR imaging head reports from 4 new neuroradiologists (10 reports each) who were not covered by our original data collection. We asked the same 3 radiologists to assign 1 of the 4 certainty categories to all the 132 sentences from the Impression sections of these 40 reports. We found that the interannotator agreement remained very high with a mean pairwise κ score of 0.761. We chose the annotations from the annotator who agreed most with the other 2 as ground truth and evaluated the performance of QC-RAD on this new data set, achieving the macro-sensitivity of 84.01%, macro-specificity of 93.59%, and a macro-AUC of 0.945. This validation experiment demonstrates great generalizability of our QC-RAD system and ground truth consensus.

CONCLUSIONS

We developed and validated a deep transfer learning system, QC-RAD, to automatically assess the level of certainty in head MR imaging reports. Experimental results demonstrated that QC-RAD can effectively unlock contextualized semantics of free-text reporting language for assessment of diagnostic certainty in radiology reports, holding the potential to facilitate precision communication of imaging findings between radiologists and referring physicians.

REFERENCES

1. American College of Radiology. **ACR practice guidelines for communication of diagnostic imaging findings. 2014. Revised 2020.** https://www.acr.org/-/media/acr/files/practice-parameters/communicationdiag.pdf. Accessed January 7, 2019

2. Swensen SJ, Johnson CD. **Radiologic quality and safety: mapping value into radiology.** *J Am Coll Radiol* 2005;2:992–1000 CrossRef Medline

3. Lafortune M, Breton G, Baudouin JL. **The radiological report: what is useful for the referring physician?** *Can Assoc Radiology J* 1988;39:140–43 Medline

4. Lee B, Whitehead MT. **Radiology reports: What YOU Think You're Saying and What THEY Think You're Saying.** *Curr Probl Diagn Radiol* 2017;46:186–95 CrossRef Medline

5. Gunn AJ, Tuttle MC, Flores EJ, et al. **Differing interpretations of report terminology between primary care physicians and radiologists.** *J Am Coll Radiol* 2016;13:1525–29.e1 CrossRef Medline

6. Rosenkrantz AB. **Differences in perceptions among radiologists, referring physicians, and patients regarding language for incidental findings reporting.** *AJR Am J Roentgenol* 2017;208:140–43 CrossRef Medline

7. Panicek DM, Hricak H. **How sure are you, doctor? A standardized lexicon to describe the radiologist's level of certainty.** *AJR Am J Roentgenol* 2016;207:2–3 CrossRef Medline

8. Hoang JK. **Do not hedge when there is certainty.** *J Am Coll Radiol* 2017;14:5 CrossRef Medline

9. Pons E, Braun LM, Hunink MG, et al. **Natural language processing in radiology: a systematic review.** *Radiology* 2016;279:329–43 CrossRef Medline

10. Peng Y, Wang X, Lu L, et al. **NegBio: a high-performance tool for negation and uncertainty detection in radiology reports.** *AMIA Jt Summits Transl Sci Proc* 2018;2017:188–96 Medline

11. Agarwal S, Yu H. **Detecting hedge cues and their scope in biomedical text with conditional random fields.** *J Biomed Inform* 2010;43:953–61 CrossRef Medline

12. Light M, Qiu XY, Srinivasan P. **The Language of Bioscience: Facts, Speculations, and Statements In Between.** HLT-NAACL 2004 Workshop: Linking Biological Literature, Ontologies and Databases. Association for Computational Linguistics 2004:17–24.

13. Shatkay H, Pan F, Rzhetsky A, et al. **Multi-dimensional classification of biomedical text: toward automated, practical provision of high-utility text to diverse users.** *Bioinformatics* 2008;24:2086–93 CrossRef Medline

14. Vincze V, Szarvas G, Móra G, et al. **Linguistic scope-based and biological event-based speculation and negation annotations in the BioScope and Genia Event corpora.** *J Biomed Semantics* 2011;2 (Suppl 5):S8 CrossRef Medline

15. Sorin V, Barash Y, Konen E, et al. **Deep learning for natural language processing in radiology: fundamentals and a systematic review.** *J Am Coll Radiol* 2020;17:639–48 CrossRef Medline

16. South BR, Shen S, Leng J, et al. **A prototype tool set to support machine-assisted annotation.** In: *Proceedings of the 2012 Workshop on Biomedical Natural Language Processing*, Montreal, Canada. June 8, 2012:130–39

17. Uzuner Ö, South BR, Shen S, et al. **2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text.** *J Am Med Inform Assoc* 2011;18:552–56 CrossRef Medline

18. South BR, Mowery D, Suo Y, et al. **Evaluating the effects of machine pre-annotation and an interactive annotation interface on manual de-identification of clinical text.** *J Biomed Inform* 2014;50:162–72 CrossRef Medline

19. Medical diagnosis. **Wikipedia. 2019.** https://en.wikipedia.org/w/index.php?title=Medical_diagnosis&oldid=900598243. Accessed June 15, 2019

20. Landis JR, Koch GG. **The measurement of observer agreement for categorical data.** *Biometrics* 1977;33:159–74 Medline

21. Young T, Hazarika D, Poria S, et al. **Recent trends in deep learning based natural language processing.** *IEEE Comput Intell Mag* 2018:13:55–75 CrossRef

22. Kim Y. **Convolutional Neural Networks for Sentence Classification.** In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar. October 25–29, 2014:1746–51 CrossRef

23. Liu P, Qiu X, Huang X. **Recurrent neural network for text classification with multi-task learning.** In: *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, New York, New York. July 9–15, 2016:2873–79

24. Tan C, Sun F, Kong T, et al. **A survey on deep transfer learning.** *arXiv* 2018 http://arxiv.org/abs/1808.01974. Accessed May 24, 2019

25. Devlin J, Chang MW, Lee K, et al. **BERT: pre-training of deep bidirectional transformers for language understanding.** In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Minneapolis, Minnesota. June 2–7, 2019:4171–86 CrossRef

26. Vaswani A, Shazeer N, Parmar N, et al. **Attention is all you need.** In: *Proceedings of the 31st Annual Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, California. December 4–9, 2017:5998–6008

27. Bird S, Klein E, Loper E. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit.* O'Reilly Media, Inc; July 21, 2009

28. Wu Y, Schuster M, Chen Z, et al. **Google's Neural Machine Translation system: bridging the gap between human and machine translation.** *arXiv* 2016. http://arxiv.org/abs/1609.08144. Accessed April 16, 2019

29. Srivastava N, Hinton G, Krizhevsky A, et al. **Dropout: a simple way to prevent neural networks from overfitting.** *J Mach Learn Res* 2014;15:1929–58

30. Lee J, Yoon W, Kim S, et al. **BioBERT: a pre-trained biomedical language representation model for biomedical text mining.** *Bioinformatics* 2020;36:1234–40 CrossRef Medline

31. Alsentzer E, Murphy J, Boag W, et al. **Publicly available clinical BERT embeddings.** *Proceedings of the 2nd Clinical Natural Language Processing Workshop*, Minneapolis, Minnesota. 2019;72–78 CrossRef

32. Johnson AE, Pollard TJ, Shen L, et al. **MIMIC-III: a freely accessible critical care database.** *Sci Data* 2016;3:160035 CrossRef Medline

33. Shinagare AB, Alper DP, Hashemi R, et al. **Early adoption of a certainty scale to improve diagnostic certainty communication.** *J Am Coll Radiol* 2020;17:1276–84 CrossRef Medline

34. Li R, Hu B, Liu F, et al. **Detection of bleeding events in electronic health record notes using convolutional neural network models enhanced with recurrent neural network autoencoders: deep learning approach.** *JMIR Med Inform* 2019;7:e10788 CrossRef Medline

35. Tang YP, Li GX, Huang SJ. **ALiPy: active learning in Python.** *arXiv* 2019. http://arxiv.org/abs/1901.03802. Accessed June 17, 2019

36. Tzeng E, Hoffman J, Saenko K, et al. **Adversarial discriminative domain adaptation.** *arXiv* 2017. https://arxiv.org/abs/1702.05464. Accessed May 20, 2019

Artificial Intelligence–Based 3D Angiography for Visualization of Complex Cerebrovascular Pathologies

S. Lang, P. Hoelter, M. Schmidt, C. Strother, C. Kaethner, M. Kowarschik, and A. Doerfler

ABSTRACT

BACKGROUND AND PURPOSE: By means of artificial intelligence, 3D angiography is a novel postprocessing method for 3D imaging of cerebral vessels. Because 3D angiography does not require a mask run like the current standard 3D-DSA, it potentially offers a considerable reduction of the patient radiation dose. Our aim was an assessment of the diagnostic value of 3D angiography for visualization of cerebrovascular pathologies.

MATERIALS AND METHODS: 3D-DSA data sets of cerebral aneurysms ($n_{CA} = 10$), AVMs ($n_{AVM} = 10$), and dural arteriovenous fistulas (dAVFs) ($n_{dAVF} = 10$) were reconstructed using both conventional and prototype software. Corresponding reconstructions have been analyzed by 2 neuroradiologists in a consensus reading in terms of image quality, injection vessel diameters (vessel diameter [VD] 1/2), vessel geometry index (VGI = VD1/VD2), and specific qualitative/quantitative parameters of AVMs (eg, location, nidus size, feeder, associated aneurysms, drainage, Spetzler-Martin score), dAVFs (eg, fistulous point, main feeder, diameter of the main feeder, drainage), and cerebral aneurysms (location, neck, size).

RESULTS: In total, 60 volumes have been successfully reconstructed with equivalent image quality. The specific qualitative/quantitative assessment of 3D angiography revealed nearly complete accordance with 3D-DSA in AVMs (eg, mean nidus size$_{3D\ angiography/3D-DSA}$= 19.9 [SD, 10.9]/20.2 [SD, 11.2] mm; $r = 0.9$, $P = .001$), dAVFs (eg, mean diameter of the main feeder$_{3D\ angiography/3D-DSA}$= 2.04 [SD, 0.65]/2.05 [SD, 0.63] mm; $r = 0.9$, $P = .001$), and cerebral aneurysms (eg, mean size$_{3D\ angiography/3D-DSA}$= 5.17 [SD, 3.4]/5.12 [SD, 3.3] mm; $r = 0.9$, $P = .001$). Assessment of the geometry of the injection vessel in 3D angiography data sets did not differ significantly from that of 3D-DSA (vessel geometry index$_{AVM}$: $r = 0.84$, $P = .003$; vessel geometry index$_{dAVF}$: $r = 0.82$, $P = .003$; vessel geometry index$_{CA}$: $r = 0.84$, $P < .001$).

CONCLUSIONS: In this study, the artificial intelligence–based 3D angiography was a reliable method for visualization of complex cerebrovascular pathologies and showed results comparable with those of 3D-DSA. Thus, 3D angiography is a promising postprocessing method that provides a significant reduction of the patient radiation dose

ABBREVIATIONS: AI = artificial intelligence; CA = cerebral aneurysm; dAVF = dural arteriovenous fistula; 3DA = 3D angiography; 3D-DSA = 3D digital subtraction angiography; VD = vessel diameter; VGI = vessel geometry index; VRT = volume rendering technique

DSA is the current criterion standard for diagnostics of the cerebral vasculature by acquisition of time-resolved 2D-DSA series and 3D rotational angiography (3DRA).[1-3] Especially for the

assessment of complex vascular anatomy, 3D-DSA improves understanding of pathologies and their spatial relationship to adjacent vascular structures.[4] Thus, 3D imaging has become an irreplaceable element in the diagnostic work-up of cerebrovascular pathologies such as cerebral aneurysms (CAs), AVMs, and dural arteriovenous fistulas (dAVFs).[5-8]

Like conventional 2D-DSA series, 3D-DSA is based on the subtraction of a non-contrast-enhanced mask volume (mask run) from a contrast-enhanced volume (fill run), thereby allowing separation of vascular from nonvascular structures. The effective patient dose for a noncollimated 3D-DSA acquisition is 0.9 mSv.[9] Of this amount, the mask run accounts for about 50% of the total radiation dose for both patients and investigators.[10] Hence, a reduction of the radiation dose required for 3D imaging of cerebral vessels is highly desirable.

There are already, however, valuable considerations for dose reduction, for example varying the scanning parameters (eg, scaling down the dose per frame) for acquisition of both mask and fill runs[11,12] or even removing the mask run and separating opacified vascular structures from soft tissue by adjusting thresholds of the contrast-enhanced 3D data set (3DRA).[13-15] While a variation of sensitive scanning parameters seems to be limited in terms of a reliable reproduction of vessels, 3DRA based on thresholding provides sufficient quality for evaluation of CAs. However, visualization of small vessels with 3DRA highly depends on the homogeneity of the applied contrast medium bolus and is thus highly susceptible to windowing.[16]

Recently, interesting approaches for dose-optimized 3D-DSA imaging of the cerebral vasculature have been proposed. Novel postprocessing techniques enable differentiation of vessels from adjacent structures using artificial intelligence (AI) and require only a single contrast-enhanced fill run.[17,18] Consequently, these AI-based algorithms provide dose savings of up to 50% and might reduce a potential decrease in image quality caused by motion artifacts due to their immunity to misregistration.[13] From a technical perspective, precise classification of different types of tissue is of central importance for these AI-based algorithms, whereby algorithms with differentiation of 3 types of tissue such as vasculature versus bone versus soft tissue[17] and 2 types of tissue such as vasculature versus nonvasculature[18] have been reported, respectively. Despite different classification strategies, both algorithms have shown promising results for generating DSA-like 3DA with diagnostic quality in preliminary assessments.

So far, these AI-based algorithms have been exclusively evaluated using data sets either with CAs or without pathologic findings. However, the diagnostic relevance of 3D-DSA applications is the highest in complex vascular pathologies like AVMs and dAVFs because 2D-DSA frequently has vascular overlap in these cases.

Hence, we present our experience with an AI-based 3DA prototype by means of a case series of CAs, AVMs, and dAVFs. In particular, our aim was to compare this prototypical 3DA technique classifying 2 types of tissue (vasculature versus nonvasculature) with the current standard 3D-DSA in terms of quantitative and qualitative parameters.

MATERIALS AND METHODS
Patient Selection
In a retrospective analysis, 3D-DSA data sets from 28 patients (age_{mean} = 58.7 [SD, 12.56] years; $n_{female/male}$ = 11:17) with unruptured AVMs (n_{AVM} = 10), unruptured CAs (n_{CA} = 10), and dural arteriovenous fistulas (n_{dAVF} = 10) were evaluated. All data sets were acquired before treatment of the underlying condition in the ICA in 17 cases (57%), the vertebral artery in 9 cases (30%), and the external carotid artery in 4 cases (13%). Informed consent was obtained from all patients enrolled.

3D Angiography
The prototypical 3D angiography (3DA) refers to a novel postprocessing approach with the aim of generating DSA-like 3D volumes with diagnostic image quality. In contrast to 3D-DSA, the AI-based 3DA requires only a single contrast-enhanced run (fill run) to differentiate vessels from adjacent structures.

From a technical perspective, the core of the 3DA technique can be seen as a 3D segmentation task and is based on a classification of contrast-enhanced vascular and nonvascular structures in fill run data by applying a deep convolutional neural network[19] specifically trained for this task before our evaluation. In particular, the convolutional neural network is trained to determine a 3D binary segmentation mask on the basis of the fill run data, which can, in turn, be applied to the fill run data to generate a DSA-like 3D volume. The chosen network architecture is conceived as a fast 3D network with a small memory footprint. Thus, a feed-forward 3D convolutional neural network using rectified linear unit activations normalized through batch normalization[20] was selected. Due to the binary nature of the classification task (ie, distinguishing vascular and nonvascular structures), the binary cross-entropy was set as a loss function. The optimization is based on the Adam optimization approach,[21] an extension of the classic stochastic gradient descent, using a batch size of 64. The base learning rate was set to 0.01. To incorporate not only local but also surrounding contextual information into the classification, we used dilated convolutions.

Analogous to an application in 1D scenarios such as generating audio waveforms (as proposed in van den Oord et al[22]), the concept of dilated convolutions can be applied to multidimensional application scenarios[23] as well. The network features 4 convolutional layers with 8 filters each, a kernel size of $5 \times 5 \times 5$ voxels, and a dilation rate of 2. The final layer features a kernel size of $3 \times 3 \times 3$ voxels with 1 result filter that constitutes the classification result. The advantage of this formulation is that the proposed algorithm is fully convolutional and allows a volume of arbitrary size (eg, $512 \times 512 \times 512$ voxels) to be used as input to the classification. Furthermore, the chosen architecture allows a reasonable trade-off, especially for larger 3D volumes between the application run time and classification quality. Key aspects of the network architecture are illustrated in Fig 1.

In total, 98 retrospectively selected 3D-DSA data sets (free from metal or motion artifacts) covering a variety of conditions (to prevent overfitting) represented the training data for the 3DA algorithm. The training process was based on the presentation of both extracted contrast-enhanced runs (fill runs) and complete 3D-DSA volumes (containing mask and fill runs) to the deep convolutional neural network, aiming at a suitable classification between contrast-enhanced vascular and nonvascular structures, ie, the segmentation of the vascular structures in a 3D volume. For each individual data set used for the training of the 3DA algorithm, vascular extraction (ie, the derivation of a 3D binary segmentation mask) was performed by thresholding the subtracted 3D-DSA volumes. To determine a reasonable threshold, we considered the filling of the vasculature as well as the exclusion of undesirable influences (such as inconsistencies, image artifacts, noise, and so forth). To cover a sufficient variety of potentially suitable thresholds (eg, comparable with the personal preferences of multiple different clinical experts for a specific volume), we also included multiple thresholding choices simultaneously in the training process. Finally, this form of augmentation is intended to provide a better overall performance of the 3DA algorithm.

For the classification, randomly selected patches with a size of $31 \times 31 \times 31$ voxels were applied from the presented 3D volumes
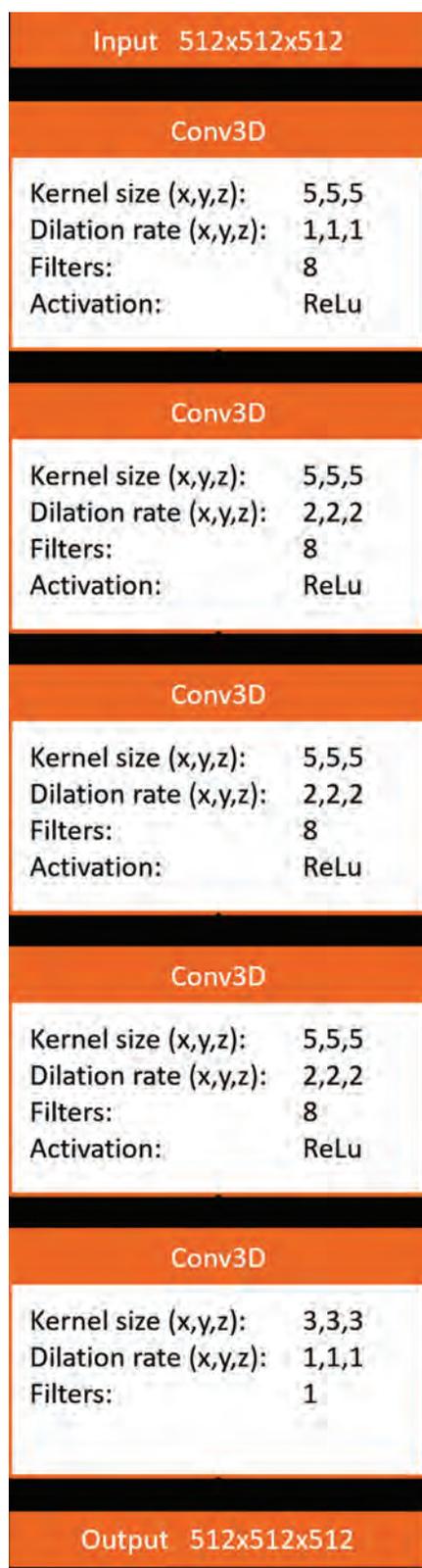
FIG 1. Key aspects of the chosen network architecture. In addition to the in- and output of the network, kernel sizes, dilation rates, number of filters, and activation function (ie, rectified linear unit [ReLu]) are shown. Conv3D indicates 3D convolutional neural network.

during the training. The chosen size of the patches is the minimum input size for the network and results in a single-voxel response. The motivation for this choice was to foster randomness during the training, which, in turn, can contribute to the robustness and quality of the resulting classifier. All parameters were determined using hyperparameter tuning. Note that during the application of the prototypical 3DA method, the processing was not patch-based but was performed on an entire volume data set. Subsequent multimodal testing and validation of the prototypical algorithm of this classification were performed with separate 3D-DSA data sets. No data set of the training or validation was part of our series for evaluation of 3DA in cases of AVMs, CAs, and dAVFs.

Data Acquisition and Postprocessing

3D-DSA was acquired using a biplane flat panel detector angiographic system (Artis zee biplane; Siemens). By means of standard angiographic methods, a 5F catheter was positioned in the proximal ICA, external carotid artery, or vertebral artery to obtain 3D-DSA data sets using the standard acquisition protocol (5-second 3D-DSA) as provided by the manufacturer. Here, an initial rotational scan (native mask run) is followed by a second rotational scan (contrast-enhanced fill run) of 5 seconds each. Each run yields 133 projections (rotational angle = 200°). The detector dose per projection image is selected as 0.36 μGy (70 kV, 1240 × 960 detector elements with 2 × 2 binning of pixels, projection on a 30 × 40 cm flat panel size, with increment of 1.5°/frame, a frame-rate of 30 frames/s). According to our protocol, a manual injection of contrast was initiated 1 second before the beginning of the fill run, maintained for 6 seconds, and stopped after the C-arm system had covered the complete rotation angle. The total contrast volume was 15 mL of Imeron 300 (iopamidol; Bracco).

Both mask and fill runs of the acquired 3D-DSA data sets were transferred to a dedicated workstation (syngo X Workplace; Siemens) running both commercially available software for conventional 3D-DSA postprocessing and an additional software prototype plug-in for the 3DA postprocessing. Reconstruction of the 3D-DSA volumes was performed using data derived from both runs, whereas reconstruction of 3DA volumes was accomplished using only data from the fill runs.

According to standardization, we used conventional reconstruction parameters for both 3DA and 3D-DSA (kernel type: edge-enhanced; characteristics: smooth; 512 × 512 image matrix).

Image Evaluation

The image quality of all data sets was evaluated for parameters that could compromise the diagnostic value using a 5-fold grading-scale (Table 1). The 3D-DSA and 3DA reconstructions were assessed in a consensus reading by 2 experienced neuroradiologists (6 and 10 years of clinical experience) blinded to the type of reconstruction (based on either the subtraction technique or the AI algorithm).

Table 1: Image quality

Grade	Characteristics
4	Excellent (high contrast, no artifacts)
3	Good (high contrast; minimal artifacts, eg, due to movement or implants)
2	Compromised (eg, noticeable movement artifacts and/or reduced homogeneity of the vessel contrast)
1	Heavily compromised (low contrast and/or strong movement artifacts)
0	Not diagnostic (vasculature is not differentiable due to heavy artifacts and/or missing contrast)
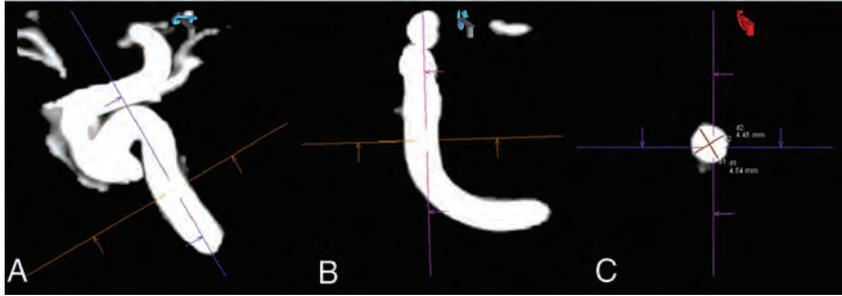
FIG 2. Sample measurement of the vascular dimensions of an ICA (C4 segment) in a 3D-DSA data set. The maximum vessel diameters have been assessed for both 3D-DSA and 3DA in 2 projections each (VD1 and VD2) using multiplanar reconstructions. *A*, Sagittal orientation of the C4 segment. *B*, Coronary orientation of the C4 segment. *C*, Finally, both vessel diameters allow the calculation of the VGI (VGI = VD1/VD2).

Assessment of 3DA and 3D-DSA Reconstructions

Vessel Geometry Index. For all 3DA- and 3D-DSA data sets, the maximum transversal diameters of the injection vessels (in millimeters; vessel diameter [VD]1/VD2) were measured in multiplanar reconstructions. Respecting the three-dimensionality of these data sets, the ratio of VD1 and VD2 was defined as the vessel geometry index (VGI = VD1/VD2).[18] Measurements of the anterior and posterior circulation were performed at the C4 segment of the ICA and the V4 segment of the vertebral artery, respectively. If branches of the external carotid artery were involved, the maximum transversal diameter of the proximal external carotid artery was measured. See Fig 2 for an illustrative measurement of VD1 and VD2 at the ICA.

AVMs. The degree of agreement between the 3DA and 3D-DSA data sets of the AVMs was evaluated with a combination of MPR/MIP/VRT images from both types of reconstructions. As qualitative parameters, involvement of eloquent brain areas, origin of the main feeder, venous drainage (superficial, deep, mixed), and the presence of venous varix or stenosis were determined. Moreover, the number of feeders, the presence and number of aneurysms on feeding arteries, the number of intranidal aneurysms, and the maximum diameter of the nidus were determined as quantitative parameters. On the basis of these data, Spetzler-Martin scores were also determined.

dAVFs. The degree of agreement between the 3DA and 3D-DSA data sets of the dAVFs was also evaluated with a combination of MPR/MIP/VRT images from both types of reconstructions. As qualitative parameters, the origin of the main feeder (eg, middle/posterior meningeal artery, ophthalmic artery, ascending pharyngeal artery), the localization of the fistulous point (anterior/middle/posterior cranial fossa), the primary vessel of venous drainage (vein or sinus), and drainage of the fistula into a dural sinus (transverse sigmoid, petrous, superior sagittal, and straight sinuses) were recorded. As a quantitative parameter, the maximum diameter of the main feeder was also obtained.

CAs. The degree of agreement between the 3DA and 3D-DSA data sets of the CAs was also evaluated with a combination of MPR/MIP/VRT images from both types of reconstructions. As qualitative parameters, the localization (parent artery) of the aneurysm and its neck configuration were recorded. As quantitative parameters, the maximum diameter of aneurysmal sac and the number of aneurysmal blebs were also noted.

Statistical Analysis

Statistical analysis was performed using commercially available software (SPSS Statistics, Version 20; IBM).

Because of their noncontinuous character, qualitative parameters (eg, location of the pathology and so forth) were analyzed using descriptive statistics.

Quantitative parameters from both groups (eg, vessel diameters, VGIs, diameter of the AVM nidus, diameter of the main dAVF feeder, and diameter of aneurysmal sac) were tested for normal distribution using the D'Agostino-Pearson test (if $P > .05$, normality was accepted) and were compared using the Pearson correlation coefficient (r) and a paired t test (P), respectively. Any other quantitative parameters (eg, number of AVM feeders, aneurysms on AVM feeding arteries, number of intranidal aneurysms, number of aneurysmal blebs, and so forth) were analyzed using descriptive statistics.

RESULTS
Image Quality
The 3DA and 3D-DSA reconstructions ($n = 60$), respectively, were all of diagnostic quality with a low number of cases with reduced image quality due to motion artifacts (3DA: $n_{excellent} = 23$, $n_{good} = 7$; 3D-DSA: $n_{excellent} = 23$, $n_{good} = 7$).

Qualitative and Quantitative Assessment of 3D-DSA and 3DA Reconstructions
VGI. In all 3DA and 3D-DSA reconstructions, measurement of vessel diameters was successfully performed ($n_{3DA} = 60$, $n_{3D-DSA} = 60$), and the acquired values for AVMs, dAVFs, and CAs did not show significant differences. Concordantly, the corresponding VGIs did not differ significantly as well. See Table 2 for details.

Table 2: VD and VGI

Parameter	3DA (Mean)	3D-DSA (Mean)	r	P
$VD1_{AVM}$	3.82 (SD, 0.47) mm	3.81 (SD, 0.55) mm	0.988	<.001
$VD2_{AVM}$	3.97 (SD, 0.60) mm	3.94 (SD, 0.59) mm	0.984	<.001
VGI_{AVM}	0.96 (SD, 0.04) mm	0.97 (SD, 0.03) mm	0.835	.003
$VD1_{CA}$	4.39 (SD, 0.84) mm	4.34 (SD, 0.87) mm	0.998	<.001
$VD2_{CA}$	4.61 (SD, 0.81) mm	4.54 (SD, 0.85) mm	0.996	<.001
VGI_{CA}	0.95 (SD, 0.04) mm	0.95 (SD, 0.03) mm	0.955	<.001
$VD1_{dAVF}$	4.24 (SD, 0.72) mm	4.22 (SD, 0.77) mm	0.993	<.001
$VD2_{dAVF}$	4.44 (SD, 0.780) mm	4.43 (SD, 0.81) mm	0.992	<.001
VGI_{dAVF}	0.95 (SD, 0.02) mm	0.95 (SD, 0.02) mm	0.824	.003
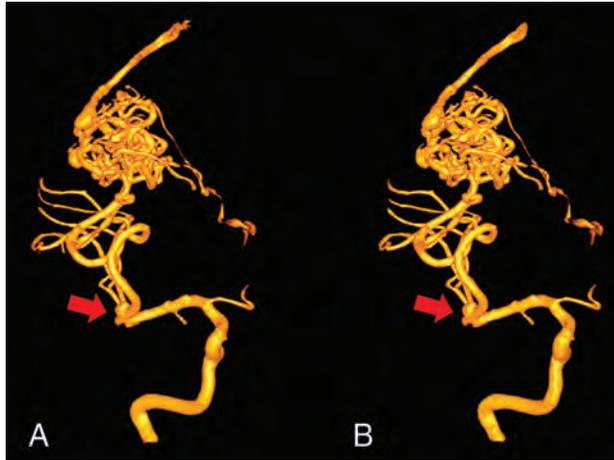
FIG 3. Illustrative case 1. Sample visualization of a right-sided AVM with 3D-DSA (*A*) and the AI-based 3DA (*B*) using VRT. 3DA (*B*) shows both feeding arteries originating from the MCA and the nidus and drainage via superficial veins, equivalent to 3D-DSA. Not only the flow-associated aneurysm of the MCA bifurcation (*red arrow*) but also the venous ectasia adjacent to the nidus is comparably visualized without loss of information using 3DA.
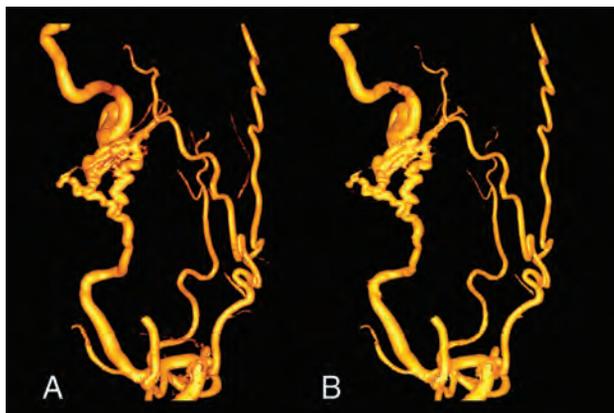
FIG 4. Illustrative case 2. Sample visualization of a left-sided dAVF with 3D-DSA (*A*) and the AI–based 3DA (*B*) using VRT. 3DA (*B*) offers comparable visualization of the influx of the fistula via the dilated middle meningeal artery and the drainage of the fistula via an ectatic cerebral vein compared with 3D-DSA (*A*).

AVMs. In total, 10 data sets acquired via the ICA ($n_{right} = 4$, $n_{left} = 1$) and vertebral artery ($n_{right} = 3$, $n_{left} = 2$) and having 10 AVMs were successfully postprocessed and evaluated. Qualitative

assessment of the corresponding 3DA and 3D-DSA data sets revealed identical results concerning the involvement of eloquent brain areas ($n_{3DA/3D-DSA} = 8$), origin of the main feeder (3DA/3D-DSA: $n_{posterior\ cerebral\ artery} = 6$, $n_{MCA} = 3$, $n_{anterior\ cerebral\ artery} = 1$), and drainage (3DA/3D-DSA: $n_{superficial} = 4$, $n_{deep} = 4$, $n_{mixed} = 2$) and pathologies of venous vessels (3DA/3D-DSA: $n_{venous\ pouch} = 4$). The number of arterial feeders ($n_{3DA/3D-DSA} = 16$) and associated flow-related ($n_{3DA/3D-DSA} = 3$) and intranidal aneurysms ($n_{3DA/3D-DSA} = 3$) was rated identically for both modalities. Measurement of the nidus size in 3DA reconstructions (mean, 19.9 [SD, 10.9] mm) showed a strong correlation ($r = 0.9/P = .001$) to 3D-DSA (mean, 20.2 [SD, 11.2] mm). The Spetzler-Martin score was rated identically for AVMs in 3DA and 3D-DSA reconstructions (3DA/3D-DSA: $n_{Spetzler-Martin\ score\ 1} = 1$, $n_{Spetzler-Martin\ score\ 2} = 4$, $n_{Spetzler-Martin\ score\ 3} = 5$). For an illustrative case see Fig 3.

dAVFs. In total, 10 data sets, acquired via the ICA ($n_{right} = 2$, $n_{left} = 1$), external carotid artery ($n_{right} = 2$, $n_{left} = 2$), and vertebral artery ($n_{right} = 1$, $n_{left} = 2$) with 10 dAVFs were successfully postprocessed and evaluated. Qualitative assessment of the corresponding 3DA and 3D-DSA data sets revealed identical results concerning the origin of the main feeder (3DA/3D-DSA: $n_{ophthalmic\ artery} = 3$, $n_{middle\ meningeal\ artery} = 3$, $n_{posterior\ meningeal\ artery} = 3$, $n_{ascending\ pharyngeal\ artery} = 1$), localization of the fistulous point (3DA/3D-DSA: $n_{anterior\ cranial\ fossa} = 2$, $n_{middle\ cranial\ fossa} = 4$, $n_{posterior\ cranial\ fossa} = 4$), primary vessel of drainage (3DA/3D-DSA: $n_{vein} = 9$, $n_{sinus} = 1$), and drainage of the fistula (3DA/3D-DSA: $n_{transverse\ sigmoid\ sinus} = 3$, $n_{superior\ sagittal\ sinus} = 3$, $n_{straight\ sinus} = 4$). Measurement of the maximum diameter of the main feeder in the 3DA reconstructions (mean, 2.04 [SD, 0.65] mm) showed a strong correlation ($r = 0.9/P = .001$) to 3D-DSA (mean, 2.05 [SD, 0.63] mm). For an illustrative case see Fig 4.

CAs. In total, 10 data sets, acquired via the ICA ($n_{right} = 4$, $n_{left} = 5$) and vertebral artery ($n_{right} = 1$) and having 10 CAs, were successfully postprocessed and evaluated. Qualitative assessment of the corresponding 3DA and 3D-DSA data sets revealed identical results concerning the identification of the parent vessel (3DA/3D-DSA: $n_{Aanterior\ communicating\ artery} = 2$, $n_{MCA} = 4$, $n_{ICA} = 3$, $n_{basilar\ artery} = 1$) and the aneurysmal neck configuration (3DA/3D-DSA: $n_{small} = 3$, $n_{medium} = 2$, $n_{large} = 5$). Measurement of the maximum diameter of the aneurysmal sac in the 3DA reconstructions (mean, 5.17 [SD, 3.4] mm) showed a strong correlation ($r = 0.9/P = .001$) to 3D-DSA (mean, 5.12 [SD, 3.3] mm). The number of aneurysmal blebs was equivalent in both types of reconstructions (3DA/3D-DSA: $n = 3$). For an illustrative case see Fig 5.

DISCUSSION

3D imaging of neurovascular pathologies is invaluable in diagnostics and therapy. So far, 3D-DSA with an established acquisition protocol and efficient postprocessing is regarded as standard. Approximately 50% of the radiation dose required for a 3D-DSA
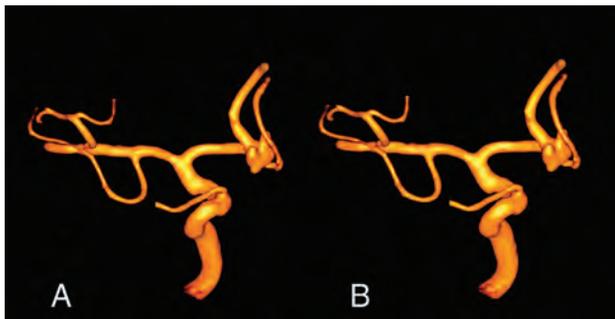
FIG 5. Illustrative case 3. Sample 3D visualization of an irregular aneurysm of the anterior communicating artery in VRT reconstructions using 3D-DSA (*A*) and 3DA (*B*). Despite differences concerning the reconstruction algorithm, AI-based 3DA (*B*) provides all relevant information on the aneurysmal configuration equivalent to 3D-DSA.

acquisition is necessary for acquisition of a mask; alone, the mask contains no clinically useful vascular information. Moreover, the subtraction technique is susceptible to artifacts due to misregistration of the mask and fill runs. Thus, AI-based algorithms like 3DA requiring only a single contrast-enhanced run have been developed for obtaining a maskless 3DA. This reduces the radiation dose and may also improve image quality by eliminating the eventuality of intersweep motion. Because AI-based algorithms have already been validated for vessels without pathologic findings[18] and for CAs,[17] the potential of this technique for generating DSA-like 3DAs with diagnostic quality has already been demonstrated. However, 3D imaging is needed most in AVMs and dAVFs. Hence, an assessment of 3DA with regard to complex pathologies is imperative before clinical implementation.

In our series, 3DA as a novel prototypical postprocessing technique was performed successfully for all data sets by differentiating 2 types of tissue (vasculature versus nonvasculature) and was suitable for visualization of CAs, AVMs, and dAVFs. When we took specific characteristics of these pathologies into account, our quantitative and qualitative analyses of 3DA demonstrated excellent agreement with standard 3D-DSA reconstructions. Even though 3DA has not left the prototype stage yet, future clinical application seems promising because no loss of information has been observed. As a result of the clinical implementation of 3DA, the effective patient dose required for 3D imaging of the cerebral vasculature would be reduced by approximately 50%: Based on usual dose values for a 3D-DSA acquisition ranging from 0.3 to 0.9 mSv (depending on the extent of collimation), 0.15–0.45 mSv per acquisition may be saved, respectively. Although an estimated dose reduction below average values of 0.35 mSv per acquisition might not substantially impact a patient's associated lifetime cancer risk, reliable dose savings would be achievable through the use of 3DA.

Previously, Montoya et al[17] have demonstrated that their AI-based algorithm (3D deep learning angiography [3D-DLA]) precisely extracts the vascular anatomy in a DSA-like manner in a series of 62 data sets. The authors analyzed the quality of vessel separation from adjacent bone and soft tissue in 3D-DLA by quantitative and qualitative parameters in cases with cerebrovascular abnormalities. 3D-DLA provided a vasculature classification accuracy of 98.7%, high-quality bone removal, and reduced misregistration artifacts caused by intersweep motion. Although

the authors provided image examples of CAs, there was no statement regarding the exact composition of their patient cohort (cases were selected in a random fashion). Moreover, a dedicated assessment of CAs concerning qualitative (eg, configuration) and quantitative (eg, size) parameters has not been performed. Nevertheless, the authors demonstrated that 3D-DLA is applicable to vascular pathologies. Because the analysis of Montoya et al[17] exclusively allows conclusions about CAs, our detailed evaluation covering CAs, AVMs, and dAVFs is another important component to assess the value of AI-based algorithms in this context and to prepare their future clinical implementation. However, all AI-based algorithms depend directly on the chosen network architecture. Provided that appropriate data sets (in particular 3D-DSA data sets of adequate quality) have been used for training the network, we fully agree that AI-based algorithms such as 3D-DLA or 3DA have the potential to simultaneously improve the image quality and reduce the radiation dose.

Most interesting, "simple" thresholding of a single contrast-enhanced data set is another remarkable option to generate DSA-like images for 3DRA.[13] Similar to AI-based approaches, 3DRA is not based on the subtraction technique and does not require a mask run. Consequently, the radiation dose can be almost halved compared with 3D-DSA. Simultaneously, 3DRA does not have susceptibility for misregistration artifacts or for intersweep motion and may improve image quality. So far, 3DRA has been evaluated in cases of CAs and was a valuable method for their preinterventional assessment. Because of its accuracy in these cases, the clinical implementation of 3DRA has already been achieved. However, 3DRA is based on modifying thresholds for the contrast-enhanced 3D data to provide differentiation of vascular from nonvascular structures.[14] As a consequence, poorly opacified vessels (eg, perforators) are potentially harmed by this reconstruction process and are in danger of not being visualized.[16] Therefore, we must assume that in general, 3DRA is limited regarding visualization of complex arteriovenous pathologies. Nevertheless, 3DRA significantly contributes to a reduction of the radiation dose required for the diagnostic work-up of CAs. However, our results indicate that 3DA as an AI-based algorithm is not restricted in terms of visualization of the angioarchitecture of an AVM or dAVF. In fact, our experience concerning visualization of very small vessels like perforating arteries with 3DA is in complete accordance with previously published data.[17,18] Hence, 3DA provides significant benefits because of its broad field of applications.

Also, direct manipulation of scanning parameters can be an effective concept for realizing significant dose reductions regarding 3D-DSA acquisitions. Pearl et al[11,12] demonstrated that simple downscaling of the applied dose is not necessarily associated with poor image quality. In the end, the authors worked out a low-dose protocol that both leads to a relevant dose reduction (up to 30%) and preserves diagnostic quality of the 3D-DSA data sets. However, any manipulation of scanning parameters caused statistically significant alterations regarding the representation of the vascular geometry and is potentially associated with loss of relevant information. In the clinical routine, this inconsistency is only tolerable to a certain extent and might be disadvantageous, especially in complex cases. Nevertheless, the work of Pearl et al is highly relevant for the evolution of AI-based algorithms. As soon

as we learn to combine low-dose protocols with AI-based algorithms, we will be able to take full advantage of both dose-saving techniques. Therefore, future research should focus on the development of clinically applicable dose-reduced single-run protocols.

Limitations

Although clinically useful 3DA was achieved from all data sets, our analysis has limitations. First, it was limited by the small sample size. Moreover, our retrospective series included only preoperative/preinterventional cases of vascular pathologies; it does not address the qualification of the prototypical 3DA to visualize metallic implants (eg, embolization material). Additionally, our series did not include cases with heavily compromised image quality (eg, motion artifacts) and contrast variations (eg, data sets with varying contrast media dilutions). Nevertheless, our analysis exclusively evaluated 1 specific AI-based algorithm for vessel recreation, classifying 2 types of tissue (vasculature versus nonvasculature). Thus, the results concerning the diagnostic value of the proposed 3DA algorithm can be transferred only to other AI-based angiographic algorithms with comparable characteristics. In fact, further investigations focusing on these issues are required to assess the full clinical applicability of this new method.

CONCLUSIONS

The investigated AI-based 3DA algorithm used in cases of complex AVMs, dAVFs, and CAs and obtaining DSA-like 3D volumes provided the same utility as conventionally reconstructed 3D-DSA volumes. Moreover, because this method does not rely on the subtraction technique, AI-based 3DA offers a significant reduction of the effective patient radiation dose. Thus, the clinical implementation of this promising postprocessing algorithm should be pursued.

REFERENCES

1. Wong SC, Nawawi O, Ramli N, et al. **Benefits of 3D rotational DSA compared with 2D DSA in the evaluation of intracranial aneurysm.** *Acad Radiol* 2012;19:701–07 Jun CrossRef Medline
2. van Rooij WJ, Sprengers ME, de Gast AN, et al. **3D rotational angiography: the new gold standard in the detection of additional intracranial aneurysms.** *AJNR Am J Neuroradiol* 2008;29:976–79 CrossRef Medline
3. Mistretta CA, Crummy AB, Strother CM. **Digital angiography: a perspective.** *Radiology* 1981;139:273–76 CrossRef Medline
4. Hochmuth A, Spetzger U, Schumacher M. **Comparison of three-dimensional rotational angiography with digital subtraction angiography in the assessment of ruptured cerebral aneurysms.** *AJNR Am J Neuroradiol* 2002;23:1199–205 Medline
5. Anxionnat R, Bracard S, Ducrocq X, et al. **Intracranial aneurysms: clinical value of 3D digital subtraction angiography in the therapeutic decision and endovascular treatment.** *Radiology* 2001;218:799–808 CrossRef Medline
6. Missler U, Hundt C, Wiesmann M, et al. **Three-dimensional reconstructed rotational digital subtraction angiography in planning treatment of intracranial aneurysms.** *Eur Radiol* 2000;10:564–68 CrossRef Medline
7. Ogilvy CS, Stieg PE, Awad I, et al. Special Writing Group of the Stroke Council, American Stroke Association. **AHA Scientific Statement: recommendations for the management of intracranial arteriovenous malformations—a statement for healthcare professionals from a Special Writing Group of the Stroke Council, American Stroke Association.** *Stroke* 2001;32:1458–71 CrossRef Medline
8. Oppenheim C, Meder JF, Trystram D, et al. **Radiosurgery of cerebral arteriovenous malformations: is an early angiogram needed?** *AJNR Am J Neuroradiol* 1999;20:475–81 Medline
9. Struffert T, Hauer M, Banckwitz R, et al. **Effective dose to patient measurements in flat-detector and multislice computed tomography: a comparison of applications in neuroradiology.** *Eur Radiol* 2014;24:1257–65 CrossRef Medline
10. Kim DJ, Park MK, Jung DE, et al. **Radiation dose reduction without compromise to image quality by alterations of filtration and focal spot size in cerebral angiography.** *Korean J Radiol* 2017;18:722–28 CrossRef Medline
11. Pearl MS, Torok CM, Messina SA, et al. **Reducing radiation dose while maintaining diagnostic image quality of cerebral three-dimensional digital subtraction angiography: an in vivo study in swine.** *J Neurointerv Surg* 2014;6:672–76 CrossRef Medline
12. Pearl MS, Torok C, Katz Z, et al. **Diagnostic quality and accuracy of low dose 3D-DSA protocols in the evaluation of intracranial aneurysms.** *J Neurointerv Surg* 2015;7:386–90 CrossRef Medline
13. Grass M, Koppe R, Klotz E, et al. **Three-dimensional reconstruction of high contrast objects using C-arm image intensifier projection data.** *Comput Med Imaging Graph* 1999;23:311–21 CrossRef Medline
14. Raabe A, Beck J, Rohde S, et al. **Three-dimensional rotational angiography guidance for aneurysm surgery.** *J Neurosurg* 2006;105:406–11 CrossRef Medline
15. Gosch D, Kurze W, Deckert F, et al. **Radiation exposure with 3D rotational angiography of the skull.** *Rofo* 2006;178:880–05 CrossRef Medline
16. Hirai T, Korogi Y, Suginohara K, et al. **Clinical usefulness of unsubtracted 3D digital angiography compared with rotational digital angiography in the pretreatment evaluation of intracranial aneurysms.** *AJNR Am J Neuroradiol* 2003;24:1067–74 Medline
17. Montoya JC, Li Y, Strother C, et al. **3D deep learning angiography (3D-DLA) from C-arm conebeam CT.** *AJNR Am J Neuroradiol* 2018;39:916–22 CrossRef Medline
18. Lang S, Hoelter P, Schmidt MA, et al. **Evaluation of an artificial intelligence-based 3D-angiography for visualization of cerebral vasculature.** *Clin Neuroradiol* 2020;30:705–12 CrossRef Medline
19. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. MIT Press; 2016
20. Ioffe S, Szegedy C. **Batch normalization: accelerating deep network training by reducing internal covariate shift.** In: *Proceedings of the 32nd International Conference on Machine Learning,* Lille, France. July 7–11, 2015;37:448–56
21. Kingma D, Ba J. **Adam: a method for stochastic optimization.** In: *Proceedings of the International Conference on Learning Representations*, San Diego, California. May 7–9, 2015
22. van den Oord A, Dieleman S, Zen H, et al. **WaveNet: a generative model for raw audio.** 2016. arXiv: 1609.03499v2[cs.SD]. https://arxiv.org/abs/1609.03499?source=post_page630afcafb9dd. Accessed April 22, 2020
23. Yu F, Koltun V. **Multi-scale context aggregation by dilated convolutions.** 2016. arXiv: 1511.07122v3[cs.CV]. http://export.arxiv.org/abs/1511.07122v3. Accessed May 8, 2020

Deep Learning–Based Software Improves Clinicians' Detection Sensitivity of Aneurysms on Brain TOF-MRA

B. Sohn, K.-Y. Park, J. Choi, J.H. Koo, K. Han, B. Joo, S.Y. Won, J. Cha, H.S. Choi, and S.-K. Lee

ABSTRACT

BACKGROUND AND PURPOSE: The detection of cerebral aneurysms on MRA is a challenging task. Recent studies have used deep learning–based software for automated detection of aneurysms on MRA and have reported high performance. The purpose of this study was to evaluate the incremental value of using deep learning–based software for the detection of aneurysms on MRA by 2 radiologists, a neurosurgeon, and a neurologist.

MATERIALS AND METHODS: TOF-MRA examinations of intracranial aneurysms were retrospectively extracted. Four physicians interpreted the MRA blindly. After a washout period, they interpreted MRA again using the software. Sensitivity and specificity per patient, sensitivity per lesion, and the number of false-positives per case were measured. Diagnostic performances, including subgroup analysis of lesions, were compared. Logistic regression with a generalized estimating equation was used.

RESULTS: A total of 332 patients were evaluated; 135 patients had positive findings with 169 lesions. With software assistance, patient-based sensitivity was statistically improved after the washout period (73.5% versus 86.5%, $P < .001$). The neurosurgeon and neurologist showed a significant increase in patient-based sensitivity with software assistance (74.8% versus 85.2%, $P = .03$, and 56.3% versus 84.4%, $P < .001$, respectively), while the number of false-positive cases did not increase significantly (23 versus 30, $P = .20$, and 22 versus 24, $P = .75$, respectively).

CONCLUSIONS: Software-aided reading showed significant incremental value in the sensitivity of clinicians in the detection of aneurysms on MRA without a significant increase in false-positive findings, especially for the neurosurgeon and neurologist. Software-aided reading showed equivocal value for the radiologist.

ABBREVIATIONS: ACA = anterior cerebral artery; CAD = computer-assisted detection

Unruptured intracranial aneurysms remain a major public health concern, and their prevalence is estimated to be 3.2% in healthy adults.[1] The annual incidence of aneurysm rupture is approximately 1% and is associated with a high risk of morbidity and mortality.[2]

TOF-MRA is a widely available technique that shows high sensitivity for the detection of aneurysms.[3] It can be performed as an initial screening because it is noninvasive and requires no contrast agent or radiation.

Detection of cerebral aneurysms on MRA is a challenging task for radiologists, neurosurgeons, and neurologists. Interpretation of both source and MIP reconstructed images is recommended to achieve good sensitivity.[4] However, detecting small lesions is a difficult and time-consuming task. Moreover, there is a relative shortage of experienced radiologists, owing to an increasing demand for imaging studies.[5]

Therefore, computer-assisted detection (CAD) of aneurysms is expected to play a key role in improving detection accuracy. Various CAD software packages for cerebral aneurysms have been investigated and have shown desirable sensitivity.[6-8] Miki et al[9] reported that routine integration of CAD with MRA for aneurysms was feasible and could help radiologists find more aneurysms without any reduction in specificity.

Recently, the use of machine learning methods has led to improvement in image-classification tasks. Along with recent advances in machine learning with the emergence of convolutional neural networks, several studies have evaluated the feasibility of using deep learning–based algorithms for the automated detection of intracranial aneurysms.[10-14] A previous study used a deep learning–based CAD software for the automated detection and localization of intracranial aneurysms on MRA and validated its high sensitivity and specificity via internal and external test sets.[15]

Due to the previously mentioned relative shortage of radiologists and increasing workload, in practice, clinicians sometimes have to meet the patient and review the MRA without an official report from a neuroradiologist. Park et al[11] investigated the performance of clinicians in identifying intracranial aneurysms using CT angiography, assisted by a deep learning model. However, to the best of our knowledge, no study has yet investigated the performance improvement of a clinician in aneurysm detection on MRA using CAD software.

The present study aimed to evaluate the effect of deep learning–based CAD software for the detection of cerebral aneurysms in MRA interpretation by 2 radiologists, a neurosurgeon, and a neurologist. Therefore, our primary end point was to investigate the improvement of accuracy with the assistance of software, including subgroup analysis based on readers and size, volume, and location of the lesions.

MATERIALS AND METHODS
Patient Cohort
This single-center, retrospective study was approved by the institutional review board (Severance Hospital, Seoul, Korea), and the requirement for informed consent was waived. We used the diagnostic cohort of a previous study, which investigated whether a deep learning model can achieve a target performance comparable with that of human radiologists. The required number of examinations was calculated to be 135 for aneurysm-containing examinations and 197 for aneurysm-free examinations using a sample-size calculation formula.[16]

TOF-MRA examinations of intracranial aneurysms were extracted from January 2018 to June 2019. The inclusion criteria were as follows: 1) older than 18 years of age; 2) MRA performed using 1.5T or 3T scanners; and 3) intracranial aneurysms on a radiology report. The number of eligible MRA examinations that met the inclusion criteria was 419. The exclusion criteria were as follows: 1) nonsaccular aneurysm, such as mycotic aneurysm, dissecting aneurysm, or pseudoaneurysm ($n = 8$); 2) giant aneurysm, >25 mm in diameter ($n = 0$); 3) ruptured aneurysm ($n = 0$); 4) aneurysms treated with surgical clipping, coil embolization, or stent insertion ($n = 118$); 5) significant displacement of the intracranial vascular structure due to intracranial hemorrhage or tumor ($n = 1$); and 6) pronounced artifacts ($n = 4$). On the basis of these criteria, 288 aneurysm-containing examinations were eligible for inclusion in the diagnostic cohort.

Reference Preparation
Three neuroradiologists (with 2, 11, and 15 years of experience in neuroradiology) independently reviewed the 288 MRA examinations in consecutive registries and evaluated the number and location of aneurysms, referring to data from other imaging modalities when available, such as CT angiography or DSA. Only examinations in which all 3 neuroradiologists concurred on the number and location of aneurysms were finally included in the diagnostic cohort. Consensus was not reached in 28 aneurysm-containing examinations; therefore, these cases were excluded. All of the excluded cases demonstrated aneurysms of <2 mm in diameter with equivocal, bulging contours, which were possibly a junctional dilation or atherosclerotic luminal irregularity. Finally, 135 aneurysm-containing examinations were prepared for the diagnostic cohort.

For the aneurysm-free examinations in the diagnostic cohort, 300 aneurysm-free MRA examinations based on radiologic reports were randomly extracted. The 3 neuroradiologists reviewed the recruited examinations in chronologic order and reached a consensus that these examinations did not show any discernible aneurysms, significant vascular steno-occlusion, or significant structural abnormalities. Finally, 197 aneurysm-free examinations were selected for the diagnostic cohort. All data were anonymized.

TOF-MRA Data Acquisition
Examinations were acquired on both 1.5T (Achieva, Philips Healthcare) and 3T scanners (Achieva and Ingenia, Philips Healthcare; Trio Tim, Siemens; Discovery MR750, GE Healthcare). The following parameters were used for 3D TOF-MRA: TR, 20–24 ms; TE, 3.4 ms; flip angle, 18°–20°; FOV, 200–250 mm; section thickness, 0.5–0.7 mm; matrix, 384×310–880×637.

CAD Software
The deep learning–based CAD software used in this study was developed by our team. The details of the algorithm have been published elsewhere.[15] To develop the original study model, we randomly extracted 600 patients from our hospital from 2014 to 2016. For validation, 110 patients from another institution were prepared for the external test set. The software was developed for classification using 3D ResNet architecture (https://github.com/kenshohara/3D-ResNets-PyTorch). The patch-wise binary classification algorithm was followed by a pixel-voting algorithm, which presents only boxes that have a higher probability of the presence of an aneurysm than a certain cutoff value to reduce the number of false-positives for aneurysm detection. Finally, the CAD software presented a 1.0×1.0 cm bounding box, which was expected to contain an aneurysm with a certain probability, on final images (Online Supplemental Data). If the algorithm predicted that there would be no aneurysm on examination, no bounding box was presented.

Study Design and Statistical Analysis
Four physicians (a neurosurgeon, neurologist, neuroradiologist, and radiology resident) interpreted MRA examinations under blinded conditions. After a 1-month washout period, MRA was interpreted again with software assistance. Interpreters could see bounding boxes using the software and freely decide whether to accept or ignore them. We measured the sensitivity and specificity per patient, sensitivity per lesion, the number of false-positives per case, and the total time required for the interpretation. For sensitivity per patient,
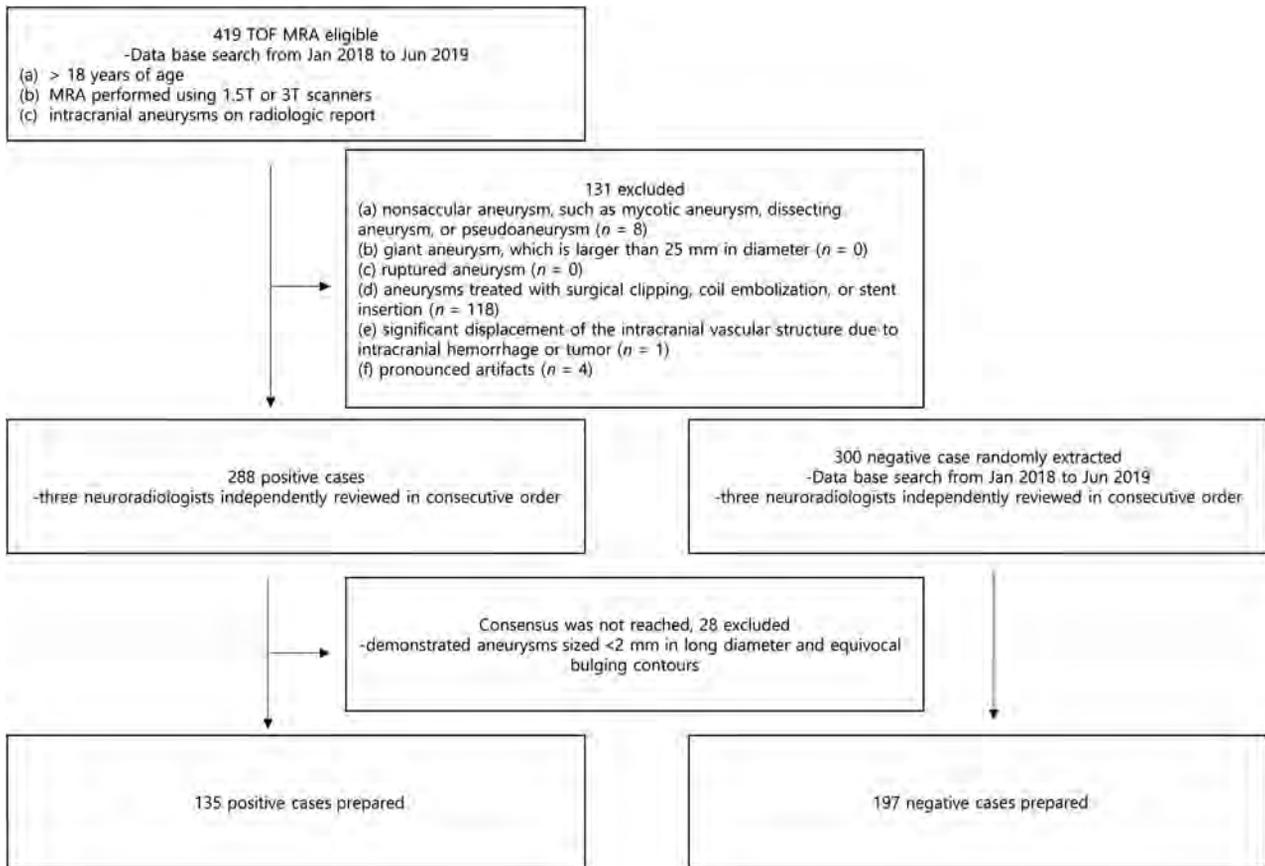
FIG 1. Patient-selection flow diagram.

Table 1: Patient demographics

	Positive for Aneurysm	Negative for Aneurysm	P Value
No. of patients	135	197	
Age (range)	62 (52–69)	62 (53–73)	$P = .5^a$
Sex (male ratio)	36 (26.7%)	84 (42.6%)	$P = .0029^b$
No. of lesions	169	0	

[a] Mann-Whitney U test.
[b] χ^2 test.

we considered true-positive cases when all true-positive lesions were detected, without any false-negative or false-positive lesions. We compared these diagnostic performances between software versus human readers, software versus humans with CAD assistance, and humans without versus humans with CAD assistance. We performed this analysis on both reader-averaged and reader-individual results. Poisson regression with the generalized estimating equation was used for the number of false-positives per case, and logistic regression with the generalized estimating equation was used for the other cases.

For characteristics analysis, we divided the 169 aneurysms into subgroups according to the diameter, volume, and location. We analyzed the sensitivity per lesion and compared it among the different subgroups. The thresholds of each subgroup in terms of size and volume were 3 and 5 mm, and 10 and 20 mL, respectively. The location was divided into 6 subgroups (Online Supplemental Data). Both reader-averaged and reader-individual comparisons

were performed. The geeglm package from the R statistical and computing software (Version 4.0.2; http://www.r-project.org) was used for the analysis (performed by K.H., statistician).

RESULTS

A total of 332 MRA examinations were collected, of which 135 had positive findings and 197 had negative findings. The 135 examinations with positive findings included 169 aneurysms, ranging from 2.0 to 17 mm in maximum diameter (mean size, 3.98 [SD, 2.11] mm) (Fig 1).

There were 84 (42.6%) and 36 (26.7%) male patients in the aneurysm-negative and aneurysm-positive cohorts, respectively. There was no statistical difference ($P = .5$, by Mann-Whitney U test) between the mean ages of the aneurysm-negative and aneurysm-positive patient cohorts (62 versus 62 years) (Table 1).

Among the 169 aneurysms, 81 lesions were <3 mm, 76 lesions were 3–5 mm, and 12 lesions were ≥5 mm. In terms of volume, 58 lesions were <10 mL, 52 lesions were 10–20 mL, and 59 lesions were ≥20 mL. In terms of location, 5 lesions were located at major branch of the anterior cerebral artery (ACA); 106, at the around-dural ring; 12, at the extradural ring; 23, at the intracranial distal ICA; 19, at the MCA major branch; and 4, at the posterior circulation.

The Online Supplemental Data present the diagnostic performance of the CAD software, reader-averaged results (with and without CAD assistance), and the statistical comparison of
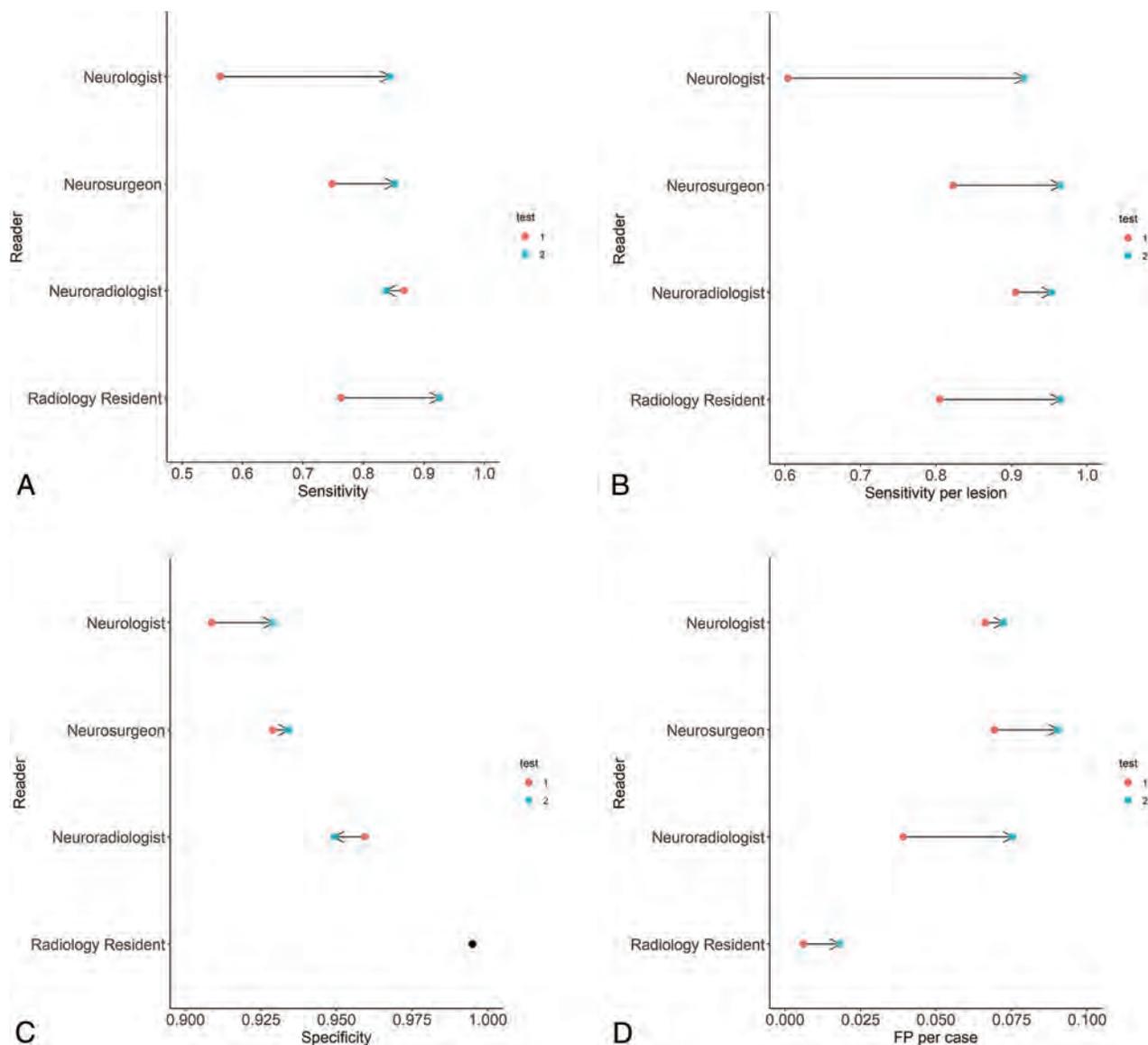
FIG 2. Change in individual diagnostic performance of clinicians. *Arrow lines* depict the change in diagnostic performance for each clinician without (*red dot*) and with (*blue dot*) software augmentation. The radiology resident showed no change in specificity between the 2 occasions (*black dot*). FP indicates false-positive.

diagnostic performance between CAD software and human readers, CAD software versus human readers (with CAD assistance), and human readers (without CAD assistance) versus human readers (with CAD assistance).

Under CAD assistance, both the patient-based sensitivity and sensitivity per lesion were statistically improved after the washout period (Fig 2). There was no difference in patient-based sensitivity between the CAD software and reader-averaged results without CAD assistance. However, when the CAD software was used, higher lesion-based sensitivity was observed compared with reader-averaged results. With CAD assistance, the patient-based sensitivity of readers was higher than that of CAD software. The number of false-positive cases was higher in readers than in the CAD software, and it increased significantly with CAD assistance in terms of reader average. There was no significant difference in specificity in any comparison.

Reader-individual diagnostic performances with/without CAD assistance are presented in Table 2.

For all readers, except the neuroradiologist, patient-based sensitivity and per-lesion sensitivity were increased significantly with CAD assistance (Fig 2). With respect to per-lesion sensitivity, neuroradiologists also showed a tendency toward increased performance with CAD assistance (95.3 versus 90.5, $P = .07$). No significant change in specificity was observed in any individual reader. The number of false-positives detected was significantly high for the neuroradiologist with CAD assistance. Other readers showed no significant changes.

The time of interpretation was 335 minutes (1.01 minute) for the neurosurgeon, 329 minutes (0.99 minute) for the neurologist, 211 minutes (0.64 minute) for the radiology resident, and 205 minutes (0.62 minute) for the neuroradiologist during the first time period. After CAD assistance, it changed to 260 minutes

Table 2: Reader-individual diagnostic performances with/without CAD assistance

	Reader	Human without CAD Assistance	Human with CAD Assistance	Human without vs with CAD Assistance (P Value)
Sensitivity (%) (95% CI)	Neurologist	56.3 (47.9–64.7)	84.4 (78.3–90.6)	<.001
	Neurosurgeon	74.8 (67.5–82.1)	85.2 (79.2–91.2)	.032
	Radiologist	86.7 (80.9–92.4)	83.7 (77.5–89.9)	.493
	Resident	76.3 (69.1–83.5)	92.6 (88.2–97)	.001
Specificity (%) (95% CI)	Neurologist	90.9 (86.8–94.9)	92.9 (89.3–96.5)	.460
	Neurosurgeon	92.9 (89.3–96.5)	93.4 (89.9–96.9)	.842
	Radiologist	95.9 (93.2–98.7)	94.9 (91.9–98)	.629
	Resident	99.5 (98.5–100)	99.5 (98.5–100)	>.999
Sensitivity per lesion (%) (95% CI)	Neurologist	60.4 (53–67.7)	91.7 (87.2–96.3)	<.001
	Neurosurgeon	82.2 (76–88.5)	96.4 (93.7–99.2)	<.001
	Radiologist	90.5 (85.7–95.3)	95.3 (92.2–98.3)	.070
	Resident	80.5 (74.6–86.4)	96.4 (93.7–99.2)	<.001
No. of false-positives per case (95% CI)	Neurologist	0.066 (0.044–0.099)	0.072 (0.048–0.108)	.751
	Neurosurgeon	0.069 (0.047–0.103)	0.09 (0.062–0.132)	.200
	Radiologist	0.039 (0.022–0.07)	0.075 (0.051–0.112)	.028
	Resident	0.006 (0.002–0.024)	0.018 (0.008–0.04)	.179
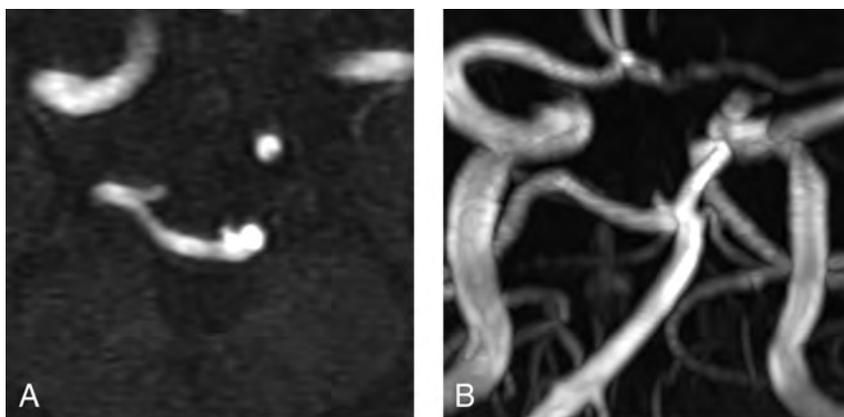
FIG 3. Source (*A*) and MIP image (*B*) from TOF-MRA of 75-year-old female patient. There was a basilar top aneurysm, <3 mm. Human readers detected this lesion, but the computer-assisted detection software did not.

(0.78 minute) for the neurosurgeon, 155 minutes (0.47 minute) for the neurologist, 184 minutes (0.55 minute) for the radiology resident, and 215 minutes (0.65 minute) for the neuroradiologist (entire dataset, with average minutes per individual case in parentheses).

The sensitivity per lesion in subgroups (the 95% confidence interval was estimated by logistic regression with the generalized estimating equation) is presented in the Online Supplemental Data. It shows a comparison of sensitivity per lesion between CAD software versus human readers, CAD software versus human readers with CAD assistance, and human readers with versus without CAD assistance. Reader-individual comparisons were also performed (Online Supplemental Data).

CAD software showed 84% lesion sensitivity for small (<3 mm and 10-mL volume) aneurysms and 100% for larger lesions. CAD software showed higher sensitivity per lesion in every subgroup of diameter and size compared with average human readers.

With CAD assistance, human readers showed improved performance in the <3 and 3–5 mm subgroups. Moreover, in every volume subgroup, human readers showed improved performance with CAD assistance.

In terms of location, average human readers showed significant increases in the sensitivity per lesion in every location, except for ACA area lesions. CAD showed higher sensitivity than an average human without CAD assistance in the subgroups of the around-dural ring, intracranial distal ICA, and MCA. In addition, CAD showed higher sensitivity than an average human with CAD assistance in the subgroups of the around-dural ring and intracranial distal ICA. However, with CAD assistance, the average human showed higher sensitivity per lesion in the subgroup of posterior circulation.

We analyzed the detection results for 169 lesions by 4 human readers (*n* = 676) and compared them with the detection results of CAD software by subgroup (Online Supplemental Data).

All lesions that the human readers detected but CAD ignored were small, <3 mm in diameter (Fig 3). Approximately 10% of the lesions that had a diameter of <3 mm that CAD missed could be detected by human readers. In terms of volume, approximately 14% of the lesions of <10 mL that CAD missed could be detected by human readers. However, lesions of >3 mm or 10 mL were never missed by CAD software, though human readers sometimes missed these lesions (Fig 4). The reader-individual comparison with CAD software is presented in the Online Supplemental Data.

DISCUSSION

Our study revealed a significant improvement in diagnostic performance of multiple physicians having different specialties in detecting cerebral aneurysms with the use of deep learning–based software. This type of software, called artificial intelligence, was developed using deep learning and achieved a target diagnostic performance comparable with that of human radiologists. Notably, this software received approval from the Korean
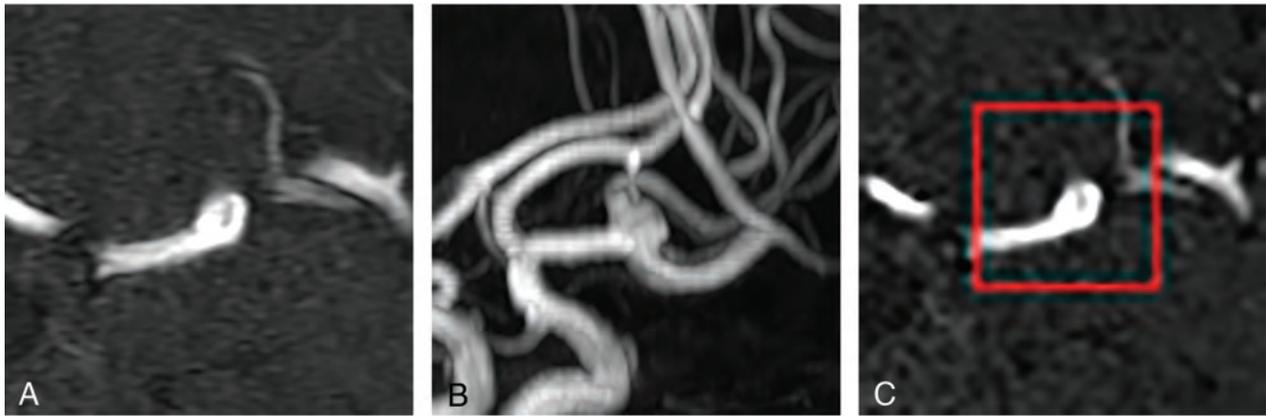
FIG 4. Source (*A*) and MIP (*B*) images from TOF-MRA of a 35-year-old female patient. Some human readers missed this lesion at first, but computer-assisted detection software detected it. *C*, During the second interpretation, human readers accepted the diagnosis of the software.

Ministry of Food and Drug Safety as an artificial intelligence–applied software.

With the assistance of this software, the neurosurgeon and neurologist involved in this study experienced significant improvement in their performance in terms of sensitivity per patient and per lesion without an increase in the number of false-positive cases. The time for interpretation was also reduced. In particular, an improvement in the detection rates for lesions of <5 mm was observed with the assistance of this software.

These results suggest a promising effect of CAD software in a real clinical environment and are compatible with the results of several previous studies. Numerous investigations have shown that deep learning–based artificial intelligence could be helpful in detecting cerebral aneurysms on CT angiography.[11] However, to the best of our knowledge, this is the first study to demonstrate the performance improvement for aneurysm detection on MRA with multiple readers.

However, the effect was unclear in the case of the neuroradiologist with software assistance. Rather, it was observed that both the number of false-positives per case and interpretation time increased. We can speculate that this is a limitation of the CAD software, which should be further improved in future work. The training set for this software was developed on the basis of the neuroradiologist's interpretation of TOF-MRA.[15] Furthermore, the previous results of deep learning–based CAD were comparable with not better than those of the radiologist.

However, it can be concluded that this type of CAD software can be helpful for neurosurgeons and neurologists. In the subgroup analysis, it was observed that the software did not miss any lesions of >3 mm. This result was consistent with that of a previous article that was validated via an external test set.[15] The accuracy of the software may not only offer time saving and safety for clinicians but may also lead to a reduction in psychological burden, because clinically significant aneurysms can be filtered by the software.

Our study is a retrospective, single-center study. We are currently working toward a prospective, multicenter study for the evaluation of the augmenting effect of this software, and it is under review by the institutional review board. Furthermore, we plan to merge this software with our PACS system and apply it real-time in a clinical environment.

Our study has several limitations. First, it was difficult to gather DSA-confirmed positive and negative cases of aneurysms with TOF-MRA. Therefore, a consensus had to be reached among the 3 neuroradiologists after consecutively gathering the MRA examinations. Although DSA is a more powerful diagnostic tool for evaluating cerebral aneurysms, a previous study has shown that in terms of sensitivity, 3T MRA is not inferior to DSA.[17]

Second, our patient cohort showed an unbalanced distribution of lesion locations. The lesions were mostly located on the ICA. The ACA, MCA, and posterior circulation region accounted for a relatively small portion. However, we gathered patients consecutively from our data base; therefore, we believe that this imbalance reflected the natural incidence of cerebral aneurysm location.

Third, the total number of false-positive predictions increased for all reviewers with CAD. While this was only significant in the neuroradiologist, considering the low incidence of aneurysms in a real clinical environment, the increase in the false-positive rate could be much higher. In our study population, the incidence of cerebral aneurysms was higher than that in the real world; therefore, the increase in false-positive rates could be underestimated.

Last, we could not apply a crossover interpretation design due to the small number of readers involved in our study. Considering the large number of cases and washout periods, the memory effect would be negligible in our study. However, we hope to have a more organized crossover design and many interpreters in our prospective, multicenter study.

CONCLUSIONS

We found that the assistance of deep learning–based software significantly improved clinicians' sensitivity, without a significant increase in false-positive cases, especially in the case of the neurosurgeon and neurologist.

REFERENCES

1. Vlak MH, Algra A, Brandenburg R, et al. **Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis.** *Lancet Neurol* 2011;10:626–36 CrossRef Medline

2. Juvela S, Korja M. **Intracranial aneurysm parameters for predicting a future subarachnoid hemorrhage: a long-term follow-up study.** *Neurosurgery* 2017;81:432–40 CrossRef Medline

3. Sailer AM, Wagemans BA, Nelemans PJ, et al. **Diagnosing intracranial aneurysms with MR angiography: systematic review and meta-analysis.** *Stroke* 2014;45:119–26 CrossRef Medline

4. Korogi Y, Takahashi M, Mabuchi N, et al. **Intracranial aneurysms: diagnostic accuracy of MR angiography with evaluation of maximum intensity projection and source images.** *Radiology* 1996;199:199–207 CrossRef Medline

5. McDonald RJ, Schwartz KM, Eckel LJ, et al. **The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload.** *Acad Radiol* 2015;22:1191–98 CrossRef Medline

6. Arimura H, Li Q, Korogi Y, et al. **Automated computerized scheme for detection of unruptured intracranial aneurysms in three-dimensional magnetic resonance angiography.** *Acad Radiol* 2004;11:1093–104 CrossRef Medline

7. Hirai T, Korogi Y, Arimura H, et al. **Intracranial aneurysms at MR angiography: effect of computer-aided diagnosis on radiologists' detection performance.** *Radiology* 2005;237:605–10 CrossRef Medline

8. Kakeda S, Korogi Y, Arimura H, et al. **Diagnostic accuracy and reading time to detect intracranial aneurysms on MR angiography using a computer-aided diagnosis system.** *AJR Am J Roentgenol* 2008;190:459–65 CrossRef

9. Miki S, Hayashi N, Masutani Y, et al. **Computer-assisted detection of cerebral aneurysms in MR angiography in a routine image-reading environment: effects on diagnosis by radiologists.** *AJNR Am J Neuroradiol* 2016;37:1038–43 CrossRef Medline

10. Nakao T, Hanaoka S, Nomura Y, et al. **Deep neural network-based computer-assisted detection of cerebral aneurysms in MR angiography.** *J Magn Reson Imaging* 2018;47:948–53 CrossRef Medline

11. Park A, Chute C, Rajpurkar P, et al. **Deep learning-assisted diagnosis of cerebral aneurysms using the HeadXNet model.** *JAMA Netw Open* 2019;2:e195600 CrossRef Medline

12. Sichtermann T, Faron A, Sijben R, et al. **Deep learning-based detection of intracranial aneurysms in 3D TOF-MRA.** *AJNR Am J Neuroradiol* 2019;40:25–32 CrossRef Medline

13. Stember JN, Chang P, Stember DM, et al. **Convolutional neural networks for the detection and measurement of cerebral aneurysms on magnetic resonance angiography.** *J Digit Imaging* 2019;32:808–15 CrossRef Medline

14. Ueda D, Yamamoto A, Nishimori M, et al. **Deep learning for MR angiography: automated detection of cerebral aneurysms.** *Radiology* 2019;290:187–94 CrossRef Medline

15. Joo B, Ahn SS, Yoon PH, et al. **A deep learning algorithm may automate intracranial aneurysm detection on MR angiography with high diagnostic performance.** *Eur Radiol* 2020;30:5785–93 CrossRef

16. Hajian-Tilaki K. **Sample size estimation in diagnostic test studies of biomedical informatics.** *J Biomed Inform* 2014;48:193–204 CrossRef Medline

17. Mine B, Pezzullo M, Roque G, et al. **Detection and characterization of unruptured intracranial aneurysms: comparison of 3T MRA and DSA.** *J Neuroradiol* 2015;42:162–68 CrossRef

Alterations in the Magnetoencephalography Default Mode Effective Connectivity following Concussion

D.D. Reddy, E.M. Davenport, F.F. Yu, B. Wagner, J.E. Urban, C.T. Whitlow, J.D. Stitzel, and J.A. Maldjian

ABSTRACT

BACKGROUND AND PURPOSE: Magnetoencephalography is sensitive to functional connectivity changes associated with concussion. However, the directional influences between functionally related regions remain unexplored. In this study, we therefore evaluated concussion-related magnetoencephalography-based effective connectivity changes within resting-state default mode network regions.

MATERIALS AND METHODS: Resting-state magnetoencephalography was acquired for 8 high school football players with concussion at 3 time points (preseason, postconcussion, postseason), as well as 8 high school football players without concussion and 8 age-matched controls at 2 time points (preseason, postseason). Time-series from the default mode network regions were extracted, and effective connectivity between them was computed for 5 different frequency bands. The default mode network regions were grouped into anterior and posterior default mode networks. The combined posterior-to-anterior and anterior-to-posterior effective connectivity values were averaged to generate 2 sets of values for each subject. The effective connectivity values were compared using a repeated measures ANOVA across time points for the concussed, nonconcussed, and control groups, separately.

RESULTS: A significant increase in posterior-to-anterior effective connectivity from preseason to postconcussion (corrected P value $= .013$) and a significant decrease in posterior-to-anterior effective connectivity from postconcussion to postseason (corrected P value $= .028$) were observed in the concussed group. Changes in effective connectivity were only significant within the delta band. Anterior-to-posterior connectivity demonstrated no significant change. Effective connectivity in the nonconcussed group and controls did not show significant differences.

CONCLUSIONS: The unidirectional increase in effective connectivity postconcussion may elucidate compensatory processes, invoking use of posterior regions to aid the function of susceptible anterior regions following brain injury. These findings support the potential value of magnetoencephalography in exploring directional changes of the brain network following concussion.

ABBREVIATIONS: EC = effective connectivity; EEG = electroencephalography; DMN = default mode network; FC = functional connectivity; GC = Granger causality; MEG = magnetoencephalography; mTBI = mild traumatic brain injury

In the United States alone, there are an estimated 1.6–3.8 million cases of sports-related concussions each year.[1] Depending on severity, concussions can lead to cognitive and other neurologic deficits.[2] The complications that arise from these injuries, which are of particular concern in the developing pediatric brain,

emphasize the need for accurate diagnosis and appropriately tailored management. Traditional practices that base treatment on self-evaluations and neuropsychological tests are limited, given the potential subjectivity of these measurements.[3] Unfortunately, conventional structural imaging modalities such as CT and MR imaging often lack sufficient sensitivity to detect abnormalities.[4]

Given these limitations, there has been increased attention in the development and use of adjunctive noninvasive tools. It is hypothesized that after a concussion, brain function may be altered through the disruption of functional brain networks.[5] Noninvasive functional neuroimaging techniques designed to interrogate these brain networks, including magnetoencephalography (MEG), electroencephalography (EEG), and fMRI, are, therefore, well-suited for such an evaluation.[3]

fMRI indirectly measures alterations in brain function on the basis of fluctuations in the blood oxygenation level caused by neuronal activity and can be corrupted by neurovascular decoupling. In contrast, MEG bypasses the hemodynamic response entirely by directly measuring the magnetic fields induced by neuronal electrical activity in the brain.[6-9] While EEG is similarly a more direct measure of neuronal activity than fMRI, the electrical signals are contaminated by interference from the conductive and inductive tissues of the brain, skull, and the scalp. The magnetic fields detected by MEG are not affected by these conductive and inductive effects and thereby maintain better signal purity.[6,7] Recent advances, including higher sensor density and enhanced source reconstruction techniques, have improved the spatial resolution of MEG relative to EEG and are comparable with those of fMRI.[7,10] Altogether, its high sensitivity and ability to localize pathologic magnetic signals generated by injured brain tissue, even with minimal changes in neuropsychologic testing, make MEG well-suited to assess changes in brain function related to sports-related concussion.[4,6,11]

Several resting-state MEG-based connectivity analyses have revealed alterations in the brain connectivity profiles of concussed subjects and those with mild traumatic brain injury (mTBI).[5,12,13] Various analytic approaches including amplitude envelope correlations and phase-locking estimates have been used in the past to derive connectivity metrics.[13-15] Analysis of functional brain architectures can be separated into functional segregation (the identification of a specific anatomic region with a specific function) and functional integration of many specialized cortical regions supporting a single function. Functional integration can be characterized by temporal correlation metrics, including functional connectivity (FC) or effective connectivity (EC). FC tests for coupling between 2 spatially remote brain regions without considering the underlying dynamic process in which causes precede consequences and thus only partly explains the interaction pattern.[16] EC, on the other hand, is defined by the causal influence that one brain region exerts over another and provides both correlation and directional information, revealing the flow of activity between 2 brain regions. Because coupling in the brain is both directed and reciprocal, functional integration can, therefore, be better understood with EC.[17]

EC can be estimated using 2 main approaches: 1) model-driven, such as dynamic causal modeling and structural equation modeling, and 2) data-driven, such as Granger causality (GC). Both methods analyze the temporal properties of signals from 2 brain regions and extract information to compute the directed causal influences between them. In this study, we used GC analysis to measure EC for 2 main reasons: 1) Unlike the structural equation or dynamic causal modeling, GC does not need an a priori hypothesis of the existing theoretic framework or EC structure to estimate the temporal precedence, and 2) because GC is applied to MEG data, its inability to account for the relationship between blood oxygen level–dependent signal and neural activity can be ignored.[17,18]

Although FC disruptions in concussed individuals have been studied using MEG,[12-14] information regarding the directional connectivity changes are rare and remain largely unexplored in this population. We were, therefore, motivated to investigate EC

changes. A suitable brain network for analysis would ideally be sensitive to physical trauma affecting different brain locations. The default mode network (DMN) is one such network, having high metabolic demands and important cognitive functions spanning multiple brain regions, and thus is likely to reflect the impact of such insults on the brain at large.[19,20] Our objective was to study the differences in the EC within the DMN between anterior and posterior DMN regions in concussed, nonconcussed, and age-matched controls. In addition to early postconcussion changes, we also sought to interrogate the late changes that may reflect recovery. Moreover, we hypothesized that like FC,[21] EC would exhibit changes within the DMN after concussion and that these changes would gradually normalize during the course of recovery.

MATERIALS AND METHODS
Protocol Summary
All participants included were part of the imaging Telemetry and Kinematic modeLing (iTAKL) study[22] approved by the institutional review board at Wake Forest Baptist Medical Center and the University of Texas Southwestern. The study was performed with written informed consent by the parents as well as the participants and was Health Insurance Portability and Accountability Act–compliant. The iTAKL study focused on the effects of subconcussive impacts on the brain of high school football players through an entire football season, including preseason and postseason MR imaging, MEG, and cognitive image acquisitions. As a part of this larger study on subconcussive impacts, players who were diagnosed with concussion during the season received an additional postconcussion MR imaging and MEG imaging.

Eight high school football players with concussion, 8 high school football players without concussion, and 8 age-matched controls were included in this study. The football players were from a local high school football team, and the controls were from local noncontact sports teams (swimming and tennis) with no history of concussion or contact sports. The concussed football players ranged between 15.3 and 17.7 years of age with a mean age of 16.1 (SD = 0.7) years; the nonconcussed football players ranged from 15.4 to 16.8 years of age with a mean age of 16.2 (SD = 0.6) years; and the controls ranged from 14.9 to 17.1 years of age with a mean age of 15.7 (SD = 0.8) years. Controls and the nonconcussed group were ±1 month age-matched and ±0.1 body mass index–matched to the concussed group. All participants in this study were male. If a concussion was suspected by the on-site certified athletic trainer, the player was sent to a sports medicine physician to confirm the diagnosis. The players underwent MR imaging and MEG scans at preseason, postseason, and a clinical diagnosis of concussion (postconcussion). The preseason scans were acquired a few days to 1 month in advance of the first practice session, and the postseason scans were acquired within 1 month of the end of football season. For the postconcussion scans, an interval of 1–8 days postinjury was allowed. All concussed players followed return-to-play protocols and were scanned postseason. The mean time between preseason and postconcussion scans was 69 (SD = 6.46) days, and the mean time between preseason and postseason scans was 135.2 (SD = 23.7) days. Controls underwent scans at 2 different
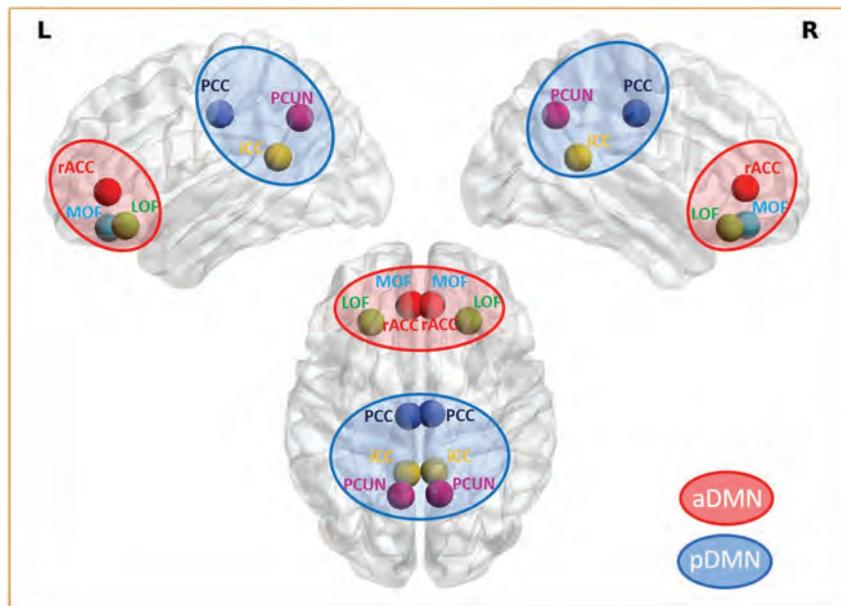
FIG 1. ROIs. DMN regions including the lateral orbitofrontal (LOF), medial orbitofrontal (MOF), and rostral anterior cingulate (rACC) regions grouped into anterior DMN regions are highlighted in red, and the precuneus (PCUN), posterior cingulate (PCC), and isthmus cingulate (iCC) regions grouped into posterior DMN regions are highlighted in blue. aDMN indicates anterior DMN; pDMN, posterior DMN.

time points approximately 4 months apart, with the mean time between scans being 127.4 (SD = 25.2) days.

Exclusion criteria included a previous history of concussion, reported neurologic disease, developmental disorders, or psychological conditions. Of the 11 concussed players identified from the iTAKL study, 3 were excluded, including 2 who did not have a pre-/postseason scan and a third who had an abnormal finding on brain MR imaging.

Data Acquisition and Preprocessing

Resting-state MEG signals sampled at a rate of 1200 Hz were recorded using a 273-channel radial gradiometer and a 29-channel reference sensor whole-head CTF Omega 2005 MEG system (VSM MedTech) for 8 minutes. Subjects had their eyes open and fixated on a point throughout the data acquisition, and the acquisition bandwidth was set to 0.25–150 Hz. A 3T Magnetom Skyra MR imaging scanner (Siemens) was used to acquire MR imaging data using a 32-channel high-resolution human head/neck coil (Siemens). A structural 3D volumetric T1-weighted MPRAGE sequence with an isotropic resolution of 0.9 mm (TR = 1900 ms; TE = 2.93 ms; TI = 900 ms; flip angle = 9°; 176 slices) was acquired. Fiducial markers were placed on the subjects' nasion and bilateral preauricular regions to enable the coregistration of MEG data with the MR imaging. The T1 images were normalized to the standard Montreal Neurological Imaging space using the Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL, part of SPM) toolbox high-dimensional warping and the SPM8[23] (http://www.fil.ion.ucl.ac.uk/spm/software/spm12) new segment procedure, as implemented in the VBM8 toolbox[23] (http://dbm.neuro.uni-jena.de/vbm.html_) and subsequently coregistered with the MEG data.

Acquired MEG data were processed using the Brainstorm[24] toolbox in Matlab (MathWorks). The MEG signals were baseline-corrected, downsampled to 250 Hz, notch-filtered at 60 Hz, and bandpass-filtered between 1 and 100 Hz. The data were further divided into 5 different frequency bands consisting of delta (2–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), and gamma (30–50 Hz). Eye blinks and muscle artifacts were identified and removed using Independent Component Analysis. Noise covariance was calculated using an empty room scan. After preprocessing, time recordings from 60 to 240 seconds (3 minutes) for each subject were used to ensure that stable and quality data were included for the subsequent processing and analysis. A cortex surface head model with overlapping spheres was then constructed for the 3 minutes of clean data for each frequency band. Sources were computed for the MEG data by projecting the signal onto the standard source space model using the whitened and depth-weighted linear-L2 minimum norm estimates algorithm.[25] Finally, ROIs were extracted for the effective connectivity analysis.

ROI

The Desikan-Killiany[26] cortical atlas partitions the whole cortex into 68 regions. Fourteen bilateral cortical DMN regions were used, including lateral orbitofrontal, medial orbitofrontal, rostral anterior cingulate, parahippocampal, isthmus cingulate, posterior cingulate, and precuneus.[27] For a better understanding of the directionality of connectivity on a large scale, the DMN regions were grouped into anterior and posterior DMN. Anterior DMN included both left and right lateral orbitofrontal, medial orbitofrontal, and rostral anterior cingulate regions, while the posterior DMN included both the left and right isthmus cingulate, posterior cingulate, and precuneus regions as shown in Fig 1. All the regions of the anterior DMN as per the Montreal Neurological Institute coordinates had Y > 16 as the centroid of those regions, and the posterior DMN consisted of regions with centroid Y < −40. The parahippocampal region, which did not fall under either the anterior DMN or posterior DMN, was not included in the connectivity analysis.[28] A mean time-series was extracted from the remaining 12 DMN regions using the Desikan automated parcellation approach, and the EC analysis was then performed between them.

Effective Connectivity and Granger Causality

EC estimating the strength and directionality of causal relationships between brain regions can be assessed using GC. GC is a
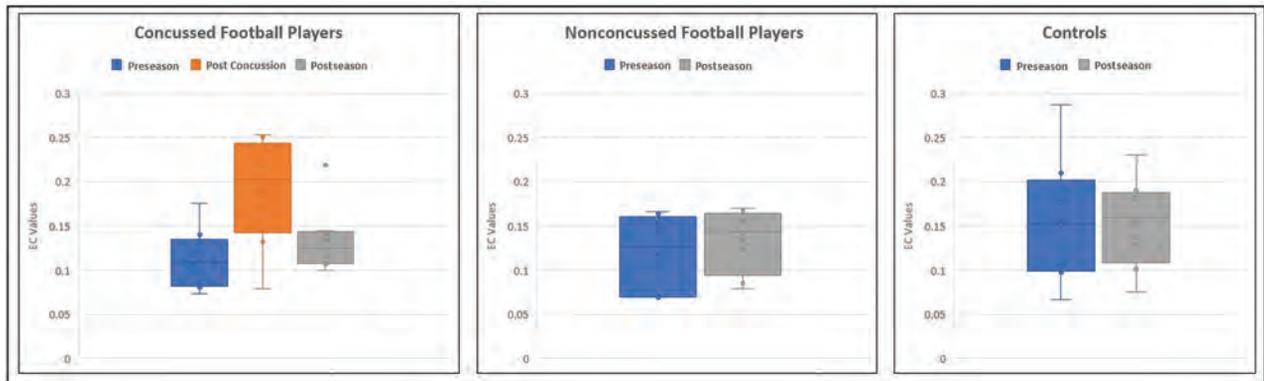
FIG 2. Comparison of effective connectivity value changes across time. EC values at preseason, postconcussion, and postseason for the concussed group (*left*), EC values at preseason and postseason for the nonconcussed group (*center*), and EC values at preseason and postseason in controls (*right*).

statistical approach that measures the extent to which one time-series signal plays a role in predicting the other. Assuming that the time-series of one ROI is represented by X and the other ROI is represented by Y, the principle of GC can be defined as Y causes X if Y is able to better predict X by including all previous information about Y compared with using previous information of X alone.[29] Two signals simultaneously acquired and of the same length can be represented as a bivariate linear autoregressive model in the time domain as follows:

$$X(t) = \sum j = 1 : pA11 \ X(t-1) + \sum j = 1$$
$$: pA12 \ Y(t-j) + E1(t),$$

$$Y(t) = \sum j = 1 : pA21 \ X(t-1) + \sum j = 1$$
$$: pA22 \ Y(t-j) + E2(t),$$

where p represents the maximum lag in the autoregressive model known as the model order, A comprises the estimated coefficients of the model, and E1 and E2 are the residuals for each time-series. On the basis of this model, if the variance of the E1 prediction error term is reduced by including the Y term in the first equation, then Y Granger causes X, and if the E2 predicted error is reduced with the inclusion of the X term in the second equation, then it indicates that X Granger causes Y. The optimum model order was set to a default of 10. The autoregressive model can be extended to the MEG time-series data from all the 12 ROIs to generate an 12×12 asymmetric values matrix, which has different measures for X causing Y and Y causing X. The GC will result in a value of zero if there are no causal influences present and as non-negative values otherwise.[30]

Statistical Analysis
An effective connectivity matrix using the GC model was obtained for each subject for all 5 different frequency bands. The normalized connectivity values for posterior-to-anterior DMN regions and anterior-to-posterior DMN regions were averaged to generate 2 sets of values for each of the 5 frequency bands per subject. The connectivity values between the ROIs within the DMN providing the same region connectivity, ie, posterior-to-other posterior and anterior-to-other anterior DMN regions,

were not considered for the analysis. A repeated measures ANOVA was used to compare the means across different time points for each combination of frequency band and connectivity orientation within the concussed, nonconcussed, and controls group separately. The resulting *P* values were adjusted using a Bonferroni multiple comparisons correction. Post hoc pair-wise testing with a Bonferroni correction between time points was performed with only the overall adjusted *P* value < .05. Differences in preseason scans were also compared for all 3 subject groups. All statistical analysis was performed using R Studio (http://rstudio.org/download/desktop).

RESULTS
A total of 36 EC values comprising the posterior-to-anterior DMN connectivity for both cerebral hemispheres were averaged. The same process was carried out to obtain the anterior-to-posterior DMN EC value for each subject. Repeated measures ANOVA revealed significant differences across different time points for the posterior-to-anterior DMN connectivity within the delta frequency band only in the concussed group. The EC values were plotted across the season for the concussed, nonconcussed, and control groups. Figure 2 is a box-and-whisker plot demonstrating the spread of EC values for all 8 subjects at each time point for the concussed, nonconcussed, and control subjects. The concussed group had an increase in the EC between the pre- and postconcussion scans for all subjects. The increased postconcussion connectivity is followed by a decrease in the connectivity postseason, partially returning to the level of the baseline (preseason) scan. Only the posterior-to-anterior connectivity showed significant differences. Of all the frequency bands, significant EC differences were noted only for the delta band. Also, the spread of EC values for all the subjects in the nonconcussed and control groups did not vary between time points.

Effective Connectivity in Concussed Subjects
Repeated measures ANOVA was performed for each combination of band and orientation, generating 10 *P* values for the concussed group. Only the delta band showed significant Bonferroni-corrected differences in the posterior-to-anterior EC (corrected *P* value = .005). The post hoc tests revealed a
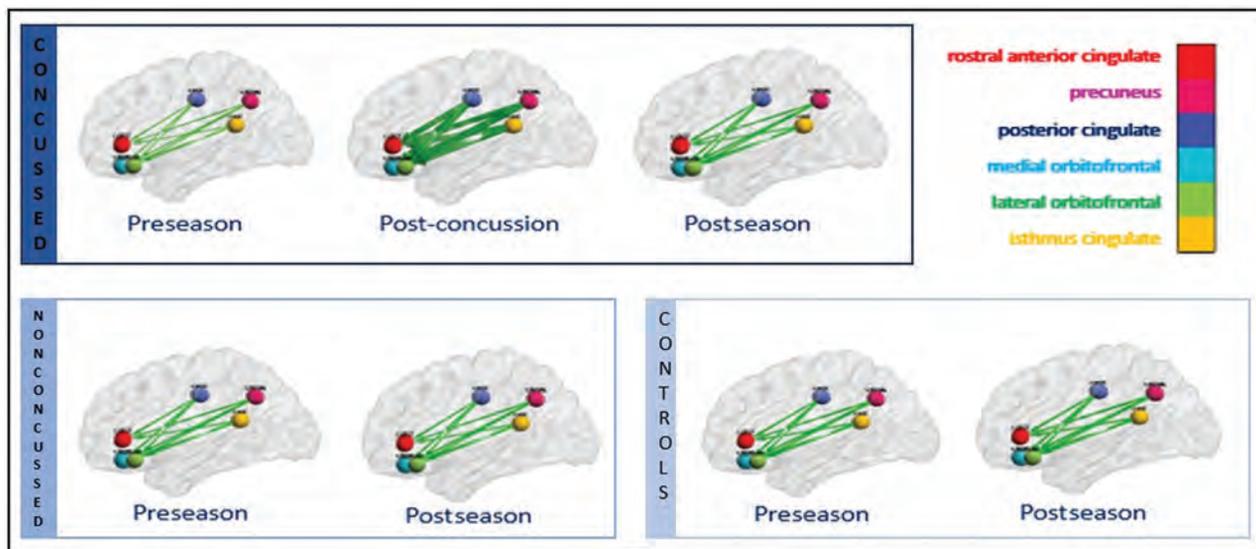
FIG 3. Visual representation of averaged EC across subjects for the concussed (*upper row*), nonconcussed (*lower left*), and control (*lower right*) groups. The concussed group demonstrates a unidirectional increase in EC from preseason to immediately postconcussion ($P = .013$) and a decrease at postseason scan compared with the postconcussion scan ($P = .028$). The nonconcussed group and controls reveal no significant difference in EC between preseason and postseason scans.

significant increase in the mean EC from preseason to postconcussion for the concussed players (corrected P value $= .013$). A consistent increase in EC was observed for each subject, with the average percentage increase for the group at approximately 8% followed by a significant decrease in the effective connectivity from postconcussion to the postseason scans (corrected P value $= .028$), at approximately 6% decrease. Despite the decreases at the postseason scan, the effective connectivity remained significantly different compared with the preseason scan (corrected P value $= .028$). Figure 3 provides a visual representation of the longitudinal unidirectional increase and decrease of EC throughout the season, averaged across all 8 subjects. Figures 1 and 3 were generated using BrainNet Viewer, a matlab-based network visualization toolbox (www.nitrc.org/projects/bnv/).

Of note, one of the concussed subjects who was not included in the study due to lack of a postseason scan exhibited similar results between the preseason baseline and postconcussion scan, with an increase of 5.9% in the total posterior-to-anterior connectivity.

Effective Connectivity in Nonconcussed Subjects and Controls

Repeated measures ANOVA revealed no significant EC difference between preseason and postseason scans for the nonconcussed football players (P value for delta $= .392$). The averaged network for the controls did not show any major changes in the connectivity strength between the pre- and postseason scans either (P value for delta $= .947$) for any combination of band and orientation.

The preseason scans compared across different subject groups using independent 2-sample t tests demonstrated no significant changes between the concussed to nonconcussed groups (P value $= .748$), the nonconcussed group to controls (P value $= .289$), and the concussed group to controls (P value $= .095$).

The Post-Concussion Symptom Inventory survey scores account for symptoms in 20 categories and were available for

only 5 of the 8 concussed subjects who were used for the analysis. The average Post-Concussion Symptom Inventory scores were 2.2 for preseason, 27 for postconcussion, and 6.4 for postseason. Although we did not have the post-concussion symptom scores for all our subjects, the correlation between the average EC changes and the average Post-Concussion Symptom Inventory scores across the season had a high R^2 value of 0.98.

DISCUSSION

Alterations in brain FC have been studied in patients with mTBI and concussions[13,31,32] using graph theory, power spectral density, and correlation analysis. However, these analyses did not use methods for providing both magnitude and directionality information. Here, we set out to explore whether an MEG-based EC method could capture variations in directed connectivity in concussed individuals, nonconcussed individuals, and matched controls. Using GC, we found that EC increased significantly in high school football players immediately after concussion followed by a significant decrease observed at the postseason time point. The EC at the postseason time point remained significantly increased compared with the preseason baseline, suggesting only a partial recovery. Additionally, EC in the DMN changed only in 1 direction, from posterior to anterior, and only affected the delta band. EC showed no significant changes in the nonconcussed individuals (both controls and football players) through the season.

Increased brain connectivity or hyperconnectivity is believed to be a fundamental adaptation of neuronal systems to physical disruptions resulting from brain injury, referred to as network reorganization.[33] Studies suggest that connectivity increases may represent a compensatory mechanism of the brain in response to the functional impairment caused by axonal injury, invoking the use of more connections to perform the same tasks and resulting in increased synaptic activity measured by MEG.[34-36] EC increases have also been attributed to the regulation of dysfunctional neural

systems and compensatory processes in conditions like bipolar disorder[37] and Alzheimer disease.[38] In some cases, EC increases have also been associated with improved corticocortical interactions in the damaged hemispheres of patients with stroke, facilitating recovery.[39] The EC increases in our concussed subjects support the hypothesis of a compensatory mechanism.

Most interesting, significant EC increases were observed only in the posterior-to-anterior direction. Neuropsychological dysfunction affecting executive, cognitive, or motor functions may be a reflection of a predominance of injuries to the frontal and anterior cortices in patients with mTBI.[40-42] However, posterior regions have also been reported as vulnerable to injuries.[43] Despite these variations in regions reported as susceptible to injury, functional imaging findings are similar. An fMRI meta-analysis found that anterior brain regions had higher activity than the posterior regions in mTBI, suggesting that the anterior regions are more prone to injury.[40] Zhou et al[43] reported FC decreases within the posterior components and increases in the anterior components of the DMN after mTBI. These findings could account for the unidirectional EC increases from posterior-to-anterior DMN regions, reflecting a compensatory mechanism. As the vulnerable anterior regions are injured postconcussion, the use of other brain regions including the posterior regions to assist in function may result in posterior-to-anterior information flow within the DMN.

Abnormal increases in low-frequency neuronal activity have been long associated with mTBI and can be measured and localized with the help of MEG.[4,11,12,42] A handful of animal and human studies have pointed to cortical deafferentation following axonal injury as the main mechanism for delta wave generation after mTBI.[4,42] Delta waves can be induced by administering atropine in animals. Atropine causes deafferentation by blocking antagonist acetylcholine within the cholinergic pathway. These animal studies have demonstrated that mTBI causes dysregulation of the cholinergic system and chronic reduction of acetylcholine.[44,45] This result suggests that dysfunction at the synaptic level can cause delta wave generation without cell death. Another explanation for the increase in EC only in the delta band is based on the bistability concept. Bistability is often demonstrated by slow oscillations, which specifically arise in the setting of functional or anatomic disconnection.[46] These stereotypical slow waves in the delta frequency range result from perturbations to the brain such as concussions or traumatic brain injuries, which may force the intact cortical circuits into a bistable state. Additional studies are needed to delineate the mechanism behind increases in delta activity after mTBI in humans.

Recovery following concussion is often a long and challenging process. Enhancing our understanding of recovery is important to help explain why debilitating symptoms persist in some individuals while resolving in others as well as to aid in personalized patient management. In this study, MEG demonstrated functional abnormalities in the concussed population even at postseason, suggesting that alterations in brain function persisted after athletes were cleared for return to play. Despite significant decreases in connectivity postseason from postconcussion, the connectivity remained significantly higher than the preseason baseline, implying that a longer recovery period may be needed. Whether a portion of these cases go on to ultimately recover to

baseline remains to be explored and highlights the need for additional large, long-term studies.

Limitations

Our study included only male adolescent football players and had a small sample size, limiting the generalizability of our results. Inclusion of orthopedically injured players can address confounding factors like pain or fatigue associated with the alterations in connectivity. The partial normalization implies that the connectivity changes may extend beyond the 3-month postseason timeframe, necessitating large-scale longitudinal studies to further clarify the course of recovery. Furthermore, analyzing the connectivity relationships, including both intra- and internetwork EC (such as frontoparietal or even task-based networks aside from the DMN), could provide additional information on compensatory pathways and network anatomy. Finally, it would also be worth exploring the relationship between measures of cognitive assessment and structural connectivity, such as with diffusion MR imaging, and FC.

CONCLUSIONS

GC analysis of MEG data demonstrates that a single episode of concussion can cause observable changes in the EC within the DMN. The EC partially returns toward the baseline level with time. The connectivity changes occur in the delta band, exhibit directionality, and are only significant from posterior to anterior DMN regions. Further large-scale, longitudinal research will be required to delineate the utility of this approach as a biomarker in concussion.

ACKNOWLEDGMENTS

REFERENCES

1. Virji-Babul N, Hilderman CG, Makan N, et al. **Changes in functional brain networks following sports-related concussion in adolescents.** J Neurotrauma 2014;31:1914–19 CrossRef Medline
2. McCrea M, Guskiewicz KM, Marshall SW, et al. **Acute effects and recovery time following concussion in collegiate football players: the NCAA Concussion Study.** JAMA 2003;290:2556–63 CrossRef Medline
3. Tormenti M, Krieger D, Puccio AM, et al. **Magnetoencephalographic virtual recording: a novel diagnostic tool for concussion.** Neurosurgical Focus 2012;33:E9 1–7 CrossRef Medline
4. Huang MX, Theilmann RJ, Robb A, et al. **Integrated imaging approach with MEG and DTI to detect mild traumatic brain injury**

in military and civilian patients. *J Neurotrauma* 2009;26:1213–26 CrossRef Medline

5. Zhu DC, Covassin T, Nogle S, et al. **A potential biomarker in sports-related concussion: brain functional connectivity alteration of the default-mode network measured with longitudinal resting-state fMRI over thirty days.** *J Neurotrauma* 2015;32:327–41CrossRef Medline

6. Fibel K, Cardinale N, Nichols S, et al. **Evaluating concussion in sport: a case for magnetoencephalography.** *J Athletic Enhancement* 2014;3:4 CrossRef

7. Brookes MJ, Woolrich M, Luckhoo H, et al. **Investigating the electrophysiological basis of resting state networks using magnetoencephalography.** *Proc Natl Acad Sci U S A* 2011;108:16783–88 CrossRef Medline

8. Dale AM, Liu AK, Fischl BR, et al. **Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity.** *Neuron* 2000;26:55–67 CrossRef Medline

9. Schwartz ES, Edgar JC, Gaetz WC, et al. **Magnetoencephalography.** *Pediatr Radiol* 2010;40:50–58 CrossRef Medline

10. Maldjian JA, Davenport EM, Whitlow CT. **Graph theoretical analysis of resting-state MEG data: Identifying interhemispheric connectivity and the default mode.** *Neuroimage* 2014;96:88–94 CrossRef Medline

11. Lee RR, Huang M. *Magnetoencephalography in the Diagnosis of Concussion.* In: *Concussion.* Vol 28. Karger Publishers; 2014:94–111

12. Dunkley B, Da Costa L, Bethune A, et al. **Low-frequency connectivity is associated with mild traumatic brain injury.** *Neuroimage Clin* 2015;7:611–21 CrossRef Medline

13. Dimitriadis SI, Zouridakis G, Rezaie R, et al. **Functional connectivity changes detected with magnetoencephalography after mild traumatic brain injury.** *Neuroimage Clin* 2015;9:519–31 CrossRef Medline

14. Dunkley BT, Urban K, Da Costa L, et al. **Default mode network oscillatory coupling is increased following concussion.** *Front Neurol* 2018;9:280 CrossRef Medline

15. Tarapore PE, Findlay AM, LaHue SC, et al. **Resting state magnetoencephalography functional connectivity in traumatic brain injury.** *J Neurosurg* 2013;118:1306–16 CrossRef Medline

16. Friston KJ. **Functional and effective connectivity: a review.** *Brain Connect* 2011;1:13–36 CrossRef Medline

17. Tana MG, Sclocco R, Bianchi AM. **GMAC: a Matlab toolbox for spectral Granger causality analysis of fMRI data.** *Comput Biol Med* 2012;42:943–56 CrossRef Medline

18. Friston K, Moran R, Seth AK. **Analysing connectivity with Granger causality and dynamic causal modelling.** *Curr Opin Neurobiol* 2013;23:172–78 CrossRef Medline

19. Sharp DJ, Beckmann CF, Greenwood R, et al. **Default mode network functional and structural connectivity after traumatic brain injury.** *Brain* 2011;134:2233–47 CrossRef Medline

20. Raichle ME, MacLeod AM, Snyder AZ, et al. **A default mode of brain function.** *Proc Natl Acad Sci U S A* 2001;98:676–82 CrossRef Medline

21. Bharath RD, Munivenkatappa A, Gohel S, et al. **Recovery of resting brain connectivity ensuing mild traumatic brain injury.** *Front Hum Neurosci* 2015;9:513 CrossRef

22. Urban JE, Davenport EM, Golman AJ, et al. **Head impact exposure in youth football: high school ages 14 to 18 years and cumulative impact analysis.** *Ann Biomed Eng* 2013;41:2474–87 CrossRef Medline

23. Ashburner J, Friston KJ. **Voxel-based morphometry: the methods.** *Neuroimage* 2000;11:805–21 CrossRef Medline

24. Tadel F, Baillet S, Mosher JC, et al. **Brainstorm: a user-friendly application for MEG/EEG analysis.** *Comput Intell Neurosci* 2011;2011:879716 CrossRef Medline

25. Hämäläinen MS, Ilmoniemi RJ. **Interpreting magnetic fields of the brain: minimum norm estimates.** *Med Biol Eng Comput* 1994;32:35–42 CrossRef Medline

26. Desikan RS, Ségonne F, Fischl B, et al. **An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest.** *Neuroimage* 2006;31:968–80 CrossRef Medline

27. Kabbara A, Falou WE, Khalil M, et al. **The dynamic functional core network of the human brain at rest.** *Sci Rep* 2017;7:2936 CrossRef Medline

28. Hillebrand A, Tewarie P, Van Dellen E, et al. **Direction of information flow in large-scale resting-state networks is frequency-dependent.** *Proc Natl Acad Sci U S A* 2016;113:3867–72 CrossRef Medline

29. Granger CW. **Investigating causal relations by econometric models and cross-spectral methods.** *Econometrica* 1969;37:424–38 CrossRef

30. Geweke J. **Measurement of linear dependence and feedback between multiple time series.** *J Am Stat Assoc* 1982;77:304–13 CrossRef

31. Li L, Pagnotta MF, Arakaki X, et al. **Brain activation profiles in mTBI: Evidence from combined resting-state EEG and MEG activity.** *Annu Int Conf IEEE Eng Med Biol Soc* 2015;2015:6963–66 CrossRef Medline

32. Mayer AR, Mannell MV, Ling J, et al. **Functional connectivity in mild traumatic brain injury.** *Hum Brain Mapp* 2011;32:1825–35 CrossRef Medline

33. Hillary FG, Rajtmajer SM, Roman CA, et al. **The rich get richer: brain injury elicits hyperconnectivity in core subnetworks.** *PLoS One* 2014;9:e104021 CrossRef Medline

34. Tang CY, Eaves E, Dams-O'Connor K, et al. **Diffuse disconnectivity in traumatic brain injury: a resting state fMRI and DTI study.** *Transl Neurosci* 2012;3:9–14 CrossRef Medline

35. Palacios EM, Sala-Llonch R, Junque C, et al. **Resting-state functional magnetic resonance imaging activity and connectivity and cognitive outcome in traumatic brain injury.** *JAMA Neurol* 2013;70:845–51 CrossRef Medline

36. Hristopulos D, Babul A, Brucar L, et al. **Disrupted information flow in resting-state in adolescents with sports related concussion.** *Front Hum Neurosci* 2019;13:419 CrossRef Medline

37. de Almeida JR, Versace A, Mechelli A, et al. **Abnormal amygdala-prefrontal effective connectivity to happy faces differentiates bipolar from major depression.** *Biol Psychiatry* 2009;66:451–59 CrossRef Medline

38. Liu Z, Zhang Y, Bai L, et al. **Investigation of the effective connectivity of resting state networks in Alzheimer's disease: a functional MRI study combining independent components analysis and multivariate Granger causality analysis.** *NMR Biomed* 2012;25:1311–20 CrossRef Medline

39. Wang L, Zhang J, Zhang Y, et al. **Conditional granger causality analysis of effective connectivity during motor imagery and motor execution in stroke patients.** *BioMed Res Int* 2016;2016:3870863 CrossRef Medline

40. Eierud C, Craddock RC, Fletcher S, et al. **Neuroimaging after mild traumatic brain injury: review and meta-analysis.** *Neuroimage Clin* 2014;4:283–94 CrossRef Medline

41. Hashimoto K, Abo M. **Abnormal regional benzodiazepine receptor uptake in the prefrontal cortex in patients with mild traumatic brain injury.** *J Rehabil Med* 2009;41:661–65 CrossRef Medline

42. Huang MX, Nichols S, Baker DG, et al. **Single-subject-based whole-brain MEG slow-wave imaging approach for detecting abnormality in patients with mild traumatic brain injury.** *Neuroimage Clin* 2014;5:109–19 CrossRef Medline

43. Zhou Y, Milham MP, Lui YW, et al. **Default-mode network disruption in mild traumatic brain injury.** *Radiology* 2012;265:882–92 CrossRef Medline

44. Gorman LK, Fu K, Hovda DA, et al. **Effects of traumatic brain injury on the cholinergic system in the rat.** *J Neurotrauma* 1996;13:457–63 CrossRef Medline

45. Schaul N, Gloor P, Ball G, et al. **The electromicrophysiology of delta waves induced by systemic atropine.** *Brain Res* 1978;143:475–86 CrossRef

46. D'Andola M, Rebollo B, Casali AG, et al. **Bistability, causality, and complexity in cortical networks: an in vitro perturbational study.** *Cereb Cortex* 2018;28:2233–42 CrossRef Medline

Comparison of Hippocampal Subfield Segmentation Agreement between 2 Automated Protocols across the Adult Life Span

A. Samara, C.A. Raji, Z. Li, and T. Hershey

ABSTRACT

BACKGROUND AND PURPOSE: The hippocampus is a frequent focus of quantitative neuroimaging research, and structural hippocampal alterations are related to multiple neurocognitive disorders. An increasing number of neuroimaging studies are focusing on hippocampal subfield regional involvement in these disorders using various automated segmentation approaches. Direct comparisons among these approaches are limited. The purpose of this study was to compare the agreement between two automated hippocampal segmentation algorithms in an adult population.

MATERIALS AND METHODS: We compared the results of 2 automated segmentation algorithms for hippocampal subfields (FreeSurfer v6.0 and volBrain) within a single imaging data set from adults ($n = 176$, 89 women) across a wide age range (20–79 years). Brain MR imaging was acquired on a single 3T scanner as part of the IXI Brain Development Dataset and included T1- and T2-weighted MR images. We also examined subfield volumetric differences related to age and sex and the impact of different intracranial volume and total hippocampal volume normalization methods.

RESULTS: Estimated intracranial volume and total hippocampal volume of both protocols were strongly correlated ($r = 0.93$ and 0.9, respectively; both $P < .001$). Hippocampal subfield volumes were correlated (ranging from $r = 0.42$ for the subiculum to $r = 0.78$ for the cornu ammonis [CA]1, all $P < .001$). However, absolute volumes were significantly different between protocols. volBrain produced larger CA1 and CA4-dentate gyrus and smaller CA2-CA3 and subiculum volumes compared with FreeSurfer v6.0. Regional age- and sex-related differences in subfield volumes were qualitatively and quantitatively different depending on segmentation protocol and intracranial volume/total hippocampal volume normalization method.

CONCLUSIONS: The hippocampal subfield volume relationship to demographic factors and disease states should undergo nuanced interpretation, especially when considering different segmentation protocols.

ABBREVIATIONS: CA = cornu ammonis; DG = dentate gyrus; HPSF = hippocampal subfield; ICV = intracranial volume; SR-SL-SM = strata radiatum-lacunosum-moleculare; THV = total hippocampal volume

The hippocampus is a major component of the limbic system, and it is affected in several neurocognitive and neuropsychiatric

disorders from Alzheimer disease to major depressive disorder.[1,2] Located in the mesial temporal lobes, the hippocampus functions as a computational hub through its extensive afferent and efferent connections with cortical and subcortical structures.[3] The hippocampus and hippocampal-related structures sustain a range of cognitive functions in the context of episodic and semantic memory, spatial navigation, planning, and learning.[4] The hippocampus is additionally divided into distinct cytoarchitectonic regions called subfields, most prominently the dentate gyrus (DG), cornu ammonis (CA) subfields 1–4, and the subiculum.[5] Distinctive cognitive functions are supported by different subfields,[6] and subfields are differentially affected in various neuropsychiatric disorders.[2,7]

An increasing number of in vivo neuroimaging studies have focused on hippocampal subfield (HPSF) involvement in neurologic and psychiatric conditions.[8,9] The ability to differentiate subfields in vivo provides a unique opportunity to identify

biomarkers for brain diseases like Alzheimer disease.[9] For example, studies have shown that the HPSFs can be impacted by aging and Alzheimer disease in a regional-specific pattern and can be used as an in vivo biomarker with diagnostic and prognostic significance.[10] Manual segmentation has limited clinical throughput due to the time requirement and the necessity of trained operators. Thus, automated approaches are needed to industrialize the clinical throughput across millions of potential brain MR imaging scans.

Two segmentation protocols that are commonly used are FreeSurfer (http://surfer.nmr.mgh.harvard.edu) and volBrain (https://volbrain.upv.es/index.php).[11,12] Between 2013 and 2019, >160 studies applying the FreeSurfer HPSF segmentation protocol in normal development and various neuropsychiatric conditions were published.[13] Although FreeSurfer is the most widely used software, some concerns about segmentation accuracy in earlier versions of FreeSurfer (v5.1, v5.2, and v5.3) have been previously raised,[14,15] leading to several improvements in the more recent versions of FreeSurfer using ex vivo and ultra-high-resolution MR imaging.[11] An increasing number of studies are using the volBrain protocol as an alternative.[16,17] The main advantage that volBrain provides over FreeSurfer is its considerably shorter processing time because the segmentation results are produced in approximately 15 minutes compared with several hours for FreeSurfer. The agreement of HPSF volumes from both protocols has never been directly compared in a single study. Such comparison is critical to allow optimal interpretation of results reported by different research groups.

The goal of the current work was to compare the output of the 2 HPSF segmentation protocols, FreeSurfer v6.0 and volBrain, in a large cohort of adults based on T1- and T2-weighted MR imaging. We selected these 2 protocols because FreeSurfer is the most popular software for hippocampal subfield segmentation and volBrain is an increasingly popular alternative due to its considerably shorter processing time. We evaluated the agreement between the 2 protocols in volumetric assessment and investigated the presence of estimation bias in measurements. We also examined qualitative and quantitative subfield differences related to age and sex and the impact of various intracranial volume (ICV) and total hippocampal volume (THV) normalization methods.

MATERIALS AND METHODS
Participants and MR Imaging Acquisition
We used the publicly available IXI Brain Development Dataset (http://brain-development.org/ixi-dataset/). This data base includes T1- and T2-weighted brain MR imaging scans of healthy adults with a wide age range. For the current analyses, we included scans that were acquired using the 3T scanner (Philips Healthcare) at Hammersmith Hospital to assess the within-subject agreement of HPSF volumes between protocols. T1-weighted imaging parameters were the following: TR = 9.6 ms, TE = 4.6 ms, number of phase encoding steps = 208, echo-train length = 208, reconstruction diameter = 240.0, acquisition matrix = 208 × 208, flip angle = 8.0°, voxel resolution = 0.9 × 0.9 × 1.2 mm. The T2-weighted parameters were the following: TR = 5725.79 ms, TE = 100.0 ms, number of phase encoding steps = 187, echo-train length = 16, reconstruction diameter = 240.0 mm,

acquisition matrix = 192 × 187, flip angle = 90.0°, voxel resolution = 0.9 × 0.9 × 1.2 mm.

FreeSurfer and volBrain Segmentation
FreeSurfer v6.0 software is one of the most widely used pipelines to obtain HPSF volumes. The FreeSurfer HPSF segmentation module generates a fully automated segmentation based on a probabilistic atlas.[11] For each scan, we used the output volume from the standard FreeSurfer processing of T1 MR imaging after correcting for motion, intensity normalization, and skull stripping. The FreeSurfer algorithm detects local variations in MR imaging contrast using a Bayesian inference algorithm and relies on a hippocampus atlas generated by combining manual labels from ex vivo and in vivo whole-brain scans.[11,18] FreeSurfer uses both T1 and T2 MR imaging to obtain a more reliable segmentation.[19] We used both T1 and T2 MR imaging in the hippocampus subfield segmentation stage to improve tissue contrast and assist in identifying landmarks of the internal hippocampal structure. FreeSurfer generates 12 subfields: parasubiculum, presubiculum, subiculum, CA1, CA3, CA4, granule cell and molecular layer of the dentate gyrus, molecular layer, hippocampus-amygdala transition area, fimbria, hippocampal tail, and hippocampal fissure (definitions of subfield boundaries are detailed in Iglesias et al[11]). CA2 is always included in the CA3 label, as CA2-CA3. We combined CA4 and the granule cell and molecular layer of the dentate gyri CA4-DG in subsequent analyses.

The volBrain hippocampal subfield segmentation protocol is a new method that consists of a fast multiatlas nonlocal patch-based label fusion.[12] This pipeline is publicly available on a web-based platform.[20] volBrain provides the option to use multimodal images (T1 and T2 MR imaging) for improved accuracy of segmentation, which we used in our analysis. The original algorithm is based on the Winterburn atlas,[21] which produces 5 subfield labels: CA1, CA2/3, CA4/DG, stratum radiatum/stratum lacunosum/stratum moleculare (SR-SL-SM), and subiculum. The processing time is about 15 minutes per scan. An example of both HPSF segmentations is shown in Fig 1. Due to the large number of scans included, we did not review each scan by visual inspection after completion, and we did not perform any manual corrections. However, as a quality control measure, we excluded individuals with >1 subfield volume as an outlying value (outlier defined as > 3 SDs).

ICV Normalization Methods
We examined the effect of different ICV normalization methods on the HPSF volumetric analysis. Total ICV estimation from each protocol was used to correct for the subfields derived by the same protocol. We performed 3 distinct approaches to account for variations in total ICV. These methods were the following: 1) the proportion method (calculated by multiplying each individual subfield-to-ICV ratio with the average ICV for the entire cohort); 2) the residual method (regressing out the effect of ICV before conducting further analysis); and 3) the covariate method (in which we included estimated ICV as a covariate in the regression analyses as described below). In addition, to evaluate regional differences in HPSF, we performed similar correction methods to account for variation in THV.[22]
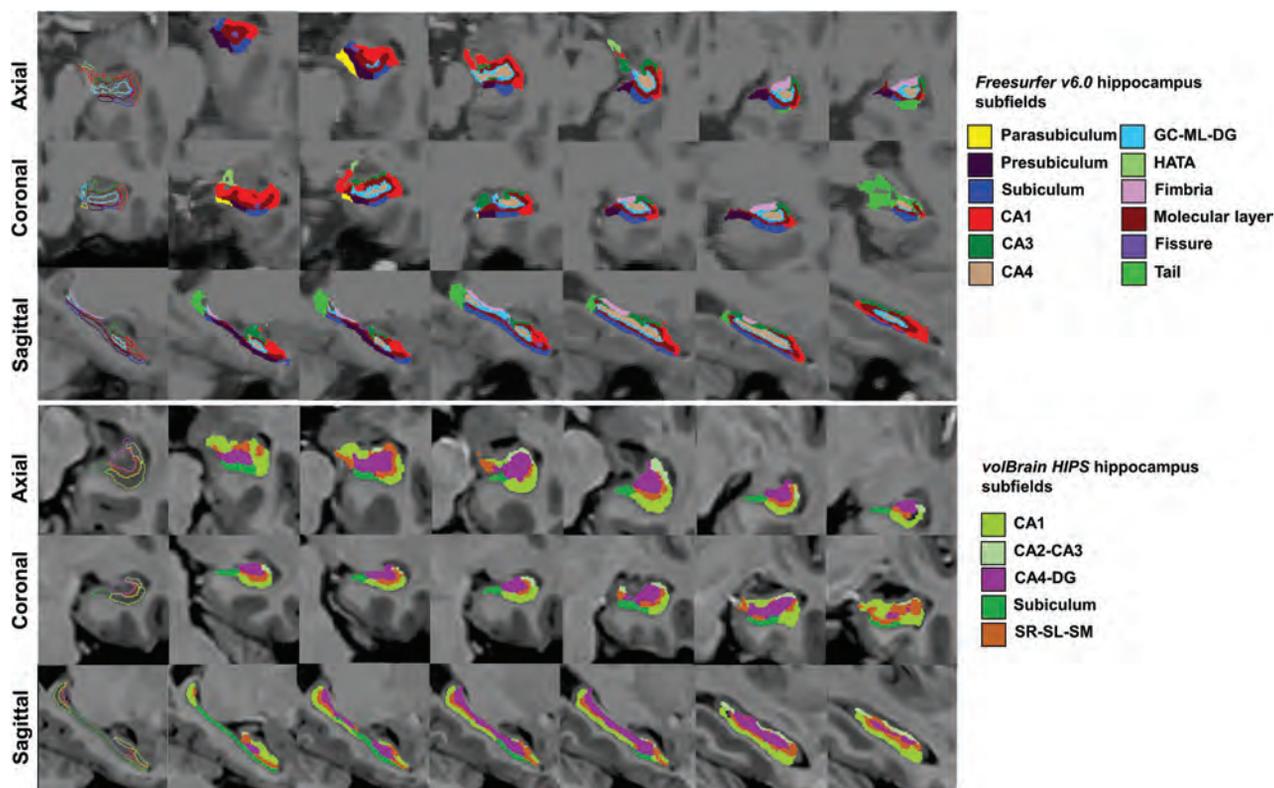
FIG 1. An example of hippocampal subfield segmentation by FreeSurfer (*upper row*) and volBrain (*lower row*) shown in axial, coronal, and sagittal sections. GC-ML-DG indicates granule cell and molecular layers of the dentate gyrus; HATA, hippocampus-amygdala transition area.

Statistical Analysis

We conducted all statistical analyses and illustrations using R statistical and computing software (Version 3.6.3; http://www.r-project.org/). We combined the right and left hemispheric volumes for each subfield. The THV for each protocol was calculated by combining all subfields (excluding the hippocampal fissure in FreeSurfer segmentation because it represents CSF). We used Pearson r correlations to relate HPSFs between the 2 protocols and paired t-tests to compare the mean differences between the 2 groups. To compare the effects of different ICV normalization methods on the relationship between HPSF and age and sex variables, we conducted multiple linear regression analyses and reported the dependent variable estimates for each ICV/THV normalization method. Additionally, we calculated the effect size of the difference in HPSFs between men and women using the Cohen's D. In regression analyses, the Bonferroni correction for multiple comparisons was used as $P < .0045$ ($P = .05/$ 11 subfields) for the FreeSurfer analyses, and $P < .01$ ($P = .05/5$ subfields) for the volBrain analyses. Finally, Bland-Altman plots were produced to visualize the disagreement between FreeSurfer and volBrain in terms of absolute, uncorrected hippocampal subfield volumes.

RESULTS

Participants

A total of 176 eligible brain MR imaging scans underwent HPSF segmentation. We subsequently excluded 4 scans because they yielded outlying volumetric values in >1 subfield (2 FreeSurfer and 2 volBrain). After exclusion, our analyzed sample included 83 men with an age range of 20–79 years (mean = 45 [SD, 16] years) and 89 women with an age range of 21–82 years (mean = 50 [SD, 17] years). A few participants had a single outlier across all subfields (FreeSurfer: CA2-CA3, 1; parasubiculum, 1; hippocampus-amygdala transition area, 1; fimbria, 1; fissure, 2; volBrain: CA1, 1; CA2-CA3, 2; SR-SL-SM, 1; subiculum, 1). We excluded these outliers, but not the entire scans, from subsequent analyses.

Hippocampal Subfield Volumes

The HPSF volumes produced by both FreeSurfer and volBrain are detailed in the Table. Further correlation analyses were applied only to subfields shared by both FreeSurfer and volBrain segmentations (ie, CA1, CA2-CA3, CA4-DG, and the subiculum). We observed significant correlations between the CA1, CA2-CA3, CA4-DG, and subiculum volumes segmented by both FreeSurfer and volBrain (Fig 2; $P < .001$ for all correlations). Correlation was strongest for CA1 ($r = 0.78$) and weakest for the subiculum volume ($r = 0.42$). However, the 2 protocols produced different average volumes. volBrain yielded larger average CA1 and CA4-DG than FreeSurfer, while FreeSurfer conversely yielded larger CA2-CA3 and subiculum volumes than volBrain ($P < .001$ for all subfields; Fig 2 and Table). The Bland-Altman plots confirmed that for almost every scan, FreeSurfer generated smaller volumes for CA1 and CA4-DG and larger volumes for CA2-CA3 and the subiculum compared with volBrain (Fig 3). Furthermore, Bland-Altman plots demonstrated that the size of

Uncorrected hippocampal subfield, total hippocampal, and intracranial volumes measured by FreeSurfer and volBrain protocols[a]

	Total (n = 172)		Male (n = 83)		Female (n = 89)		
	Mean	SD	Mean	SD	Mean	SD	Cohen's D
FreeSurfer v6.0							
CA1	1251	193	1321	169	1184	159	0.83
CA2-CA3	472	71	497	51	449	54	0.91
CA4-DG	1113	171	1173	141	1057	135	0.84
Subiculum	807	102	841	101	775	87	0.7
Parasubiculum	121	21	126	16	117	21	0.48
Presubiculum	553	72	581	75	527	59	0.8
Molecular layer	1205	168	1241	136	1172	150	0.48
HATA	136	31	142	19	130	23	0.57
Fimbria	178	67	188	62	168	48	0.36
Tail	999	160	1049	155	952	118	0.7
Fissure[b]	269	71	276	49	262	70	0.23
THV	6839	957	7163	846	6537	758	0.78
eICV	1,563,261	146,983	1,677,235	158,967	1,456,885	231,310	1.11
volBrain, HIPS segmentation protocol							
CA1	1486	286	1574	267	1405	234	0.67
CA2-CA3	264	37	276	43	253	32	0.61
CA4-DG	1420	221	1462	207	1382	260	0.34
SR-SL-SM	938	158	988	173	892	145	0.6
Subiculum	566	87	593	86	541	93	0.58
THV	4689	771	4908	776	4484	665	0.59
eICV	1,374,768	133,467	1,462,729	139,905	1,292,671	126,129	1.28

Note:—HATA indicates hippocampus-amygdala transition area; eICV, estimated intracranial volume; HIPS, hippocampal subfield segmentation protocol.
[a] Uuits are cubic millimeters.
[b] Not significantly different. Men had significantly larger values than females in all volumes, except hippocampus fissure ($P < .05$, Student t test). The average values of total and subfield volumes reported here are the sum of right and left hemispheres combined.
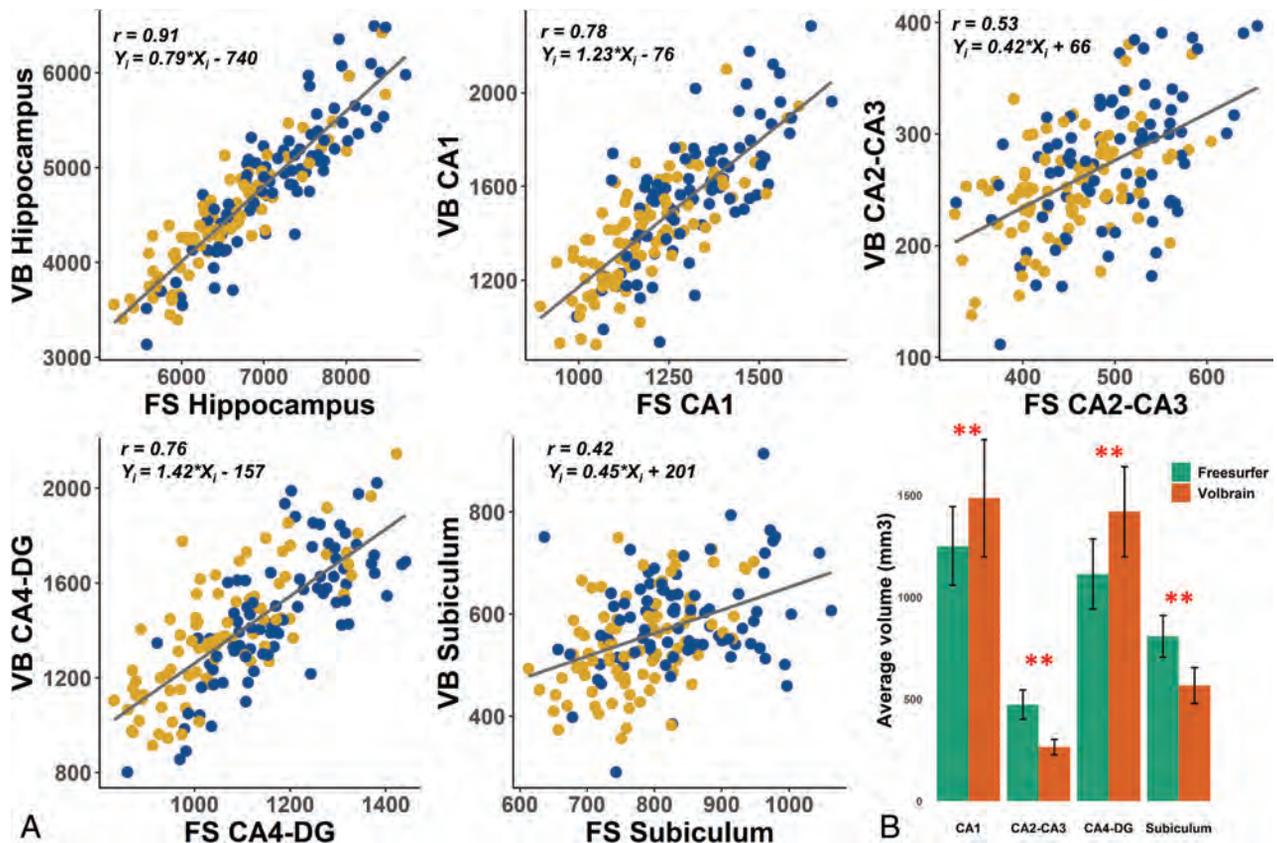
FIG 2. *A*, Comparison of uncorrected total hippocampus, CA1, CA2-CA3, CA4-DG, and subiculum volumes (cubic millimeters) between FreeSurfer and volBrain (yellow, women; blue, men). *Regression lines* relating volBrain to FreeSurfer volumes are shown for each subfield. The average values of subfield volumes reported here are the sum of right and left hemisphere volumes combined. All Pearson r correlations are significant ($P < .001$). *B*, Bar graphs show means (SDs). *Double asterisks* indicate statistical significance ($P < .001$); FS, FreeSurfer; VB, volBrain.
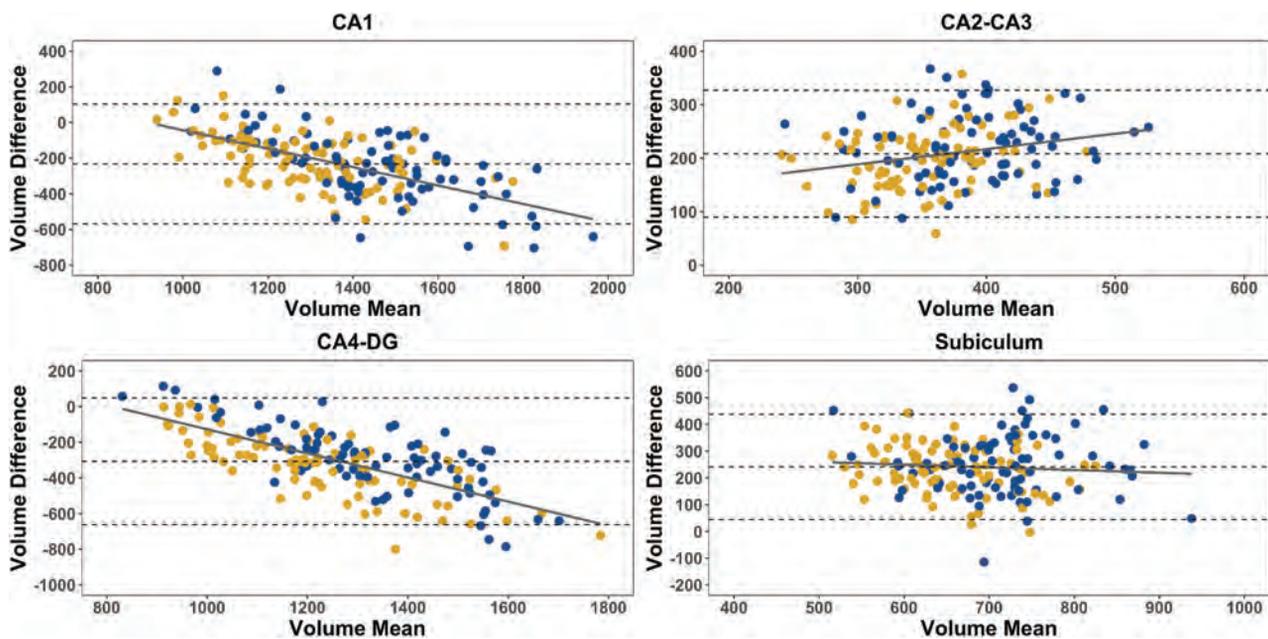
FIG 3. Bland-Altman plots for uncorrected subfield volumes (CA1, CA2-CA3, CA4-DG, and subiculum volumes [cubic millimeters]) generated by FreeSurfer and volBrain (yellow, women; blue, men). Intrasubject volume difference (y-axis) is defined as (FreeSurfer volume) – (volBrain volume). Mean volume is represented on the x-axis. The mean (SD, 1.96) volume difference and 95% confidence intervals are plotted as *dashed horizontal lines*. Except for the subiculum, all Pearson *r* correlations are significant ($P < .001$).

the disagreement between the 2 protocols increased for larger volume estimates of CA1 ($r = -0.61$; 95% CI, -0.70 to -0.51), CA2-CA3 ($r = 0.25$; 95% CI, $0.11-0.39$), and CA4-DG ($r = -0.71$; 95% CI, -0.78 to -0.63). No such relationship was found for the subiculum ($r = -0.08$; 95% CI, $-0.23-0.07$; $P = .30$).

THV and total ICV estimations were strongly correlated between FreeSurfer and volBrain ($r = 0.91$ and $r = 0.93$, respectively; $P < .001$). Compared to volBrain, FreeSurfer produced higher THV and ICV volumes ($P < .001$ for both) (Table). Both FreeSurfer and volBrain produced larger uncorrected HPSF volumes in men compared with women, except for the hippocampal fissure (Table).

ICV and THV Normalization Methods
For each HPSF segmented by either FreeSurfer or volBrain, volume values were normalized for both ICV and THV derived by the same protocol using the covariate, proportion, and residual methods. These normalized volumes were then entered, separated by normalization method, into multiple linear regression models with participant age and sex as covariates (Online Supplemental Data).

Both age and sex showed different associations with HPSF volumes depending on the segmentation protocol and normalization method. Marked inconsistency in the statistical significance and magnitude of the regression estimates could be observed in multiple HPSFs. Specifically, in FreeSurfer, CA1, CA2-CA3, CA4-DG, and the presubiculum showed significant negative correlations with age only per ICV covariate and residual methods. However, for the molecular layer and subiculum, the significant positive correlation with age could only be established with THV normalization, but not ICV normalization. Regardless of the ICV/THV normalization method, age positively correlated with

hippocampal fissure and negatively correlated with the fimbria and hippocampal tail. HPSF volumes were consistently higher in women than in men only when the ICV proportion method was used. In volBrain, CA4-DG volume negatively correlated with age across all ICV/THV normalization methods. CA1 and the subiculum negatively correlated with age only when ICV normalization methods were applied. The SR-SL-SM subfield positively correlated with age when THV methods were used. CA4-DG volumes were significantly higher in women using all normalization methods, except for the ICV covariate method.

Moreover, some contradictory findings emerged when comparing the results between the 2 segmentation protocols. Most strikingly, in CA2-CA3, significant regression estimates for age had negative findings in FreeSurfer but positive findings in volBrain. In the subiculum, the estimates were positive in FreeSurfer but negative in volBrain. Additionally, by means of the ICV proportion method, CA2-CA3 and subiculum volumes were significantly larger in women than in men for FreeSurfer but not for volBrain, which showed no significant sex effect.

DISCUSSION
Our study is the first to directly compare the results of 2 commonly used HPSF segmentation protocols, providing new insight to help compare results across different neuroimaging studies. Although the HPSF volumes produced using the 2 protocols were well-correlated, significant differences were observed in absolute volumes. Specifically, volBrain produced larger CA1 and CA4-DG volumes, while FreeSurfer produced larger CA2-CA3 and subiculum volumes. We also observed systematic biases in the HPSF estimations because the absolute volume difference between the 2 protocols increased for larger volume estimates for

CA1, CA2-CA3, and CA4-DG. Finally, we found that different segmentation protocols and ICV/THV normalization methods could lead to inconsistent and sometimes contradictory conclusions regarding the regional effects of age and sex on HPSF volumes.

While absolute volumetric differences exist across the 2 protocols, their results being correlated indicates that they may be interchangeably used for correlational analyses. Some of the inconsistencies between protocols are likely due to differences in the number of HPSF labels (FreeSurfer, $n = 12$; volBrain, $n = 5$) and how the 2 protocols define the HPSF anatomic boundaries. For example, FreeSurfer produces specific labels for the hippocampal tail, fimbria, hippocampus-amygdala transition area, parasubiculum, and presubiculum, while these subfield labels do not exist in volBrain. Yushkevich et al[23] compared the results of 21 HPSF labeling protocols (which did not include volBrain) and concluded that the greatest disagreement was along the CA1/subiculum anatomic boundary and anterior hippocampus. This observation could potentially explain the larger CA1 and smaller subiculum produced by volBrain compared with FreeSurfer.[23] Finally, the correlation between the 2 protocols was more robust for THV than for any HPSF, suggesting greater agreement in the outer hippocampal boundaries than in HPSF labels.

Differences in the age range and acquisition parameters in each algorithm training data set might have also contributed to the observed differences. The generative model for the FreeSurfer protocol was based on 15 ex vivo postmortem brain hemispheres obtained from individuals 60–91 years of age, with some individuals who had mild Alzheimer disease or mild cognitive impairment.[11] The brain tissue was scanned using 7T MR imaging at 0.13-mm isotropic resolution on average. On the other hand, the volBrain segmentation protocol relied on the Winterburn atlas data base obtained from 5 healthy individuals 29–57 years of age and acquired at 0.6-mm isotropic resolution.[12] On the basis of the age range differences included in each dataset, it is reasonable to suggest that volBrain might provide more accurate segmentations when applied in younger age groups. In fact, when FreeSurfer segmentation is applied to standard resolution MR imaging (1 mm isotropic), the molecular layer would not be clearly visible and atlas deformation is unlikely to be influenced by this anatomic feature. In this case, fitting of the atlas to internal structure relies on prior encoded information alone.[11,24]

This issue will introduce bias in a way that underestimates CA1 and CA4-DG volumes in younger individuals because these 2 subfields are susceptible to age-related changes.[15] The Bland-Altman plots support this explanation and show that between-protocol differences in CA1 and CA4-DG volumes increased with higher mean volumes (ie, in younger individuals), while the differences decreased with lower mean volumes (ie, in older individuals). Iglesias et al[11] explicitly stated in their original article that the FreeSurfer atlas might include hippocampal atrophy because it was built using delineations in elderly individuals, which could compromise its applicability in younger populations. Nevertheless, we acknowledge the possibility that the differences observed in the Bland-Altman plots could also be attributed to differing segmentation boundaries between the 2 protocols or higher error variances in one segmentation method than in the other.

When the volBrain protocol was compared with manual segmentation from the Winterburn data base at a standard resolution (0.9 mm isotropic), the average Dice similarity score was 0.66 (ranging between 0.52 for CA2-CA3 and 0.76 for CA4-DG).[12] These findings highlight the inherent limitations of the volBrain protocol. On the other hand, Iglesias et al[11] performed a qualitative assessment of the multimodal FreeSurfer segmentation on the Winterburn atlas data base. The authors suggested that direct spatial overlap evaluation (eg, using Dice similarity scores) between the Winterburn manual and FreeSurfer automated segmentations is not possible due to labeling protocol differences. Although the agreement between both segmentations was fair in general, prominent differences were observed in areas poorly supported by the MR imaging contrast (eg, the medial digitation) and regions where the definitions of HPSF boundaries were not similar (eg, the inferior parts). For example, the FreeSurfer subiculum was mostly part of the Winterburn atlas CA1 subfield, while the presubiculum and parasubiculum approximately corresponded to the Winterburn atlas subiculum.[11]

ICV normalization is a commonly used procedure in neuroimaging research, and it is an important step to account for sex differences and intersubject variations in head size. This step is necessary because relative, rather than absolute, differences in volumes better described the structure-function relationships. Several ICV normalization methods had been described in the literature, including the covariate, proportion, and residual methods. In addition, van Eijk et al[22] reported sex-dependent regional differences in HPSF volumes after adjusting for THV. They suggested that the THV normalization could provide additional value over ICV normalization alone.[22] When we applied different ICV/THV normalization methods in our study, the most noticeable finding was the marked impact of the choice of normalization method on both the direction and magnitude of estimates of age and sex-related differences. Previous studies have demonstrated marked effects of ICV normalization methods on volumetric assessment of cortical and subcortical structures.[25,26] For example, different ICV normalization methods produce contradictory results regarding the presence of sex-related volumetric differences.[25] Some of these studies have also suggested that the residual method generally provides greater advantages over the proportion and covariate methods,[25-27] and these recommended residual ICV normalization for volumetric studies of neuroanatomic structures.[25] We also noticed a clear trend of larger HPSF volumes in women compared with men when the ICV proportion method was used. This finding is consistent with those in prior studies showing that women have proportionately larger gray matter regions than men,[26,28] and these differences are potentially due to ICV differences rather than sex.[25]

The main limitations of this study include a focus on quantitative values for the HPSF volumes without looking at spatial overlap metrics and comparing label segmentations on a voxel-by-voxel basis. Also, we could not directly compare the reliability of HPSF segmentations across different scanners, voxel resolutions, and field strengths (1.5T versus 3T). How this would affect the comparison across protocols remains to be determined. However, prior work had shown that understanding the performance of HPSF

segmentation software at this field strength carries potentially greater clinical significance.[29] Additionally, future studies should compare the segmentation results in pathologic conditions like Alzheimer disease. The performance of both protocols could drastically change if applied to MR imaging of patients, when tissue damage could reduce the contrast between the tissues and, consequently, lead to greater variability in segmentation.

CONCLUSIONS

Although automatic segmentation of HPSFs has enabled large-scale in vivo analysis and has increased in popularity in recent years, it is important to interpret the results of these studies with caution. Although volumetric assessment of HPSF derived from FreeSurfer and volBrain are well-correlated, we showed significant differences between the 2 protocols in terms of absolute volumes and estimation bias. These differences could lead to different conclusions about HPSF changes across the adult life span. We also showed that the specific ICV normalization method used could influence the conclusions. Researchers should also be very careful when combining data across different protocols in any meta-analyses. Finally, the findings of our study highlight the need for a standard unified approach for HPSFs in neuroimaging studies.

REFERENCES

1. Raji CA, Lopez OL, Kuller LH, et al. **Age, Alzheimer disease, and brain structure.** *Neurology* 2009;73:1899–905 CrossRef Medline
2. Han KM, Kim A, Kang W, et al. **Hippocampal subfield volumes in major depressive disorder and bipolar disorder.** *Eur Psychiatry* 2019;57:70–77 CrossRef Medline
3. Battaglia FP, Benchenane K, Sirota A, et al. **The hippocampus: hub of brain network communication for memory.** *Trends Cogn Sci* 2011;15:310–18 CrossRef Medline
4. Buzsáki G, Moser EI. **Memory, navigation and theta rhythm in the hippocampal-entorhinal system.** *Nat Neurosci* 2013;16:130–38 CrossRef Medline
5. Amaral DG, Witter MP. **The three-dimensional organization of the hippocampal formation: a review of anatomical data.** *Neuroscience* 1989;31:571–91 CrossRef Medline
6. Dimsdale-Zucker HR, Ritchey M, Ekstrom AD, et al. **CA1 and CA3 differentially support spontaneous retrieval of episodic contexts within human hippocampal subfields.** *Nat Commun* 2018;9:294 CrossRef Medline
7. Mueller SG, Weiner MW. **Selective effect of age, Apo e4, and Alzheimer's disease on hippocampal subfields.** *Hippocampus* 2009;19:558–64 CrossRef Medline
8. Apostolova LG, Dutton RA, Dinov ID, et al. **Conversion of mild cognitive impairment to Alzheimer disease predicted by hippocampal atrophy maps.** *Arch Neurol* 2006;63:693–99 CrossRef Medline
9. Hanseeuw BJ, Van Leemput K, Kavec M, et al. **Mild cognitive impairment: differential atrophy in the hippocampal subfields.** *AJNR Am J Neuroradiol* 2011;32:1658–61 CrossRef Medline
10. Hett K, Ta VT, Catheline G, et al. Alzheimer's Disease Neuroimaging Initiative. **Multimodal hippocampal subfield grading for Alzheimer's disease classification.** *Sci Rep* 2019;9:13845 CrossRef Medline
11. Iglesias JE, Augustinack JC, Nguyen K, et al. Alzheimer's Disease Neuroimaging Initiative. **A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: application to adaptive segmentation of in vivo MRI.** *Neuroimage* 2015;115:117–37 CrossRef Medline
12. Romero JE, Coupé P, Manjón JV. **HIPS: a new hippocampus subfield segmentation method.** *Neuroimage* 2017;163:286–95 CrossRef Medline
13. Sämann PG, Inglesias JE, Gutman B, et al. **FreeSurfer-based segmentation of hippocampal subfields: a review of methods and applications, with a novel quality control procedure for ENIGMA studies and other collaborative efforts.** *Hum Brain Mapp* 2020 Dec 27. [Epub ahead of print] CrossRef Medline
14. Wisse LE, Biessels GJ, Geerlings MI. **A critical appraisal of the hippocampal subfield segmentation package in FreeSurfer.** *Front Aging Neurosci* 2014;6:261 CrossRef Medline
15. de Flores R, La Joie R, Landeau B, et al. **Effects of age and Alzheimer's disease on hippocampal subfields: comparison between manual and FreeSurfer volumetry.** *Hum Brain Mapp* 2015;36: 463–74 CrossRef Medline
16. Moghaddam HS, Aarabi MH, Mehvari-Habibabadi J, et al. **Distinct patterns of hippocampal subfield volume loss in left and right mesial temporal lobe epilepsy.** *Neurol Sci* 2021;42:1411–21 CrossRef Medline
17. Kim S, Greene DJ, D'Andrea CB, et al. **Hippocampal volume in provisional tic disorder predicts tic severity at 12-month follow-up.** *J Clin Med* 2020;9:1715 CrossRef Medline
18. Van Leemput K, Bakkour A, Benner T, et al. **Automated segmentation of hippocampal subfields from ultra-high resolution in vivo MRI.** *Hippocampus* 2009;19:549–57 CrossRef Medline
19. Despotović I, Goossens B, Philips W. **MRI segmentation of the human brain: challenges, methods, and applications.** *Comput Math Methods Med* 2015;2015:450341 CrossRef Medline
20. Manjón JV, Coupé P. **volBrain: an online MRI brain volumetry system.** *Front Neuroinform* 2016;10:30 CrossRef Medline
21. Winterburn JL, Pruessner JC, Chavez S, et al. **A novel in vivo atlas of human hippocampal subfields using high-resolution 3 T magnetic resonance imaging.** *Neuroimage* 2013;74:254–65 CrossRef Medline
22. van Eijk L, Hansell NK, Strike LT, et al. **Region-specific sex differences in the hippocampus.** *Neuroimage* 2020;215:116781 CrossRef Medline
23. Yushkevich PA, Amaral RS, Augustinack JC, et al. Hippocampal Subfields Group (HSG). **Quantitative comparison of 21 protocols for labeling hippocampal subfields and parahippocampal subregions in in vivo MRI: towards a harmonized segmentation protocol.** *Neuroimage* 2015;111:526–41 CrossRef Medline
24. Wisse LE, Chételat G, Daugherty AM, et al. **Hippocampal subfield volumetry from structural isotropic 1 mm^3 MRI scans: a note of caution.** *Hum Brain Mapp* 2021;42:539–50 CrossRef Medline
25. Pintzka CW, Hansen TI, Evensmoen HR, et al. **Marked effects of intracranial volume correction methods on sex differences in neuroanatomical structures: a HUNT MRI study.** *Front Neurosci* 2015;9:238 CrossRef Medline
26. Voevodskaya O, Simmons A, Nordenskjöld R, et al, Alzheimer's Disease Neuroimaging Initiative. **The effects of intracranial volume adjustment approaches on multiple regional MRI volumes in healthy aging and Alzheimer's disease.** *Front Aging Neurosci* 2014;6:264 CrossRef Medline
27. Sanfilipo MP, Benedict RH, Zivadinov R, et al. **Correction for intracranial volume in analysis of whole brain atrophy in multiple sclerosis: the proportion vs. residual method.** *Neuroimage* 2004;22:1732–43 CrossRef Medline
28. Gur RC, Turetsky BI, Matsui M, et al. **Sex differences in brain gray and white matter in healthy young adults: correlations with cognitive performance.** *J Neurosci* 1999;19: 4065–72 CrossRef Medline
29. Chow N, Hwang KS, Hurtz S, et al. **Comparing 3T and 1.5T MRI for mapping hippocampal atrophy in the Alzheimer's Disease Neuroimaging Initiative.** *AJNR Am J Neuroradiol* 2015;36:653–60 CrossRef Medline

The Brain Metabolic Signature in Superagers Using In Vivo ^{1}H-MRS: A Pilot Study

L.L. de Godoy, A. Studart-Neto, M. Wylezinska-Arridge, M.H. Tsunemi, N.C. Moraes, M.S. Yassuda, A.M. Coutinho, C.A. Buchpiguel, R. Nitrini, S. Bisdas, and C. da Costa Leite

ABSTRACT

BACKGROUND AND PURPOSE: Youthful memory performance in older adults may reflect an underlying resilience to the conventional pathways of aging. Subjects having this unusual characteristic have been recently termed "superagers." This study aimed to explore the significance of imaging biomarkers acquired by ^{1}H-MRS to characterize superagers and to differentiate them from their normal-aging peers.

MATERIALS AND METHODS: Fifty-five patients older than 80 years of age were screened using a detailed neuropsychological protocol, and 25 participants, comprising 12 superagers and 13 age-matched controls, were statistically analyzed. We used state-of-the-art 3T ^{1}H-MR spectroscopy to quantify 18 neurochemicals in the posterior cingulate cortex of our subjects. All ^{1}H-MR spectroscopy data were analyzed using LCModel. Results were further processed using 2 approaches to investigate the technique accuracy: 1) comparison of the average concentration of metabolites estimated with Cramer-Rao lower bounds <20%; and 2) calculation and comparison of the weighted means of metabolites' concentrations.

RESULTS: The main finding observed was a higher total *N*-acetyl aspartate concentration in superagers than in age-matched controls using both approaches ($P = .02$ and $P = .03$ for the weighted means), reflecting a positive association of total *N*-acetyl aspartate with higher cognitive performance.

CONCLUSIONS: ^{1}H-MR spectroscopy emerges as a promising technique to unravel neurochemical mechanisms related to cognitive aging in vivo and providing a brain metabolic signature in superagers. This may contribute to monitoring future interventional therapies to avoid or postpone the pathologic processes of aging.

ABBREVIATIONS: CRLB = Cramer-Rao lower bounds; Glu = glutamate; GCP = glycerophosphocholine; NAAG = *N*-acetyl aspartylglutamate; PCho = phosphocholine; PCr = phosphocreatine

Multifactorial neurobiologic mechanisms appear to underlie the complex phenomenon of superior cognitive performance in older adults.[1] Subjects exhibiting this outstanding phenotype are newly described as "superagers"[2] and have been studied by imaging, through structural[3] and functional MR imaging.[4] To date, it is known that superagers show selective cortical preservation involving regions of the default mode and salience networks, which also exhibit strong functional connectivity.[5] There is a paucity of data, in particular generated by ^{1}H-MRS, related to the metabolic profile of the brains of superagers.[6]

Previous studies have reported the correlation between ^{1}H-MR spectroscopy parameters of metabolite concentration with measures of intelligence,[7-9] affect,[10] creativity,[11] and personality[12] in cohorts of healthy human individuals, mostly younger adults. Such findings suggest that neurometabolites may be involved in healthy brain aging trajectories, such as those observed among superagers. Therefore, ^{1}H-MR spectroscopy, which noninvasively probes several brain metabolites, can be used to provide information on

biochemical pathways associated with the cognitive status of the aging brain.[13]

Physiologic aging changes involve specific cellular mechanisms, such as bioenergetics, oxidative stress, inflammation, cell membrane turnover, and endogenous neuroprotection revealed by neurochemicals in [1]H-MR spectroscopy.[14] Overall, the aging brain shows a reduced concentration of NAA and glutamate (Glu) and an increased concentration of choline (Cho) and myoinositol (mIns).[6] Few studies showed a positive correlation between the NAA concentration and better memory performance in healthy older adults.[6] Nevertheless, there are no studies to determine whether these age-related metabolite changes correlate with neuropsychological performance in superagers using age-matched controls for comparison.

[1]H-MR spectroscopy under standardized conditions is necessary to obtain an accurate representation of the metabolic profile in vivo. First, the VOI must be placed in the least inhomogeneous area of gray or white matter. A high field strength (at least 3T) is needed to achieve better spectral resolution and a higher signal-to-noise ratio for metabolites present in tissue at a much lower concentration than in water or those affected by J-coupling.[15] Metabolite concentrations as relative ratios of creatine (Cr) should be avoided in older adults[16] because creatine seems to vary in the aging brain.[17] Finally, optimal shimming methods and implementation of post hoc correction based on tissue and CSF segmentation within the VOI are mandatory.[18]

In the present study, we sought to investigate any relationship between brain metabolite concentrations obtained by [1]H-MR spectroscopy with the late-life cognitive performance in a selected population of superagers and age-matched controls. We hypothesized that there is a different neurochemical signature between these 2 subject groups. Our results may be useful to address the resilience process of aging that underlies the superager's profile.

MATERIALS AND METHODS
Selection of Participants
Initially, 55 participants (Online Supplemental Data) were screened from different centers in the city of Sao Paulo, Brazil, including the outpatient clinic of the Geriatrics Department of Hospital das Clinicas da Faculdade de Medicina da Universidade de Sao Paulo; the Open University Program for Senior Citizens at the School of Arts, Sciences and Humanities Universidade de Sao Paulo; and the Development Center for the Promotion of Healthy Aging. We also screened community elderly volunteers through social media and newspaper campaigns. Informed consent was obtained from each participant, and the research project was approved by the Ethics Committee of the University of Sao Paulo (No. 62047616.0.0000.0068). The study was designed and conducted according to the Declaration of Helsinki.

The inclusion criteria for the participants were the following: 1) 80 years of age or older; 2) education ≥4 years; 3) Mini-Mental State Examination findings normal for their education status;[19,20] 4) Functional Activity Questionnaire score ≤ 4;[21] 5) Clinical Dementia Rating score equal to zero; and 6) the 15-question version of the Geriatric Depression Scale result of ≤5.

The exclusion criteria included the following: 1) diagnosis of dementia or mild cognitive impairment according to the criteria of the National Institute on Aging and Alzheimer's Association;[22,23] 2) diagnosis of a major psychiatric disorder by the *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition; 3) a history of alcohol or psychoactive drug abuse; 4) current or previous diagnosis of diseases of the CNS (ie, stroke or seizure); 5) the presence of structural lesions in the CNS at image examination that could distort the brain parenchyma (ie, tumor or brain malformation); and 6) visual and/or auditory limitations that impair the performance of cognitive tests.

Neurocognitive Screening
The first assessment consisted of a semi-structured interview with the collection of sociodemographic data; cognitive assessment using the Mini-Mental State Examination, the Montreal Cognitive Assessment, and the Brief Cognitive Screening Battery;[24] screening for depressive symptoms and anxiety using the Geriatric Depression Scale-15 and the Geriatric Anxiety Inventory, respectively; and a functional assessment with the Functional Activity Questionnaire score and the Clinical Dementia Rating.

Subsequently, the subjects who met the inclusion criteria underwent neuropsychological tests. The tests included Digit Span Forward and Backward, Trail-Making Tests A and B, category (animals) and letter verbal fluency tests, Rey-Osterrieth Complex Figure (copy and delayed recall), Logical Memory of the Wechsler Memory Scale, Rey Auditory Verbal Learning Test, the 60-item version of the Boston Naming Test, and the estimated intelligence quotient, which was measured with the Wechsler Adult Intelligence Scale, Third Edition. Those who performed equal or less than −1.5 SDs from average normative values adjusted by age and education for any cognitive test aforementioned were excluded.

Healthy Older Adults Grouping
Participants were separated into 2 groups: namely superagers and age-matched controls. Superagers were defined as the participants who presented a delayed recall score (30 minutes) of the Rey Auditory Verbal Learning Test, a measure of episodic memory, equal to or greater than average normative values for individuals 50–60 years of age (≥9 words), according to the criteria established by the Northwestern SuperAging Research Program.[25] In addition, to fulfill this superager definition, they had to perform within or above 1 SD of the average range for their age and demographics for cognitive function in the nonmemory domains tests, including the Digit Span Forward and Backward, the Boston Naming Test-60, Trail-Making Tests A and B, the Rey-Osterrieth Complex Figure, and category (animals) and letter verbal fluency, according to the published normative values.[26,27] The age-matched controls performed within 1 SD of the average range for their age and demographics in memory and nonmemory domains, which means that they were average older adults according to their cognitive status.

MR Imaging
All MR imaging studies were performed on a 3T Signa PET/MR scanner (GE Healthcare) using a multichannel receiver (8HRBRAIN) radiofrequency coil. First, 3D volumetric T1-weighted images were acquired (TR =7.7 ms; TE = 3.1 ms; TI = 600ms; in-plane resolution = 0.5 × 0.5 mm;[2] section thickness =
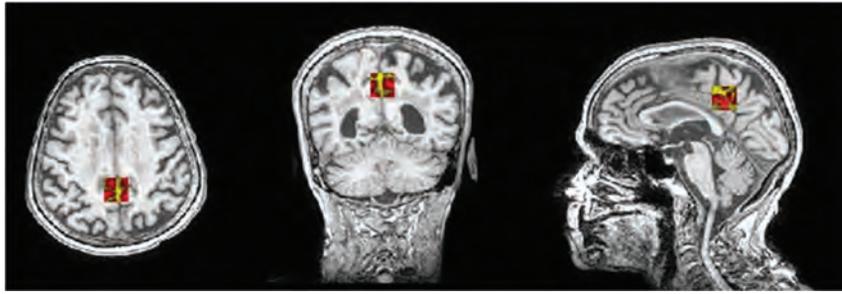
FIG 1. Voxel placement for ^{1}H-MR spectroscopy on the posterior cingulate cortex. CSF (yellow) and gray and white matter (red).

1 mm, isotropic) and were used for ^{1}H-MR spectroscopy voxel placement and to enable subsequent voxel segmentation to address the partial volume averaging. ^{1}H-MR spectroscopy data were acquired with a single-voxel point-resolved spectroscopy sequence, with a bandwidth of 5000 Hz. The standardized VOI measuring 20 × 20 × 20 mm^3 was always positioned on the posterior cingulate cortex as shown in Fig 1. ^{1}H-MR spectroscopy preacquisition adjustments included high-order shimming (automatic and manual) within the selected region and optimization of water suppression. Water-suppressed spectra were then acquired with TE = 35ms; TR = 1500ms; 4096 complex points; and 128 transients. A reference spectrum without water suppression was also obtained.

Data Analysis

All spectroscopic data were analyzed using LCModel (Version 6.3-1; http://www.lcmodel.com/).[28] Model spectra of 18 metabolites were included in the basis data set together with model spectra for macromolecules and lipids.[29] Metabolite levels were estimated using internal water as a reference. A typical value for tissue water content in the gray matter of 43,300 mM was applied.[30] The parameter was adjusted using T2 ∼ 80 ms for the major tissue water component, as previously reported.[31] Metabolite levels were then corrected for partial volume of CSF in the ^{1}H-MR spectroscopy VOI using the following correction factor (CSFcor), CSFcor = 1.0/(1.0-CSF fraction). The CSF fraction was calculated according to standard practice.[32] Briefly, the steps were the following: 1) registration of the ^{1}H-MR spectroscopy VOI to 3D T1-weighted images; 2) segmentation of the 3D T1-weighted volume using the SPM 8 software (http://www.fil.ion.ucl.ac.uk/spm/software/spm12);[33] and 3) use of segmentation results to determine the brain tissue and CSF fractions in the ^{1}H-MR spectroscopy VOI.

Quality control of ^{1}H-MR spectroscopy spectra was performed before the inclusion of LCModel results for further analysis. We used the following criteria: signal-to-noise ratio > 12 (SNR obtained from LCModel analysis and defined as the maximum in a spectrum minus baseline over the Analysis Window to twice the root-mean-square of Residuals), and linewidths at full width at half maximum of 10 Hz (full width at half maximum obtained from LCModel analysis). The cutoff values for SNR and full width at half maximum were established experimentally as conservative thresholds to provide adequate quantitative assessment and consistent data quality across the 2 groups.[34]

Statistical Analysis

For metabolite concentrations, only those estimated values with relative Cramer-Rao lower bounds (CRLB) <20% were retained in the statistical analysis. We also used an alternative approach based on the "weighted mean method."[14]

Descriptive statistics including mean, SD, standard error, median, and quartiles were generated for all study variables and groups. Distribution normality of the variables was assessed using the Shapiro-Wilks test. The means and medians among groups were compared using the 2-sided t test for independent samples and/or the Mann-Whitney U test according to the normality assessment at $P < .05$. The comparison of means between groups was also performed using the weighted t test (weighted linear model) at $P < .05$.

For the weighted mean approach, the weights (Wi) for each metabolite estimate were calculated as the following equation:

$$w_i = \frac{1}{\left[C_i\left(\frac{R_i}{100}\right)\right]^2},$$

where C_i is the estimated concentration of a metabolite, and R_i, the corresponding CRLB (%) measure. When the reliability is very low, CRLB% > 80% in the LCModel, the weight was defined as zero.

Then the weighted mean (wM) and weighted SD (wSD) were defined as

$$wM = \frac{\sum\limits_{i=1}^{N} w_i c_i}{\sum\limits_{i=1}^{N} w_i},$$

and

$$wSD = \sqrt{\frac{\sum\limits_{i=1}^{N} w_i (C_i - wM)^2}{\sum\limits_{i=1}^{N} w_i}},$$

where i is the measurement index and N is the number of measurements.

Qualitative Examination

In addition, a qualitative examination was performed on the basis of visual comparison of the spectral patterns shown by averaging spectra for each group (Fig 2). The difference spectrum was also calculated to highlight any differences among groups. The averaged spectra were calculated following normalization and spectral registration/alignment of individual within-group spectra using the FID-A tool (Matlab; MathWorks).[35]
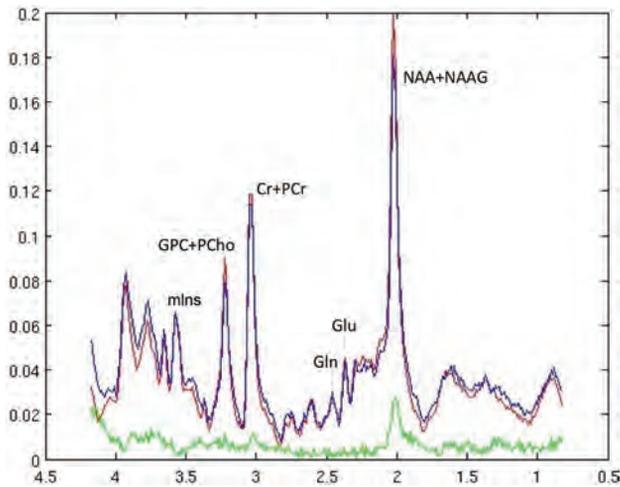
FIG 2. Averaged spectra for superagers (red), age-matched controls (blue), and the difference between groups (green). The most prominent difference is for NAA+NAAG (2.02 ppm), and there is a small difference for total Cr (3.03 ppm) and mIns (~3.56 ppm) metabolites.
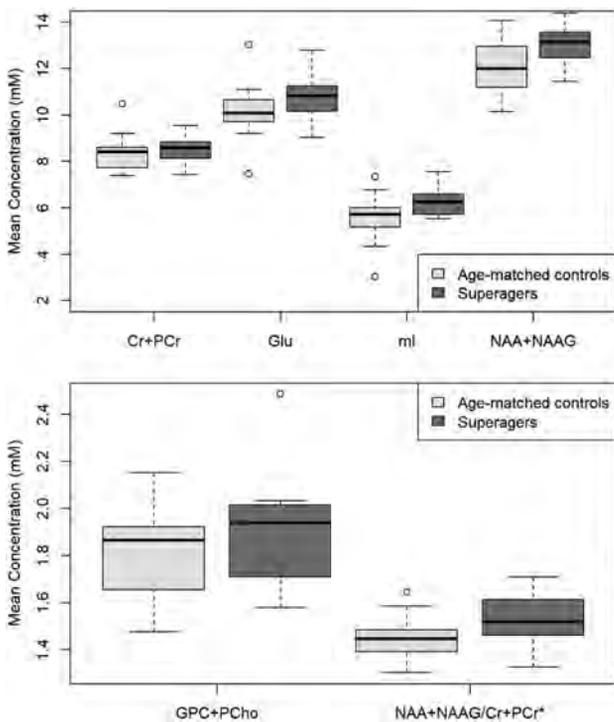
FIG 3. Boxplots showing metabolite concentration differences between superagers and age-matched controls. NAA+NAAG was statistically significantly more elevated in the posterior cingulate cortex of superagers than in age-matched controls ($P = .02$), and mIns tended to be higher in superagers ($P = .06$). Cr+PCr, Glu, GCP+PCho, and NAA+NAAG/Cr+PCr were not statistically significantly different across groups. The *asterisk* means do not consider the unit millimolar (mM) for the ratio NAA+NAAG/Cr+PCr.

RESULTS

Demographics and Neuropsychological Performance of Participants

Fourteen superagers and 15 age-matched controls were assessed through [1]H-MR spectroscopy, but 2 superagers and 2 age-matched

controls were excluded due to a low SNR. Twelve superagers and 13 age-matched controls were included in the statistical analyses (Online Supplemental Data). On the basis of the selection criteria, superagers and age-matched controls did not differ in terms of age, education, or sex distribution (Online Supplemental Data). Superagers had statistically significantly better performance compared with age-matched controls in the Montreal Cognitive Assessment and some episodic memory tests (delayed recall of the Brief Cognitive Screening Battery, delayed recall of the Rey Auditory Verbal Learning Test, and delayed recall of the Logical Memory II). See the Online Supplemental Data for demographic and neuropsychological testing information.

[1]H-MR spectroscopy

The location and size of the standardized voxel used for [1]H-MR spectroscopy are shown in Fig 1. Averaged spectra illustrate the spectral quality consistently obtained in the posterior cingulate cortex (Fig 2). On average, the SNR in the superagers spectra was 22.7 (SD, 3.8) and 19.8 (SD, 4.2) in controls. In the superagers, cortex linewidths full width at half maximum were 5.22 (SD, 0.64) Hz compared with 5.09 (SD, 0.77) Hz in the age-matched controls. Of 18 metabolites, 9 had CRLB% < 20% (a frequently used inclusion criterion indicating acceptable fitting reliability; Online Supplemental Data). Four of them were sums of metabolites, including creatine+phosphocreatine (Cr+PCr); glycerophosphocholine+phosphocholine (GCP+PCho); glutamate+glutamine (Glu+Gln) and *N*- acetyl aspartate+*N*-acetylaspartyl glutamate (NAA+NAAG).

The results from the LCModel analysis were further processed using 2 approaches as aforementioned. First, only the concentrations of metabolites estimated with CRLB% < 20% were considered for statistical analysis and comparisons; in the second, weighted means of the metabolite concentrations with CRLB% < 20% were calculated and used to compare the 2 groups. The motivation for the second approach was to reflect the variable accuracy of fitted concentrations, providing rigor to the former findings. Superagers presented with statistically significantly higher NAA+NAAG concentrations than age-matched controls in both methods ($P = .02$ and $P = .03$ for the weighted means). However, the ratio NAA+NAAG/Cr+PCr did not reach statistical significance ($P = .06$) (Fig 3 and Online Supplemental Data).

mIns tended to be more increased in superagers than in average-age controls when considering metabolite concentration ($P = .06$), but it was not statistically significant when analyzing the weighted mean ($P = .09$). The remaining metabolites quantified reliably, including glycerophosphocholine+phosphocholine (GCP+PCho), Glu, and total creatine (PCr + Cr), did not demonstrate statistically significant differences between groups for both methods of measurement analysis (Fig 3 and Online Supplemental Data).

Tissue Composition

Superagers and age-matched controls did not differ significantly in the means of CSF and gray matter components in the VOIs; however, the groups showed statistically significant differences in the white matter content of the VOI composition, greater in superagers ($P = .04$) (Online Supplemental Data).

DISCUSSION

Our study highlighted and compared the metabolic differences in the posterior cingulate cortex of superagers and age-matched controls to find associations between superior memory performance and neurometabolites detected by [1]H-MR spectroscopy. We evaluated cognitive-related differences independent of age showing higher levels of total NAA (NAA+NAAG) in superagers using 2 different methods of analysis, suggesting a positive association of total NAA with higher cognitive function.

Although some studies have investigated cognitive-related metabolic changes in the normal aging brain,[6,36-41] to our knowledge, this is the first examination of metabolite concentrations in superagers. Our cohorts were made up of subjects older than 80 years of age because the concept of superior cognition as an index of resilience and resistance becomes more critical with age,[42] and subjects between 60 and 80 years of age may be biased to obtain meaningful assertions about "youthful" memory performance.[43,44] Furthermore, the reported literature of this age category is still limited.

We focused on the posterior cingulate cortex because it is one of the primary brain regions to display volume loss and reduced functional integrity during healthy aging and in neurodegenerative disorders.[45,46] Compared with the hippocampal formation, the [1]H-MR spectroscopy in the posterior cingulate cortex shows better homogenization of the magnetic field and fewer artifacts, providing more reliable and reproducible results.[47] The observed differences in neurochemical concentrations in this region appeared to be not noticeably susceptible to differences in the VOI composition between the 2 groups. Even though there was a statistically significant difference in the white matter fraction (superagers = 0.15; age-matched controls = 0.12), the results of the VOI analysis of the posterior cingulate cortex were predominantly driven by the gray matter fraction (superagers = 0.65; age-matched controls = 0.64) (Online Supplemental Data). Moreover, the effect of the CSF content on metabolite concentration was considered during quantification, and no substantial differences were detected in the total tissue fraction content.

Total NAA is the sum of NAA and NAAG. The separation between NAA and NAAG using the 3T [1]H-MR spectroscopy point-resolved spectroscopy sequence is inherently inaccurate; however, the sum of these metabolites can be estimated with good accuracy. NAA is a marker of neuronal and axonal function[38] and may be implicated in mitochondrial activity, representing a potential source of neuronal metabolic efficiency.[48] NAA can also be involved in myelin synthesis and maintenance.[49]

Previous findings in young adults have demonstrated that NAA levels are strongly associated with measures of intelligence,[7-9] creativity,[11] and, more recently, emotional constructs such as agency and flexibility.[50] Superagers, by definition, have superior episodic memory performance, which is an important component of intelligence, according to more recent theoretic models.[51] Therefore, our results support the relevance of the NAA levels in surrogate biomarkers of cognitive performance. Further studies are still necessary to investigate whether superagers may score higher on measures of creativity, emotion, and personality and whether such characteristics can be associated with the neurometabolic markers.

There is an agreement void in the literature regarding NAA levels in the normal-aging brain; however, studies using state-of-the-art single-voxel [1]H-MR spectroscopy showed an overall reduction in NAA,[52-56] notably in the frontal lobes and hippocampus.[6] Whole-brain [1]H-MR spectroscopy studies also found decreased NAA concentrations with aging.[57,58] In line with the literature, our age-matched controls presented with an expected reduction in the NAA concentration, further explaining some divergences among the previous aging studies in which the cognition status was not taken into account.

Even though different cognitive tasks and different VOIs were applied reporting cognitive-related metabolic changes in normal aging, our results in superagers are broadly in agreement with previous studies,[36-41] which have shown a consistent positive association between NAA levels and cognitive performance. The previous studies have demonstrated higher NAA concentrations associated with improved performance on executive function tasks, digit-span tasks, composite processing speeds, memory tasks, and psychomotor processing speeds, supporting our findings in superagers.

Because gray matter loss is a component of aging,[59] the reduction in NAA may be due to neuronal death or shrinkage.[60] The age-related reduction in synaptic density can also explain the decline in NAA concentration and underlie cognitive changes in older adults.[61] It is also known that levels of NAA are reduced in neurodegenerative diseases[62,63] such as Alzheimer disease, and its reduction has been associated with increased amyloid-β markers in normal aging[64] and in subjects with mild cognitive impairment.[65] Thus, [1]H-MR spectroscopy can be a noninvasive tool in vivo to diagnose, prognose, and monitor neurodegenerative disorders.

NAAG, a product of NAA, is a neurotransmitter that might be part of a compensatory neuroprotective mechanism through its actions on the presynaptic group II metabotropic glutamate receptors.[66,67] Harris et al,[14] in 2014, showed significantly higher NAAG in both the hippocampus and cortex of older animals, suggesting a protective response to the constitutively elevated glucose in aged brains. Moreover, NAAG has been shown to protect neurons in vitro from cell death after exposure to high glucose,[68,69] a metabolite that seems to play a key role in dementia pathology.[70]

In a previous study reporting cognitive-related metabolic changes in healthy older adults, Kochunov et al,[39] in 2010, showed a positive relationship between higher Cr and Cho and better memory performance. Another study demonstrated a negative association between higher Cho/Cr and Cho/mIns concentrations and a composite score of global cognition.[41] Our investigation did not find any differences in total Cho (GCP+PCho) in superagers compared with age-matched controls. Cho concentration seems to be elevated with age independent of cognitive status, notably in the posterior cingulate cortex and centrum semiovale,[17,53,71] reflecting, most likely, increased cell membrane turnover and breakdown.[14,56]

Our results suggest a lack of association between cognitive performance and the concentration of total Cr demonstrated in normal aging.[36] In addition, when considering total Cr as a denominator for the relative NAA+NAAG concentration (NAA+NAAG/Cr+PCr), we could not detect statistically significant differences between superagers and age-matched controls,

supporting previous findings that Cr changes are inconsistently associated with age.[6] On the basis of this inconsistent concentration pattern, Cr must be avoided to determine metabolite ratios in elderly samples and can be misleading when considered as a biomarker in this population.[16,36]

Our analysis suggested a trend toward increased mIns in superagers compared with the age-matched controls. mIns has been described as a marker of glial proliferation.[72,73] Nevertheless, this is questionable because brain tissue histopathology studies have not shown a statistically significant relationship between astrogliosis and mIns.[74] Ultimately, mIns has been proposed to function as an osmolyte in the brain and is potentially implicated in brain cell signaling.[75,76]

Glu is a neurotransmitter involved in cognition, learning, and memory.[56,77] Previous studies in dementia showed decreased Glu levels in mild cognitive impairment and notably in Alzheimer disease compared with healthy older adults.[78,79] Other studies have shown Glu reduction with age.[53,80] Among healthy older adults, when controlling for age, the literature lacks Glu examination for cognitive performance, and our study did not find differences in Glu concentration between the 2 groups. Nonetheless, more studies are still necessary to characterize the Glu role in cognition among the healthy elderly population.

The main strengths of our study were the use of a detailed and validated neuropsychological protocol to stratify older adults (older than 80 years of age) into superagers and age-matched controls and the implementation of standardized [1]H-MR spectroscopy acquisition.[17] Second, the VOI standardization and placement in the posterior cingulate cortex contribute to the generation of robust results that allow them to be easily validated in another setting. Finally, the implemented postprocessing methods, including measures of spectra quality and tissue/CSF segmentation to address the partial volume averaging in the [1]H-MR spectroscopy measurements, enhanced the quality and reproducibility of our results.[81]

Nevertheless, this study also presents some weaknesses. The published data concerning the prevalence of superagers in the global population are still insufficient[82] and did not allow us to undertake a power sample analysis before study conduction. In previous studies using the same inclusion criteria, the sample size of superagers varied from 12[2] to 56.[83] The examined metabolite concentrations were limited to those sampled in the posterior cingulate cortex; however, metabolite concentrations differ among distinct portions of the brain and between gray and white matter,[84] possibly having an unpredictable impact on normal aging and the superagers' metabolic signatures. Also, no inferences regarding the impact of aging on several metabolite levels could be drawn. Future studies with whole-brain [1]H-MR spectroscopy and control groups with younger volunteers would be useful to understand better metabolic profile changes during aging in different brain regions.

CONCLUSIONS

The present single-voxel [1]H-MR spectroscopy study provides in vivo evidence that superior memory performance in late life is positively associated with total NAA in the posterior cingulate cortex. These findings indicate that higher total NAA can contribute to the resilience process of the conventional pathways of aging present in superagers. Expanding on the current results, future evidence accumulation will probably qualify [1]H-MR spectroscopy as a diagnostic means for the quantification of neurochemical biomarkers in the aging population and as a prognostic tool including monitoring interventional therapies to preserve or enhance cognition in later life.

REFERENCES

1. de Godoy LL, Alves CA, Saavedra JS, et al. **Understanding brain resilience in superagers: a systematic review.** *Neuroradiology* 2021;63:663–83 CrossRef Medline
2. Harrison TM, Weintraub S, Mesulam MM, et al. **Superior memory and higher cortical volumes in unusually successful cognitive aging.** *J Int Neuropsychol Soc* 2012;18:1081–85 CrossRef Medline
3. Gefen T, Peterson M, Papastefan ST, et al. **Morphometric and histologic substrates of cingulate integrity in elders with exceptional memory capacity.** *J Neurosci* 2015;35:1781–91 CrossRef Medline
4. Wang X, Ren P, Baran TM, et al. Alzheimer's Disease Neuroimaging Initiative. **Longitudinal functional brain mapping in supernormals.** *Cereb Cortex* 2019;29:242–52 CrossRef Medline
5. Zhang J, Andreano JM, Dickerson BC, et al. **Stronger functional connectivity in the default mode and salience networks is associated with youthful memory in superaging.** *Cereb Cortex* 2020;30:72–84 CrossRef Medline
6. Cleeland C, Pipingas A, Scholey A, et al. **Neurochemical changes in the aging brain: a systematic review.** *Neurosci Biobehav Rev* 2019;98:306–19 CrossRef Medline
7. Jung RE, Gasparovic C, Chavez RS, et al. **Imaging intelligence with proton magnetic resonance spectroscopy.** *Intelligence* 2009;37:192–98 CrossRef Medline
8. Jung RE, Yeo RA, Chiulli SJ, et al. **Myths of neuropsychology: intelligence, neurometabolism, and cognitive ability.** *Clin Neuropsychol* 2000;14:535–45 CrossRef Medline
9. Jung RE, Brooks WM, Yeo RA, et al. **Biochemical markers of intelligence: a proton MR spectroscopy study of normal human brain.** *Proc Bio Sci* 1999;266:1375–79 CrossRef Medline
10. Jung RE, Yeo RA, Love TM, et al. **Biochemical markers of mood: a proton magnetic resonance spectroscopy study of normal human brain.** *Biol Psychiatry* 2002;51:224–29 CrossRef Medline
11. Jung RE, Gasparovic C, Chavez RS, et al. **Biochemical support for the "Threshold" theory of creativity: a magnetic resonance spectroscopy study.** *J Neurosci* 2009;29:5319–25 CrossRef Medline
12. Ryman SG, Gasparovic C, Bedrick EJ, et al. **Brain biochemistry and personality: a magnetic resonance spectroscopy study.** *PLoS One* 2011;6:e26758 CrossRef Medline

13. Lee MR, Denic A, Hinton DJ, et al. **Preclinical (1)H-MRS neurochemical profiling in neurological and psychiatric disorders.** *Bioanalysis* 2012;4:1787–804 CrossRef Medline

14. Harris JL, Yeh HW, Swerdlow RH, et al. **High-field proton magnetic resonance spectroscopy reveals metabolic effects of normal brain aging.** *Neurobiol Aging* 2014;35:1686–94 CrossRef Medline

15. Schmitt F, Grosu D, Mohr C, et al. **3 Tesla MRI: successful results with higher field strengths.** *Radiologe* 2004;44:31–47 CrossRef Medline

16. Jansen JF, Backes WH, Nicolay K, et al. **1H MR spectroscopy of the brain: absolute quantification of metabolites.** *Radiology* 2006;240:318–32 CrossRef Medline

17. Chiu PW, Mak HK, Yau KK, et al. **Metabolic changes in the anterior and posterior cingulate cortices of the normal aging brain: proton magnetic resonance spectroscopy study at 3T.** *Age (Dordr)* 2014;36:251–64 CrossRef Medline

18. Oz G, Alger JR, Barker PB, et al. MRS Consensus Group. **Clinical proton MR spectroscopy in central nervous system disorders.** *Radiology* 2014;270:658–79 CrossRef Medline

19. Folstein MF, Folstein SE, McHugh PR. **Mini-Mental State: a practical guide for grading the mental state of patients for the clinician.** *J Psych Res* 1975;12:189–98 CrossRef Medline

20. Brucki S, Nitrini R, Caramelli P, et al. **Suggestions for utilization of the mini-mental state examination in Brazil** [in Portuguese]. *Arq Neuropsiquiatr* 2003;61:777–81 CrossRef Medline

21. Pfeffer RI, Kurosaki TT, Harrah CH Jr, et al. **Measurement of functional activities in older adults in the community.** *J Gerontol* 1982;37:323–29 CrossRef Medline

22. McKhann GM, Knopman DS, Chertkow H, et al. **The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease.** *Alzheimers Dement* 2011;7:263–69 CrossRef Medline

23. Albert MS, DeKosky ST, Dickson D, et al. **The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease.** *Alzheimers Dement* 2011;7:270–79 CrossRef Medline

24. Nitrini R, Caramelli P, Porto CS, et al. **Brief cognitive battery in the diagnosis of mild Alzheimer's disease in subjects with medium and high levels of education.** *Dement Neuropsychol* 2007;1:32–36 CrossRef Medline

25. Rogalski EJ, Gefen T, Shi J, et al. **Youthful memory capacity in old brains: anatomic and genetic clues from the Northwestern SuperAging Project.** *J Cogn Neurosci* 2013;25:29–36 CrossRef Medline

26. Heaton RK, Miller SW, Taylor MJ, et al. Revised comprehensive norms for an expanded Halstead-Reitan Battery: Demographically Adjusted Neuropsychological Norms For African American and Caucasian adults. 2004. http://www4.parinc.com/Products/Product.aspx?ProductID=RCNAAC. Accessed September 10, 2020

27. Shirk SD, Mitchell MB, Shaughnessy LW, et al. **A web-based normative calculator for the Uniform Data Set (UDS) neuropsychological test battery.** *Alzheimers Res Ther* 2011;3:32 CrossRef Medline

28. Provencher SW. **Estimation of metabolite concentrations from localized in vivo proton NMR spectra.** *Magn Reson Med* 1993;30:672–79 CrossRef Medline

29. Seeger U, Klose U, Mader I, et al. **Parameterized evaluation of macromolecules and lipids in proton MR spectroscopy of brain diseases.** *Magn Reson Med* 2003;49:19–28 CrossRef Medline

30. Kreis R, Ernst T, Ross BD. **Absolute quantitation of water and metabolites in the human brain, II: metabolite concentrations.** *J Magn Reson B* 1993;102:9–19 CrossRef

31. Ernst T, Kreis R, Ross BD. **Absolute quantitation of water and metabolites in the human brain, I: compartments and water.** *J Magn Reson B* 1993;102:1–8 CrossRef

32. Edden RA, Puts NA, Harris AD, et al. **Gannet: a batch-processing tool for the quantitative analysis of gamma-aminobutyric acid–edited MR spectroscopy spectra.** *J Magn Reson Imaging* 2014;40:1445–52 CrossRef Medline

33. Friston KJ, Holmes AP, Worsley KJ, et al. **Statistical parametric maps in functional imaging: a general linear approach.** *Hum Brain Mapp* 1994;2:189–210 CrossRef

34. Pedrosa de Barros N, Slotboom J. **Quality management in in vivo proton MRS.** *Anal Biochem* 2017;529:98–116 CrossRef Medline

35. Simpson R, Devenyi GA, Jezzard P, et al. **Advanced processing and simulation of MRS data using the FID appliance (FID-A): an open source, MATLAB-based toolkit.** *Magn Reson Med* 2017;77:23–33 CrossRef Medline

36. Charlton RA, McIntyre DJO, Howe FA, et al. **The relationship between white matter brain metabolites and cognition in normal aging: the GENIE study.** *Brain Res* 2007;1164:108–16 CrossRef Medline

37. Driscoll I, Hamilton DA, Petropoulos H, et al. **The aging hippocampus: cognitive, biochemical and structural findings.** *Cereb Cortex* 2003;13:1344–51 CrossRef Medline

38. Ross AJ, Sachdev PS, Wen W, et al. **Cognitive correlates of 1H MRS measures in the healthy elderly brain.** *Brain Res Bull* 2005;66:9–16 CrossRef Medline

39. Kochunov P, Coyle T, Lancaster J, et al. **Processing speed is correlated with cerebral health markers in the frontal lobes as quantified by neuroimaging.** *Neuroimage* 2010;49:1190–99 CrossRef Medline

40. Erickson KI, Weinstein AM, Sutton BP, et al. **Beyond vascularization: aerobic fitness is associated with N-acetylaspartate and working memory.** *Brain Behav* 2012;2:32–41 CrossRef Medline

41. Gomar JJ, Gordon ML, Dickinson D, et al. **APOE genotype modulates proton magnetic resonance spectroscopy metabolites in the aging brain.** *Biol Psychiatry* 2014;75:686–92 CrossRef Medline

42. Rogalski EJ. **Don't forget: age is a relevant variable in defining SuperAgers.** *Alzheimers Dement (Amst)* 2019;11:560–61 CrossRef Medline

43. Nyberg L, Lövdén M, Riklund K, et al. **Memory aging and brain maintenance.** *Trends Cogn Sci* 2012;16:292–305 CrossRef Medline

44. Borelli WV, Carmona KC, Studart-Neto A, et al. **Operationalized definition of older adults with high cognitive performance.** *Dement Neuropsychol* 2018;12:221–27 CrossRef Medline

45. Minoshima S, Giordani B, Berent S, et al. **Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease.** *Ann Neurol* 1997;42:85–94 CrossRef Medline

46. Lehmann M, Rohrer JD, Clarkson MJ, et al. **Reduced cortical thickness in the posterior cingulate gyrus is characteristic of both typical and atypical Alzheimer's disease.** *J Alzheimers Dis* 2010;20:587–98 CrossRef Medline

47. Park EJ, Lyra KP, Lee HW, et al. **Correlation between hippocampal volumes and proton magnetic resonance spectroscopy of the posterior cingulate gyrus and hippocampi in Alzheimer's disease.** *Dement Neuropsychol* 2010;4:109–13 CrossRef Medline

48. Scavuzzo CJ, Moulton CJ, Larsen RJ. **The use of magnetic resonance spectroscopy for assessing the effect of diet on cognition.** *Nutr Neurosci* 2018;21:1–15 CrossRef Medline

49. Chakraborty G, Mekala P, Yahya D, et al. **Intraneuronal N-acetylaspartate supplies acetyl groups for myelin lipid synthesis: evidence for myelin-associated aspartoacylase.** *J Neurochem* 2001;78:736–45 CrossRef Medline

50. White TL, Gonsalves MA, Cohen RA, et al. **The neurobiology of wellness: 1H-MRS correlates of agency, flexibility and neuroaffective reserves in healthy young adults.** *Neuroimage* 2021;225:117509 CrossRef Medline

51. Jewsbury PA, Bowden SC, Duff K. **The Cattell–Horn–Carroll Model of cognition for clinical assessment.** *J Psychoeduc Assess* 2017;35:547–67 CrossRef

52. Hädel S, Wirth C, Rapp M, et al. **Effects of age and sex on the concentrations of glutamate and glutamine in the human brain.** *J Magn Reson Imaging* 2013;38:1480–87 CrossRef Medline

53. Marjańska M, Riley McCarten J, Hodges J, et al. **Region-specific aging of the human brain as evidenced by neurochemical profiles measured noninvasively in the posterior cingulate cortex and the**

occipital lobe using 1H magnetic resonance spectroscopy at 7 T. *Neuroscience* 2017;354:168–77 CrossRef Medline

54. Sailasuta N, Ernst T, Chang L. **Regional variations and the effects of age and gender on glutamate concentrations in the human brain.** *Magnetic Resonance Imaging* 2008;26:667–75 CrossRef Medline

55. Yang ZY, Yue Q, Xing HY, et al. **A quantitative analysis of (1)H-MR spectroscopy at 3.0 T of three brain regions from childhood to middle age.** *Br J Radiol* 2015;88:20140693 CrossRef Medline

56. Zahr NM, Mayer D, Rohlfing T, et al. **In vivo glutamate measured with magnetic resonance spectroscopy: behavioral correlates in aging.** *Neurobiol Aging* 2013;34:1265–76 CrossRef Medline

57. Eylers VV, Maudsley AA, Bronzlik P, et al. **Detection of normal aging effects on human brain metabolite concentrations and microstructure with whole-brain MR spectroscopic imaging and quantitative MR imaging.** *AJNR Am J Neuroradiol* 2016;37:447–54 CrossRef Medline

58. Ding, XQ, Maudsley, AA, Sabati, M, et al. **Physiological neuronal decline in healthy aging human brain: an in vivo study with MRI and short echo-time whole-brain 1H MR spectroscopic imaging.** *Neuroimage* 2016;137:45–51 CrossRef Medline

59. Drachman DA. **Aging of the brain, entropy, and Alzheimer disease.** *Neurology* 2006;67:1340–52 CrossRef Medline

60. Peters A, Morrison JH, Rosene DL, et al. **Feature article: are neurons lost from the primate cerebral cortex during normal aging?** *Cereb Cortex* 1998;8:295–300 CrossRef Medline

61. Woods DL, Wyma JM, Herron TJ, et al. **The effects of aging, malingering, and traumatic brain injury on computerized trail-making test performance.** *PLoS One* 2015;10:e0124345 CrossRef Medline

62. Graff-Radford J, Kantarci K. **Magnetic resonance spectroscopy in Alzheimer's disease.** *Neuropsychiatr Dis Treat* 2013;9:687–96 CrossRef Medline

63. Coutinho A, Porto FH, Zampieri PF, et al. **Analysis of the posterior cingulate cortex with [18 F] FDG-PET and NAA/mI in mild cognitive impairment and Alzheimer's disease: correlations and differences between the two methods.** *Dement Neuropsychol* 2015;9:385–93 CrossRef Medline

64. Nedelska Z, Przybelski SA, Lesnick TG, et al. **1H-MRS metabolites and rate of β-amyloid accumulation on serial PET in clinically normal adults.** *Neurology* 2017;89:1391–99 CrossRef Medline

65. Voevodskaya O, Poulakis K, Sundgren P, et al. Swedish BioFINDER Study Group. **Brain myoinositol as a potential marker of amyloid-related pathology: a longitudinal study.** *Neurology* 2019;92:e395–405 CrossRef

66. Sanabria ER, Wozniak KM, Slusher BS, et al. **GCP II (NAALADase) inhibition suppresses mossy fiber-CA3 synaptic neurotransmission by a presynaptic mechanism.** *J Neurophysiol* 2004;91:182–93 CrossRef Medline

67. Zhong C, Zhao X, Van KC, et al. **NAAG, peptidase inhibitor increases dialysate NAAG and reduces glutamate, aspartate and GABA levels in the dorsal hippocampus following fluid percussion injury in the rat.** *J Neurochem* 2006;97:1015–25 CrossRef Medline

68. Berent-Spillson A, Robinson AM, Golovoy D, et al. **Protection against glucose-induced neuronal death by NAAG and GCP II inhibition is regulated by mGluR3.** *J Neurochem* 2004;89:90–99 CrossRef Medline

69. Berent-Spillson A, Russell JW. **Metabotropic glutamate receptor 3 protects neurons from glucose-induced oxidative injury by increasing intracellular glutathione concentration.** *J Neurochem* 2007;101:342–54 CrossRef Medline

70. Kuehn BM. **In Alzheimer research, glucose metabolism moves to center stage.** *JAMA* 2020;323:297–99 CrossRef Medline

71. Gruber S, Pinker K, Riederer F, et al. **Metabolic changes in the normal ageing brain: consistent findings from short and long echo time proton spectroscopy.** *Eur J Radiol* 2008;68:320–27 CrossRef Medline

72. David JP, Ghozali F, Fallet-Bianco C, et al. **Glial reaction in the hippocampal formation is highly correlated with aging in human brain.** *Neurosci Lett* 1997;235:53–56 CrossRef Medline

73. Sandhir R, Onyszchuk G, Berman NE. **Exacerbated glial response in the aged mouse hippocampus following controlled cortical impact injury.** *Exp Neurol* 2008;213:372–80 CrossRef Medline

74. Duarte JM, Do KQ, Gruetter R. **Longitudinal neurochemical modifications in the aging mouse brain measured in vivo by 1H magnetic resonance spectroscopy.** *Neurobiol Aging* 2014;35:1660–68 CrossRef Medline

75. Novotny EJ Jr, Fulbright RK, Pearl PL, et al. **Magnetic resonance spectroscopy of neurotransmitters in human brain.** *Ann Neurol* 2003;54(Suppl 6):S25–31 CrossRef Medline

76. Hoyer C, Gass N, Weber-Fahr W, et al. **Advantages and challenges of small animal magnetic resonance imaging as a translational tool.** *Neuropsychobiology* 2014;69:187–201 CrossRef Medline

77. Ramadan S, Lin A, Stanwell P. **Glutamate and glutamine: a review of in vivo MRS in the human brain.** *NMR Biomed* 2013;26:1630–46 CrossRef Medline

78. Zeydan B, Deelchand DK, Tosakulwong N, et al. **Decreased glutamate levels in patients with amnestic mild cognitive impairment: an sLASER proton MR spectroscopy and PiB-PET study.** *J Neuroimaging* 2017;27:630–36 CrossRef Medline

79. Wong D, Atiya S, Fogarty J, et al. **Reduced hippocampal glutamate and posterior cingulate N-acetyl aspartate in mild cognitive impairment and Alzheimer's disease is associated with episodic memory performance and white matter integrity in the cingulum: a pilot study.** *J Alzheimers Dis* 2020;73:1385–405 CrossRef Medline

80. Suri S, Emir U, Stagg CJ, et al. **Effect of age and the APOE gene on metabolite concentrations in the posterior cingulate cortex.** *Neuroimage* 2017;152:509–16 CrossRef Medline

81. Reyngoudt H, Claeys T, Vlerick L, et al. **Age-related differences in metabolites in the posterior cingulate cortex and hippocampus of normal ageing brain: a 1H-MRS study.** *Eur J Radiol* 2012;81:e223–31 CrossRef

82. Kim BR, Kwon H, Chun MY, et al. **White matter integrity is associated with the amount of physical activity in older adults with super-aging.** *Front Aging Neurosci* 2020;12:549983 CrossRef Medline

83. Huentelman MJ, Piras IS, Siniard AL, et al. **Associations of MAP2K3 gene variants with superior memory in SuperAgers.** *Front Aging Neurosci* 2018;10:155 CrossRef Medline

84. Kreis R, Slotboom J, Hofmann L, et al. **Integrated data acquisition and processing to determine metabolite contents, relaxation times, and macromolecule baseline in single examinations of individual subjects.** *Magn Reson Med* 2005;54:761–68 CrossRef Medline

Rescan Time Delays in Ischemic Stroke Imaging: A Retrospective Observation and Analysis of Causes and Clinical Impact

J.M. Katz, J.J. Wang, A.T. Boltyenkov, G. Martinez, J. O'Hara, C. Feizullayeva, M. Gribko, A. Pandya, and P.C. Sanelli

ABSTRACT

BACKGROUND AND PURPOSE: Delays to reperfusion negatively impact outcomes of patients with ischemic stroke, yet current guidelines recommend selective sequential imaging for thrombectomy candidates. We aimed to quantify and analyze time delays associated with rescanning in sequential acute stroke imaging.

MATERIALS AND METHODS: This was a retrospective cohort study of consecutive patients with acute ischemic stroke who underwent imaging for treatment decision-making from January 1, 2017, to June 30, 2020. Rescan time delay was defined as ≥10-minute difference between initial NCCT and CTA ± CTP. Mean rescan time delays in comprehensive and primary stroke centers were compared. Bivariate and multivariable regression analyses assessed clinical and imaging factors associated with rescanning time delays and early outcomes.

RESULTS: A total of 588 patients with acute ischemic were included in statistical analyses. Rescanning occurred in 27.9% (164/588 patients), with a mean time delay of 53.7 (SD, 43.4) minutes. For patients presenting at primary compared with comprehensive stroke centers, rescan time delays were more common (59.6% versus 11.8%, $P < .001$), with longer delays (65.4 [SD, 45.4] minutes versus 23.6 [SD, 14.0] minutes, $P < .001$). Independent predictors of rescan time delays included primary stroke center presentation, intravenous thrombolysis administration, black race, admission NIHSS ≥10, baseline independent ambulation, and onset-to-comprehensive stroke center arrival in ≥6 hours. Protocols for early simultaneous comprehensive CT (NCCT + CTA + CTP) were associated with lower odds of time delays (OR = 0.34; 95% CI, 0.21–0.55). Rescanning was associated with lower odds of home discharge (OR = 0.53; 95% CI, 0.30–0.95).

CONCLUSIONS: A sequential approach to CT-based imaging may be significantly associated with prolonged acute stroke evaluations. Adoption of early simultaneous comprehensive CT could minimize treatment delays and improve outcomes.

ABBREVIATIONS: CTAP = CTA perfusion; CSC-G = comprehensive stroke center imaging subgroup; PSC-G = primary stroke center imaging subgroup

Treating select patients with acute ischemic stroke with endovascular thrombectomy up to 24 hours from last known well increases the proportion of patients with large-vessel occlusion who have an independent functional outcome.[1-7] The robust benefits of endovascular stroke treatment have been shown to be time-dependent in numerous analyses of randomized clinical trials and a large registry, in which the benefits decrease as time to reperfusion increases in a nonlinear fashion.[8-11] Stroke systems of care initiatives focus on improving the efficiency of acute stroke management to maximize favorable clinical outcomes. For instance, the American Heart Association Target: Stroke Phase III initiative (https://www.heart.org/en/professional/quality-improvement/target-stroke/introducing-target-stroke-phase-iii) challenges comprehensive stroke centers to achieve door-to-device times of <60 minutes for transfer patients and <90 minutes for direct emergency department arrivals.[12]

The most recent American Heart Association guidelines for the management of patients with acute ischemic stroke recommend NCCT of the head as the initial imaging technique for

triage of patients with acute stroke.[13] Comprehensive CT imaging with head and neck CTA and perfusion CTA (CTAP) is recommended specifically for patients who may be endovascular candidates, with CTP only for patients in the >6- to 24-hour time window. Consequently, imaging practices have evolved to perform initial NCCT first to rapidly determine eligibility for intravenous thrombolysis; then, patients are returned to the scanner for CTA ± CTP, particularly when thrombectomy is considered.[13-15] The main concern is that these inefficient imaging practices have potentially led to unnecessary time delays to endovascular treatment, translating to worse clinical outcomes.[16-19] While there have been several studies that report the mean time it takes to rescan patients with acute stroke with urgent MR imaging following an initial NCCT or CTAP in an acute trial setting,[20-23] no study to date has quantified the time delay associated with rescanning for advanced CT-based imaging following an initial NCCT in real-world practice.

In this study, we aimed to quantify the time it takes to rescan patients with acute ischemic stroke with CTA ± CTP following NCCT at comprehensive and primary stroke centers within the same stroke network. We then assessed the clinical and imaging factors associated with rescan time delays and analyzed whether these delays impacted early clinical outcomes. Our hypothesis was that patients with rescan time delays will have worse clinical outcomes.

MATERIALS AND METHODS

This was a retrospective cohort study of consecutive patients with acute ischemic stroke admitted to the North Shore University Hospital comprehensive stroke center between January 1, 2017, and June 30, 2020. Institutional review board approval was obtained with a waiver of consent because of the retrospective nature of this study. Demographic, clinical, treatment, and outcome data were imported from the Get-With-The-Guidelines-Stroke data base (https://www.clinicaltrials.gov/ct2/show/NCT02693223) into a separate Stroke Health Outcomes database. The Stroke Health Outcomes database incorporates additional patient-level details regarding imaging performed during the index stroke admission, identified by unique identifiers in the health system's PACS. Imaging examinations were collected chronologically from 48 hours prearrival from another institution (ie, primary stroke center) until the discharge date, with each imaging examination assigned as a separate entry in the patient's record in the Stroke Health Outcomes database.

Because rescan time delays predominantly affect patients with acute stroke who undergo treatment with intravenous thrombolysis and/or endovascular thrombectomy, the study inclusion aimed to select patients with the following criteria: patients with suspected acute stroke who were likely considered potential treatment candidates and underwent early comprehensive CT imaging with CTA ± CTP, with the likely intent of evaluating for acute stroke intervention and within treatment decision-making timeframes sufficient to meet the Target: Stroke Phase III initiative's 90-minute door-to-device goal.[12] Our imaging timeframe criteria further aimed to exclude patients being imaged only for management purposes in secondary stroke prevention. Therefore, taking a more conservative approach, we chose different imaging timeframes, depending on the mode of arrival to the

comprehensive stroke center, to increase the accuracy of patient inclusion in the study: 1) Direct arrivals had to have the first NCCT and CTA or CTAP performed within 1 hour of hospital arrival to differentiate patients who were potential treatment candidates from those being imaged for secondary stroke prevention management decisions. 2) Transfer patients from a primary stroke center had to be transferred for thrombectomy evaluation and to have undergone CTA at the referring site. To confirm this process, we required that transfer patients meet additional criteria, including the following: 1) NCCT and CTA were performed at the referring site to assess thrombectomy eligibility within the 6 hours before comprehensive stroke center door time, and 2) CTP (only performed for thrombectomy evaluations at our institution) was the first imaging completed at the comprehensive stroke center within 1 hour of arrival. Direct arrivals defined the comprehensive stroke center imaging subgroup (CSC-G), and transfer patients constituted the primary stroke center imaging subgroup (PSC-G). Exclusion criteria were patients with acute stroke younger than 18 years of age and incomplete imaging data. A flow diagram illustrating the final selection of patients for inclusion in the statistical analysis is shown in Fig 1.

Data collection included the type of imaging examination (NCCT, CTA, CTP) and the date and earliest imaging timestamp in the PACS. The recorded timestamps reference the start of the imaging performance. However, when multiple imaging examinations were completed in a single scanning session and assigned the same accession number, the earliest timestamp was used for all examinations within the imaging session. Only the pretransfer imaging timestamps were used for the PSC-G to exclude the transfer time to the comprehensive stroke center in the rescan time. Imaging during off-hours was defined as overnight (5:00 PM to 8:00 AM) and weekend days. Other data included demographic characteristics (age, sex, race [black, white, Asian, unknown/other], health insurance [private, Medicare/Medicaid, uninsured/self-pay], median household income [<$80,000, $80,000–$120,000, and >$120,000, based on ZIP code and census data]);[24] clinical factors (admission NIHSS score, arrival mode [ambulance, private transportation] and time [year, day, time], onset-to-arrival time [defined as last known well to comprehensive stroke center door time], treatment [intravenous thrombolysis, thrombectomy, none], length of hospital stay [0–7 days and ≥8 days], and the time interval between NCCT and CTA ± CTP imaging); and early outcomes (discharge disposition [home, rehabilitation facility, died/hospice], and discharge mRS). On the basis of the discharge mRS scores, good clinical outcome was defined as functional independence (mRS 0–2) at discharge. Age was dichotomized as younger than 80 years versus 80 years of age or older because of the negative impact of older age (80 years of age or older) on stroke outcomes.[25,26] Data collection on race, based on self-reported identification of racial categories, and analyses of racial disparities in healthcare delivery, was included in this study according to National Institutes of Health policy[27] and the American Heart Association position statement recommendations.[28]

During the study period, CTP was rarely performed at the primary stroke centers in our healthcare system. At the comprehensive stroke center, CTP was first adopted for evaluating selected patients for thrombectomy candidacy in late 2017 on the basis of

FIG 1. Flowchart of patient selection. "Arrival *asterisk*" refers to door time at the comprehensive stroke center.

separately; and 3) based on our experience, a 10-minute time interval allows the additional time that may be required in some patients to set up the next scanning protocol, intravenous access issues, and administration of intravenous thrombolytics on the table when appropriate. Our selection of 10 minutes as the threshold defining time delay is further supported by previous literature reporting median time intervals of 6 minutes[29] (CT + CTA) and 12 minutes[30] (CTAP) for acquisition of single-session comprehensive CT imaging in patients with acute ischemic stroke, with the latter study including time for CTP processing and analysis. For the CSC-G, the rescan time delay was calculated as the time difference between the initial NCCT and either CTA or CTAP. When NCCT and CTA were performed in the same session and CTP was in a separate session (within 60 minutes from NCCT), then the difference between NCCT-CTA and CTP was used. For the PSC-G, rescan time delay was defined as the difference between NCCT and CTA at the primary stroke center before transfer.

For the primary analyses, we combined the CSC-G and PSC-G data to stratify patients with and without rescan time delays (≥10 minutes). The demographic characteristics, clinical factors, and early outcomes in patients with and without rescan time delays were compared using bivariate analyses. Independent-samples *t* tests and rank sum tests analyzed continuous variables, and χ^2 tests analyzed categoric variables. Multivariable logistic regression was performed to separately analyze the association between demographic/clinical parameters and rescan time delays, and rescan time delays plus demographic/clinical parameters and early outcomes. All statistically significant variables from the bivariate analyses and clinically important factors, such as age, sex, median household income, and onset-to-arrival time of >6 hours, were included in the regression models. When >2 subcategories of a categoric variable were available to move forward into the multivariable logistic regression analysis, we selected either the variable with the largest difference between groups (from the bivariate analysis), or when numeric differences were small, we selected the subcategory hypothesized to be most clinically relevant, with the other subcategories combined into the reference category. The Cochran-Mantel-Haenszel shift test was used to perform mRS and discharge disposition shift analyses for different time-delay intervals.

Secondary analyses were then performed to assess selection bias that may have influenced our primary results. We applied a uniform 1- or 6-hour door-to-imaging window limit to select

the results of the first extended time window thrombectomy trial.[6] In 2019, acute stroke imaging with NCCT and CTAP during the same session became protocolized at the comprehensive stroke center as routine imaging for virtually all patients with stroke presenting within 24 hours of last known well. Similarly, in 2019, owing to adoption of local early comprehensive CT imaging protocols at many of the primary stroke centers, CTA performance in tandem with the NCCT at these referring sites became more routine. To reflect these protocol changes in our analyses, admission year was dichotomized into pre- (2017 and 2018) and postprotocol (2019 and 2020) implementation time periods.

Statistical Analysis

For comparative analyses, we classified patients with and without rescan time delays, defined as a difference of ≥10 minutes between the initial NCCT and CTA ± CTP. Ten minutes was chosen as a cutoff threshold because of the following: 1) Patients with simultaneous comprehensive CT imaging (CTAP) have a single accession number for all studies and will have a zero-minute time difference between the initial NCCT and CTA ± CTP examinations; 2) we wanted to exclude single-session imaging with different accession numbers from the time-delay group, a situation that may occur when multiple orders are placed

Multivariable logistic regression analysis of clinical and imaging factors associated with rescan time delay and home discharge

Parameter	Rescan Time Delay		Home Discharge	
	OR (95% CI)	P Value	OR (95% CI)	P Value
Age, younger than 80 yr	0.71 (0.42–1.20)	.200	1.53 (0.96–2.43)	.075
Sex, female	0.94 (0.59–1.48)	.772	1.46 (0.96–2.23)	.078
Race, black	2.34 (1.31–4.17)	.004	1.16 (0.66–2.05)	.615
MHI <$80,000	0.71 (0.45–1.14)	.160	1.25 (0.82–1.90)	.296
NIHSS ≥10	1.72 (1.01–2.92)	.045	0.10 (0.06–0.17)	<.001
Baseline ambulatory	4.34 (1.77–10.66)	.001	1.41 (0.69–2.86)	.349
Onset-to-arrival >6 hr	1.98 (1.14–3.43)	.015	0.64 (0.39–1.03)	.063
2019–2020	0.34 (0.21–0.55)	<.001	0.73 (0.47–1.11)	.141
Off-hours	1.14 (0.71–1.82)	.588	0.80 (0.52–1.23)	.294
Arrival by ambulance	1.15 (0.66–1.98)	.628	0.49 (0.32–0.77)	.002
PSC-G	9.40 (5.76–15.34)	<.001	0.85 (0.49–1.49)	.571
IV rtPA	2.27 (1.33–3.87)	.003	1.80 (1.07–3.03)	.026
Thrombectomy	0.68 (0.37–1.25)	.215	0.54 (0.24–1.20)	.129
Rescan delay	NA	NA	0.53 (0.30–0.95)	.032

Note:—MHI indicates median household income, NA; not applicable.

patients for both the CSC-G and PSC-G. The univariate and regression analyses were performed separately for these alternate scenarios (1- and 6-hour time limits). Furthermore, distinct bivariate and regression analyses were conducted for the CSC-G and PSC-G separately. Statistical analyses were performed using SAS, Version 9.4 (SAS Institute). A P value < .05 was considered statistically significant.

RESULTS

During the study period, a total of 2216 consecutive patients with ischemic stroke were admitted to the comprehensive stroke center. Of these, 26.5% (588/2216 patients) met the inclusion criteria, comprising 66.3% (390/588) of patients in the CSC-G and 33.7% (198/588) of patients in the PSC-G (Fig 1). The mean age of the study cohort was 71.7 (SD, 15.0) years with range of 22–101 years. The cohort comprised 46.4% women, 57.7% white, 17.0% black, and 12.4% Asian patients. The Online Supplemental Data provide the demographic and clinical characteristics of the study population.

The cohort was stratified into patients with and without rescan time delays (≥10 minutes), of which 27.9% (164/588) were classified with a rescan time delay with a mean of 53.7 (SD, 43.4) minutes and range of 10–224 minutes. Comparison of patients with and without rescan time delays (Online Supplemental Data) revealed statistically significant differences in race, arrival year, arrival time and mode, imaging location, baseline function, and stroke severity. Specifically, the rescan time delay group had a significantly higher proportion of black patients, arrival at off-hours, arrival by ambulance, PSC-G imaging location, prestroke independent ambulation (mRS ≤3), and severe stroke presentation (NIHSS ≥10). Insurance status and median household income were not significantly different among groups. Acute stroke treatment, particularly with intravenous thrombolysis at the comprehensive stroke center and thrombectomy, was significantly more frequent in the rescan time delay group. Patients with rescan time delays were significantly more likely to have longer hospital admissions (length of hospital stay, ≥8 days), while discharge home was significantly more likely in those without rescan time delays (46.5% versus 26.8%, P < .001). Finally, the frequency of

rescan delays significantly decreased in 2019–2020 compared with 2017–2018 (16.9% [49/290 patients] versus 38.6% [115/298 patients], P < .001).

The mean overall time difference calculated between NCCT and CTA ± CTP was significantly higher in PSC-G compared with CSC-G (n = 588; 40.4 [SD, 46.4] minutes versus 3.9 [SD, 9.0] minutes, P < .001). Similarly, in the subset of patients with rescan time delays (n = 164), the mean delay duration was significantly higher in PSC-G compared with CSC-G (65.4 [SD, 45.4] minutes versus 23.6 [SD, 14.0] minutes, P < .001). Secondary analyses of the alternate scenarios applying uniform inclusion criteria of 1- and 6-hour time limits for both CSC-G and PSC-G showed findings consistent with the primary analysis (Online Supplemental Data). Figure 2 plots the distribution of overall time differences between NCCT and CTA ± CTP, comparing PSC-G and CSC-G in the primary and secondary inclusion scenarios.

Several factors were found to be independently associated with rescan time delay in the multivariable logistic regression analysis (Table). Acute stroke imaging at a primary stroke center was the most significant factor associated with rescan time delay (OR = 9.40; 95% CI, 5.76–15.34; P < .001). Other factors associated with increased odds of rescan time delay included intravenous thrombolytic administration (OR = 2.27; 95% CI, 1.33–3.87; P = .003), black race (OR = 2.34; 95% CI, 1.31–4.17; P = .004), baseline independent ambulation (OR = 4.34; 95% CI, 1.77–10.66; P = .001), admission NIHSS ≥10 (OR = 1.72; 95% CI, 1.01–2.92; P = .045), and onset-to-arrival time of >6 hours (OR = 1.98; 95% CI, 1.14–3.43; P = .015). However, admission in 2019–2020 was associated with significantly decreased odds of rescan time delay (OR = 0.34; 95% CI, 0.21–0.55; P < .001).

When we controlled for factors associated with outcome, such as advanced age, baseline NIHSS, and treatment, additional analyses revealed that rescan time delays were significantly associated with decreased odds of home discharge (OR = 0.53; 95% CI, 0.30–0.95; P = .032). Furthermore, the odds of home discharge were significantly increased by association with intravenous thrombolysis and were decreased by association with an admission NIHSS score of ≥10 and hospital arrival by ambulance. More details on this logistic regression analysis may be seen in
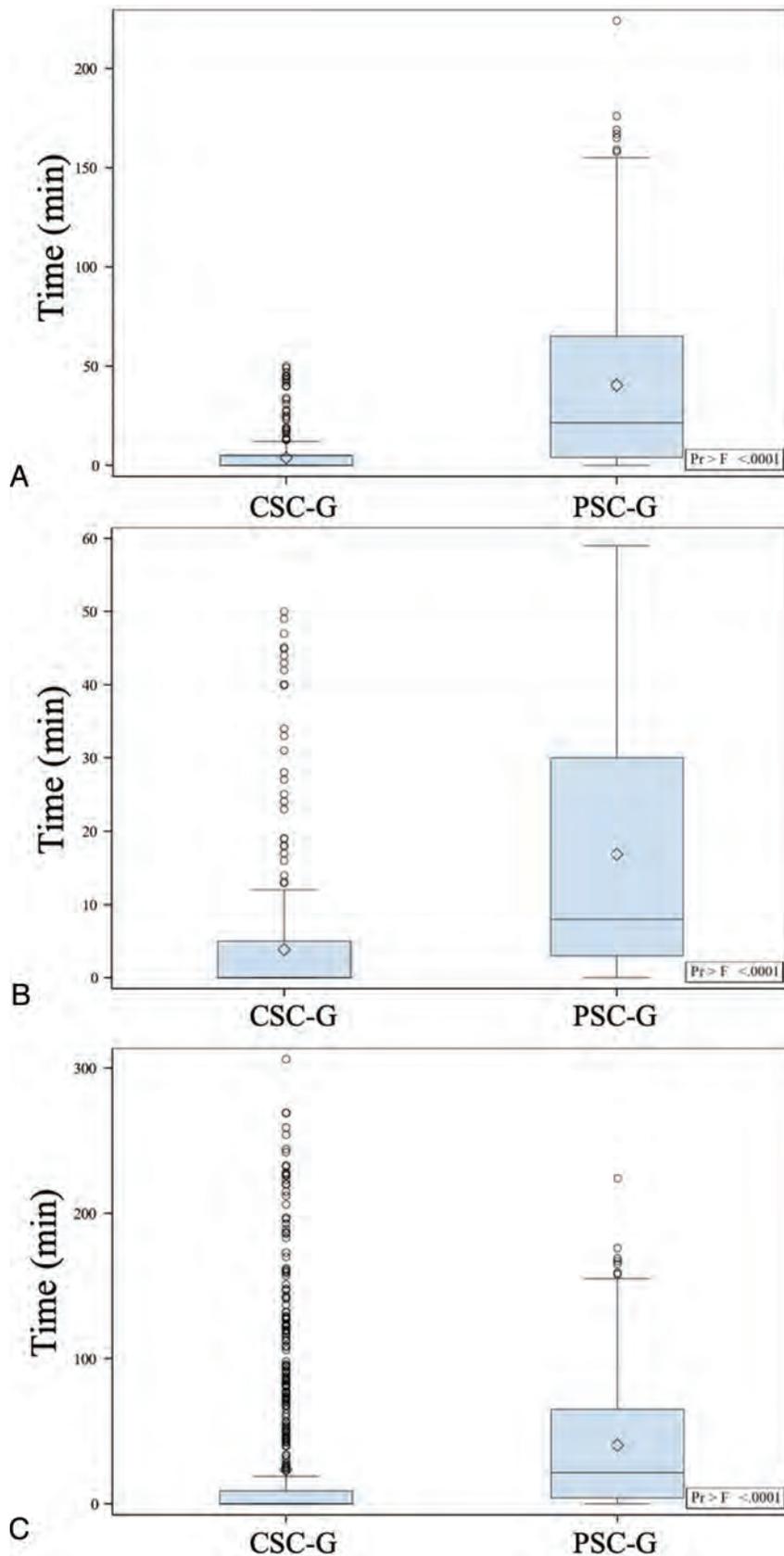
FIG 2. Distribution of overall time differences between NCCT and CTA ± CTP scans stratified by CSC-G and PSC-G. A, Primary cohort. B, One-hour imaging window scenario. C, Six-hour imaging window scenario.

the Table. Figure 3 shows the shift of discharge mRS scores and disposition by rescan time-delay intervals. Rescan time delays of <20 minutes were significantly associated with an increased proportion of patients who were functionally independent (mRS 0–2; $P = .030$) at discharge. Furthermore, no rescan time delay was significantly associated with more patients discharged home ($P = .002$).

After analyzing the alternate inclusion scenarios, we found similar findings with only a few differences for the rescan time delay and early outcome regression models compared with the primary analyses. These differences are presented in the Online Supplemental Data.

DISCUSSION

To our knowledge, this is the first study to quantify time delays associated with rescanning for advanced CT-based imaging following an initial NCCT in patients with acute ischemic stroke undergoing evaluation for acute treatment. Because a sequential approach to early CT-based imaging is recommended in current guidelines for patients being considered for thrombectomy, understanding the time delays associated with this approach is important because successful reperfusion and good clinical outcomes are time-dependent. While almost three-quarters of our acute ischemic stroke cohort underwent NCCT and CTA ± CTP simultaneously, not doing so was associated with an average of 53.7 (SD, 43.4) minutes of time delay between imaging. Rescanning was significantly more common and the mean delay 2.8 times longer at primary stroke centers compared with the comprehensive stroke center, a finding that remained significant after controlling for other demographic and clinical factors. Even the lower mean time-delay duration observed in CSC-G (23.6 minutes) accounts for 26% of the 90-minute door-to-device goal for direct arrivals in the Target: Stroke Phase III initiative.

Nonetheless, the frequency of rescanning significantly decreased during the study period, owing to the
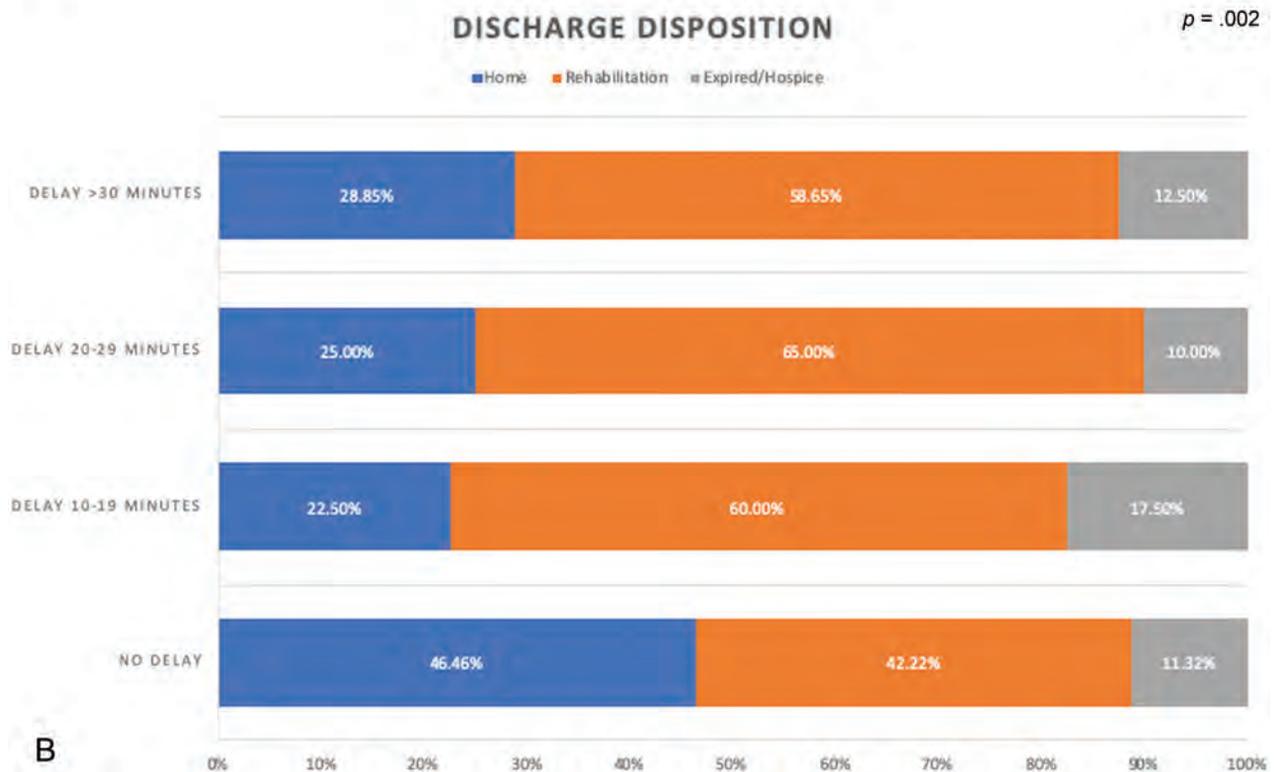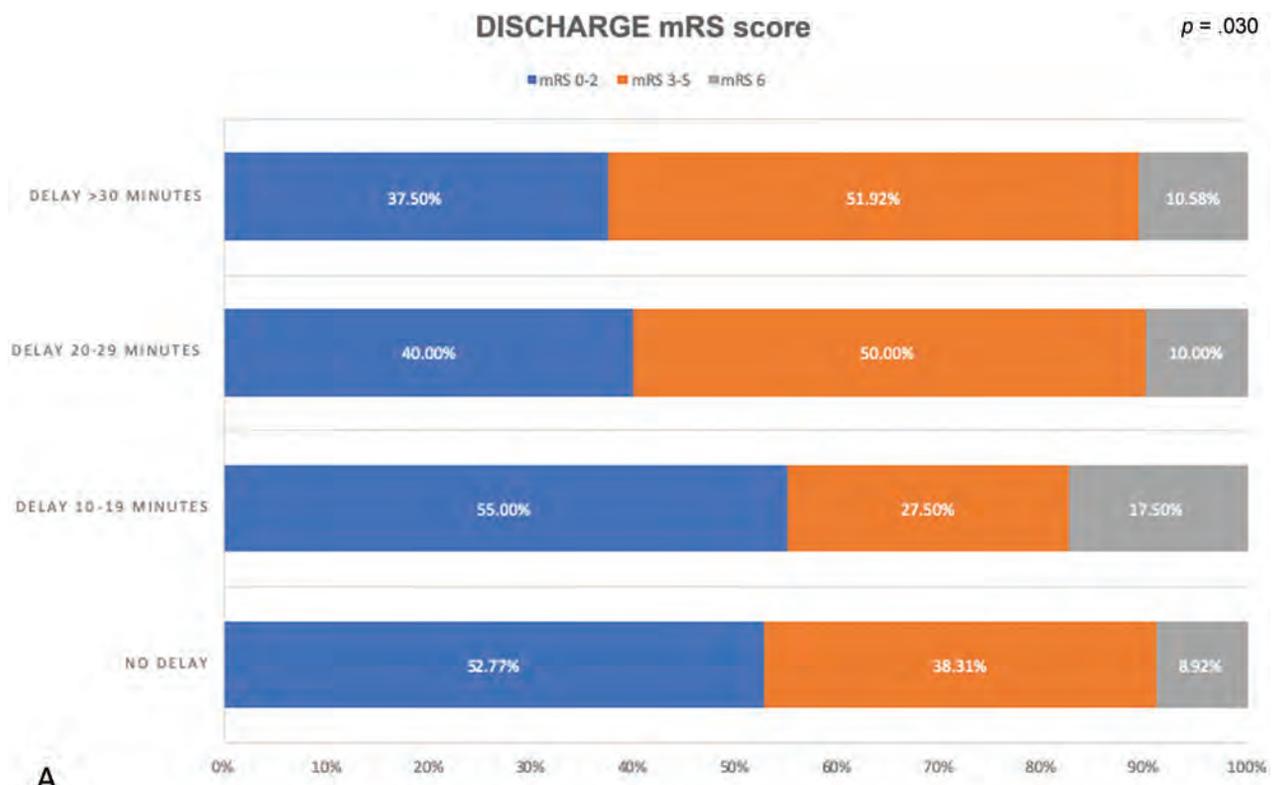
FIG 3. Distribution of mRS score (*A*) and discharge disposition (*B*) by time-delay intervals.

adoption of a new workflow at the comprehensive stroke center for early CTAP in all patients with stroke presenting in the 24-hour treatment window and protocolized early CTA use at many primary stroke centers. To improve efficiencies even further, some have recently advocated a direct-to-angiography suite workflow, bypassing the CT scanner in an emergency department entirely,

with CTA ± CTP performed in the interventional suite using conebeam CT technology.[31-33] Two observational studies[31,32] have shown improved time to treatment with this novel paradigm; however, improved functional outcomes, or even workflow efficiencies, were not proved in a randomized trial.[34]

As might be expected, patients administered intravenous thrombolytics or presenting >6 hours from last known well were significantly more likely to have time delays to CTA ± CTP. In the multivariable logistic regression analysis, endovascular treatment was not associated with time delays after controlling for the other factors listed in the Table. The observed independent association between intravenous thrombolytic administration and time delay was likely a result of workflow when thrombolytics were administered immediately post-NCCT, but before CTA ± CTP, especially when treatment was delivered outside the CT suite. Less-expected findings included racial disparity of increased time delays in black patients, particularly in the PSC-G, and increased time delays in patients with higher admission NIHSS scores and those ambulatory at baseline. The time delays in black patients were consistent across both secondary analyses using identical inclusion windows for the CSC-G and PSC-G and even when controlling for median household income. Delays in stroke care experienced by black patients in emergency departments have been reported by others.[35] Bias in care delivery, as well as delayed stroke-symptom recognition and hospital arrival, and less use of ambulance services by black patients may be contributory.[28] The relation to higher admission NIHSS scores may result from causes of delay that we could explain, such as blood pressure instability, airway management, or treatment for agitation that may occur in patients with more severe deficits.

Delay between NCCT and CTA ± CTP was associated with a near 50% reduction in the likelihood of home discharge, and scan time differences of <20 minutes were associated with a higher proportion of patients who were functionally independent at discharge. The magnitude of time delay found in our study would be expected to significantly impact outcome following successful reperfusion. For instance, based on data from the Interventional Management of Stroke (IMS) III Trial, the mean delay we found translated to a 21% reduction in the relative likelihood of 90-day functional independence.[10] Furthermore, according to data from SWIFT-STAR trials[9] and HERMES meta-analysis,[8] our finding would yield 12%–14% additional patients having more disabled outcomes post-successful reperfusion and 9% fewer patients achieving 90-day functional independence.[8] Similarly, per the Multicenter Randomized Clinical Trial of Endovascular Treatment of Acute Ischemic Stroke in the Netherlands (MR CLEAN) registry, delay associated with rescanning would reduce the probability of 90-day functional independence by 7%.[11]

Our study has several strengths and limitations. We aimed to focus our inclusion criteria on patients who were most likely undergoing evaluation for endovascular treatment and not purely for secondary stroke-prevention evaluation; this feature is a strength of this study because delayed imaging is not a main concern to patients outside endovascular treatment considerations. Nonetheless, this goal necessitated different inclusion time windows for patient selection in the PSC-G and CSC-G, which could have introduced selection bias into our results. We attempted to mitigate this limitation by re-analyzing our data using uniform criteria applied to both the PSC-G and CSC-G. These alternate inclusion scenarios (1- and 6-hour time limits) showed very similar patterns of results compared with the primary analysis, though the magnitude of the time delay was certainly influenced by our inclusion time limits. Another strength is our inclusion and analysis of data from both primary and comprehensive stroke centers. This feature enabled us to compare time delays at different sites with varied resources and levels of stroke care. However, because our findings are from a single stroke network, broad applicability is limited and similar studies from other health care networks and geographic areas may be useful.

Our study has other important limitations. First, it has a retrospective study design; thus, we used discharge disposition and discharge mRS, instead of 90-day mRS as outcome measures because they were readily available in our database. Several studies support discharge disposition being a good predictor of a longer-term stroke outcome.[36,37] Our retrospective design also limited our ability to include imaging results in our analyses, such as the presence of a large-vessel occlusion or a perfusion defect. Second, our sample size of 588 patients is relatively small and may have resulted in not detecting statistically small differences between groups. This sample size also precluded us from drawing conclusions from important subgroup analyses, such as those who underwent endovascular thrombectomy or had large-vessel occlusion, and from stratifying by infarction vascular territory, because, for instance, patients with posterior circulation stroke may have more unstable conditions and may need airway management, resulting in increased time delays. Nonetheless, the ratio of the number of participants to the number of variables in our regression model is comparable with that in other studies.

CONCLUSIONS

Obtaining NCCT and CTA ± CTP separately was associated with significant time delays in acquiring comprehensive CT imaging information for treatment decision-making in patients with acute ischemic stroke. Time delays to CTA ± CTP were associated with poorer early clinical outcomes. Both intravenous thrombolytic administration and delayed onset-to-arrival times were independently associated with rescanning imaging time delays, while implementing protocols to obtain NCCT and CTAP simultaneously was associated with avoiding the observed time delays associated with rescanning. Considering that this imaging paradigm has been shown to be the most cost-effective imaging strategy for patients with acute ischemic stroke,[38] the current study provides further support for simultaneous NCCT and CTA ± CTP imaging to minimize time delays to treatment.

REFERENCES

1. Goyal M, Demchuk AM, Menon BK, et al. **Randomized assessment of rapid endovascular treatment of ischemic stroke.** *N Engl J Med* 2015;372:1019–30 CrossRef Medline

2. Campbell BC, Mitchell PJ, Kleinig TJ, et al. **Endovascular therapy for ischemic stroke with perfusion-imaging selection.** *N Engl J Med* 2015;372:1009–18 CrossRef Medline

3. Berkhemer OA, Fransen PSS, Beumer LA, et al. MR CLEAN Investigators. **A randomized trial of intraarterial treatment for acute ischemic stroke.** *N Engl J Med* 2015;372:11–20 CrossRef Medline

4. Jovin T, Chamorro A, Cobo E, et al. REVASCAT Trial Investigators. **Thrombectomy within 8 hours after symptom onset in ischemic stroke.** *N Engl J Med* 2015;372:2296–2306 CrossRef Medline

5. Saver JL, Goyal M, Bonafe A, et al. **Stent-retriever thrombectomy after intravenous t-PA vs. t-PA alone in stroke.** *N Engl J Med* 2015;372:2285–95 CrossRef Medline

6. Nogueira RG, Jadhav AP, Haussen DC, et al. **Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct.** *N Engl J Med* 2018;378:11–21 CrossRef Medline

7. Albers GW, Marks MP, Kemp S, et al. **Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging.** *N Engl J Med* 2018;378:708–18 CrossRef Medline

8. Saver JL, Goyal M, van der Lugt A, et al. HERMES Collaborators. **Time to treatment with endovascular thrombectomy and outcomes from ischemic stroke: a meta-analysis.** *JAMA* 2016;316:1279–88 CrossRef Medline

9. Sheth SA, Jahan R, Gralla J, et al. for the SWIFT-STAR Trialists. **Time to endovascular reperfusion and degree of disability in acute stroke.** *Ann Neurol* 2015;78:584–93 CrossRef Medline

10. Khatri P, Yeatts SD, Mazighi M, et al. IMS III Trialists. **Time to angiographic reperfusion and clinical outcome after acute ischaemic stroke: an analysis of data from the Interventional Management of Stroke (IMS III) phase 3 trial.** *Lancet Neurol* 2014;13:567–74 CrossRef Medline

11. Mulder MJH, Jansen IG, Goldhoorn RB, et al. MR CLEAN Registry Investigators. **Time to endovascular treatment and outcome in acute ischemic stroke: MR CLEAN registry results.** *Circulation* 2018;138:232–40 CrossRef Medline

12. American Heart Association. **Target: Stroke Phase III.** https://www.heart.org/en/professional/quality-improvement/target-stroke/introducing-target-stroke-phase-iii. Accessed January 5, 2021

13. Powers WJ, Rabinstein AA, Ackerson T, et al. **Guidelines for the Early Management of Patients With Acute Ischemic Stroke: 2019 Update to the 2018 Guidelines for the Early Management of Acute Ischemic Stroke—A Guideline for Healthcare Professionals from the American Heart Association/American Stroke Association.** *Stroke* 2019;50:e344–418 CrossRef Medline

14. Pierot L, Jayaraman MV, Szikora I, et al. Asian-Australian Federation of Interventional and Therapeutic Neuroradiology (AAFITN), Australian and New Zealand Society of Neuroradiology (ANZSNR), American Society of Neuroradiology (ASNR), Canadian Society of Neuroradiology (CSNR), European Society of Minimally Invasive Neurological Therapy (ESMINT), European Society of Neuroradiology (ESNR), European Stroke Organization (ESO), Japanese Society for NeuroEndovascular Therapy (JSNET), The French Society of Neuroradiology (SFNR), Ibero-Latin American Society of Diagnostic and Therapeutic Neuroradiology (SILAN), Society of NeuroInterventional Surgery (SNIS), Society of Vascular and Interventional Neurology (SVIN), World Stroke Organization (WSO), World Federation of Interventional Neuroradiology (WFITN). **Standards of practice in acute ischemic stroke intervention: international recommendations.** *J Neurointerv Surg* 2018;10:1121–26 CrossRef Medline

15. Wintermark M, Sanelli PC, Albers GW, et al. Society of NeuroInterventional Surgery. **Imaging recommendations for acute stroke and transient ischemic attack patients: a joint statement by the American Society of Neuroradiology, the American College of Radiology and the Society of NeuroInterventional Surgery.** *J Am Coll Radiology* 2013;10:828–32 CrossRef Medline

16. Kwei KT, Liang J, Wilson N, et al. **Stroke severity affects timing: time from stroke code activation to initial imaging is longer in patients with milder strokes.** *Neurologist* 2018;23:79–82 CrossRef Medline

17. Mowla A, Doyle J, Lail NS, et al. **Delays in door-to-needle time for acute ischemic stroke in the emergency department: a comprehensive stroke center experience.** *J Neurol Sci* 2017;376:102–05 CrossRef Medline

18. Prabhakaran S, Khorzad R, Brown A, et al. **Academic-community hospital comparison of vulnerabilities in door-to-needle process for acute ischemic stroke.** *Circ Cardiovasc Qual Outcomes* 2015;8:S148–54 CrossRef Medline

19. Koge J, Matsumoto S, Nakahara I, et al. **Improving treatment times for patients with in-hospital stroke using a standardized protocol.** *J Neurol Sci* 2017;381:68–73 CrossRef Medline

20. Campbell BC, Christensen S, Levi CR, et al. **Comparison of computed tomography perfusion and magnetic resonance imaging perfusion-diffusion mismatch in ischemic stroke.** *Stroke* 2012;43:2648–53 CrossRef Medline

21. Eastwood JD, Lev MH, Wintermark M, et al. **Correlation of early dynamic CT perfusion imaging with whole-brain MR diffusion and perfusion imaging in acute hemispheric stroke.** *AJNR Am J Neuroradiol* 2003;24:1869–75 Medline

22. Saur D, Kucinski T, Grzyska U, et al. **Sensitivity and interrater agreement of CT and diffusion-weighted MR imaging in hyperacute stroke.** *AJNR Am J Neuroradiol* 2003;24:878–85 Medline

23. Urbach H, Flacke S, Keller E, et al. **Detectability and detection rate of acute cerebral hemisphere infarcts on CT and diffusion-weighted MRI.** *Neuroradiol* 2000;42:722–27 CrossRef Medline

24. United States Census Bureau. **B19013 median household income in the past 12 months (in 2019 inflation-adjusted dollars).** https://censusreporter.org/tables/B19013/. Accessed September 15, 2020

25. Saposnik G, Cote R, Phillips S, et al. **Stroke outcome in those over 80: a multicenter cohort study across Canada.** *Stroke* 2008;39:2310–17 CrossRef Medline

26. Arora R, Salamon E, Katz JM, et al. **Use and outcomes of intravenous thrombolysis for acute ischemic stroke in patients ≥90 years of age**. *Stroke* 2016;47:2347–54 CrossRef Medline

27. **NIH policy on reporting race and ethnicity data: subjects in clinical research. Release Date. August 8, 2001.** https://grants.nih.gov/grants/guide/notice-files/not-od-01-053.html. Accessed April 22, 2021

28. Cruz-Flores S, Rabinstein A, Biller J, et al. Council on Quality of Care and Outcomes Research. **Racial-ethnic disparities in stroke care: the American experience—a statement for healthcare professionals from the American Heart Association/American Stroke Association.** *Stroke* 2011;42:2091–116 CrossRef Medline

29. Goyal M, Almekhlafi MA, Fan L, et al. **Evaluation of interval times from onset to reperfusion in patients undergoing endovascular therapy in the IMS III Trial.** *Circulation* 2014;130:265–72 CrossRef Medline

30. Campbell BC, Weir L, Desmond PM, et al. **CT perfusion improves diagnostic accuracy and confidence in acute ischaemic stroke.** *J Neurol Neurosurg Psychiatry* 2013;84:613–18 CrossRef Medline

31. Ribó M, Boned S, Rubiera M, et al. **Direct transfer to angiosuite to reduce door-to-puncture time in thrombectomy for acute stroke.** *J Neurointerv Surg* 2018;10:221–24 CrossRef Medline

32. Mendez B, Requena M, Aires A, et al. **Direct transfer to angio-suite to reduce workflow times and increase favorable clinical outcome.** *Stroke* 2018;49:2723–27 CrossRef Medline

33. Brehm A, Tsogkas I, Maier IL, et al. **One-stop management with perfusion for transfer patients with stroke due to a large-vessel occlusion: feasibility and effects on in-hospital times.** *AJNR Am J Neuroradiol* 2019;40:1330–34 CrossRef Medline

34. Pfaff JA, Schönenberger S, Herweh C, et al. **Direct transfer to angio-suite versus computed tomography-transit in patients receiving mechanical thrombectomy: a randomized trial.** *Stroke* 2020;51:2630–38 CrossRef Medline

35. Karve SJ, Balkrishnan R, Mohammad YM, et al. **Racial/ethnic disparities in emergency department waiting time for stroke patients in the United States.** *J Stroke Cerebrovasc Dis* 2011;20:30–40 CrossRef Medline

36. Belagaje SR, Sun CH, Nogueira RG, et al. **Discharge disposition to skilled nursing facility after endovascular reperfusion therapy predicts a poor prognosis.** *J Neurointerv Surg* 2015;7:99–103 CrossRef Medline

37. Bautista AE, Meyer DM, Meyer BC. **Novel definition of stroke "good responders" predicts 90-day outcome after thrombolysis.** *J Stroke Cerebrovasc Dis* 2019;28:104422 CrossRef Medline

38. Martinez G, Katz JM, Pandya A, et al. **Cost-effectiveness study of initial imaging selection in acute ischemic stroke care.** *J Am Coll Radiol* 202031362–64 CrossRef Medline

Contrast Bolus Interference in a Multimodal CT Stroke Protocol

E. Kellner, A. Rau, T. Demerath, M. Reisert, and H. Urbach

ABSTRACT

BACKGROUND AND PURPOSE: Whether CTP is performed before or after CTA varies within multimodal CT stroke protocols. CTA after CTP might show venous filling, and CTP metrics might be disturbed by prior CTA. Therefore, we compared CTP metrics conducted before and after CTA in a large cohort of patients with stroke and analyzed interferences of the CTA bolus with the CTP measurement.

MATERIALS AND METHODS: We analyzed 1980 patients (368 patients with CTP performed before CTA [group A] versus 1612 patients with CTP performed after [group B]) in a retrospective study. Mean curves, histograms of CTP baseline Hounsfield units, CBF, CBV, time-to-maximum, hypoperfusion, and core volumes were calculated using the software VEOcore. CTA and CTP interferences were analyzed, and a detection and correction method was proposed.

RESULTS: Mean CTP baseline values were significantly different in both groups (41 versus 45 HU within the groups A and B, respectively). However, perfusion metrics, hypoperfusion, and core volumes yielded no significant differences. In 49 patients, the descending flank of the CTA bolus interfered with the baseline of the CTP measurement, leading to erroneously low CBV values. These errors vanished when a correction method was applied.

CONCLUSIONS: CTP can be reliably performed after CTA without a relevant net effect on perfusion metrics. However, when measuring CTP after CTA, either a short pause on the order of 30 seconds should be observed or an appropriate correction method should be applied. It may help to avoid excluding patients from mechanical thrombectomy by overestimating infarct cores.

ABBREVIATION: Tmax = time-to-maximum

Multimodal CT stroke imaging protocols comprise NCCT followed by CTA and, at least beyond the 6-hour window, CTP.[1-4] Technically, CTP can be performed either before or after CTA. When CTP is performed first, processing or postprocessing delays, CTA acquisition, and CTA image quality might be disturbed by venous contrast.[5,6] When CTA is performed first, CTP time density curves may be altered.[7]

From a theoretic point of view, CTA before CTP should not alter the perfusion metrics as long as a sufficiently long time has passed between the measurements and the CTA bolus has completely dissolved to a well-mixed and constant state in the vascular system, because constant enhancement cancels out during perfusion calculation, as described below.

To optimize our multimodal CT stroke protocol, we compared CTP before CTA and vice versa and performed a systematic comparison in a large collection of 1980 patients. We analyzed perfusion metrics of CTP performed either before or after CTA and bolus interferences between CTA and CTP. We finally proposed a correction algorithm for cases in which an overlap of the descending CTA bolus with the CTP acquisition (CTA flank effect) had relevant influence on perfusion metrics.

MATERIALS AND METHODS
Patient Cohort
In a retrospective analysis, a PACS query was conducted for patients having undergone both CTA and CTP within the same
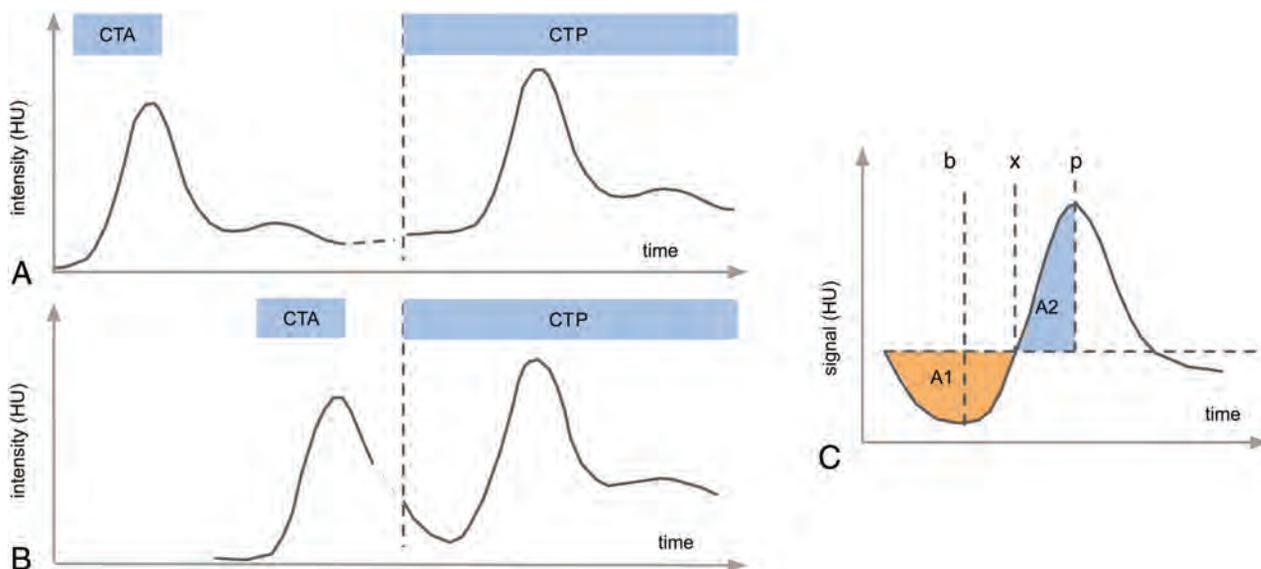
FIG 1. Measurement scheme of CTA performed before CTP. *A*, If there is a sufficiently long pause between the measurements (>30 seconds), the contrast bolus from the CTA measurement is well-mixed in the vascular system and causes a constant elevation of the CTP baseline on the order of some Hounsfield units. *B*, If the time gap between CTP and CTA is very short, the trailing flank of the CTA bolus might interfere with the CTP baseline, leading to errors in perfusion calculation. *C*, We suggest the following simple-but-robust method for detection and correction: The outstanding feature of the effect is the presence of a negative lobe at the curve start before the actual bolus. This can be measured by comparing the areas under the curve A1 and A2. The thus required relevant points can be robustly determined as the following: p as the position of the curve maximum, and the turning point x as the minimum of the integral of the integral/stem function (the area grows larger after this point). If A1/A2 exceeds a certain threshold, the artifact is considered significant and is corrected by dropping all points before the minimum b of the baseline.

examination between January 2017 and October 2020. For data management and image processing, all images were exported and anonymized to a local instance of the imaging platform Nora (www.nora-imaging.org). Information on patients' biologic sex, age, study date, and time interval between CTA and CTP were extracted from the DICOM headers. On the basis of the timing information, cases were assigned to the groups with CTP conducted before (group A) and after CTA (group B). The study was approved by the local ethics committee, and informed written consent was waived.

Quality Control and Visual Rating

All raw perfusion series and processed results as described below were checked by 2 neuroradiologists (T.D. and A.R.) to identify cases with obvious technical issues (severe motion, erroneous contrast agent injection, and so forth). Furthermore, the presence of interferences between time-density curves of CTA and CTP as illustrated in Fig 1 was assessed during this process.

The overall hemodynamics might be influenced by factors such as extracranial carotid stenosis or occlusion. We rated the presence of a high-grade stenosis or occlusion (>70% according to NASCET criteria) from DSA available in 387 cases, to examine equal balance among groups.

Multimodal CT Protocol

CT scans were performed on a 128–detector row (Somatom Definition Flash; Siemens) or a 64–detector row multidetector CT scanner (Somatom Definition 64 AS; Siemens). NCCTs were acquired in the spiral mode covering the skull base to the vertex.

The parameters for Somatom Definition Flash were the following: reference tube voltage = 120 kV, reference tube current = 300 mAs, CARE Dose4D (Siemens), collimation = 20 × 0.6 mm ("alternating spot"), rotation speed = 1.0 seconds, resolution = 0.75 × 0.75 × 3 mm, gantry tilt if needed. The parameters for Somatom Definition 64 AS were the following: reference tube voltage = 120 kV, reference tube current = 280 mAs, CARE Dose4D, collimation = 64 × 0.6 mm (alternating spot), rotation speed = 0.3 seconds, resolution = 0.75 × 0.75 × 3 mm, gantry tilt if needed.

CTP series were acquired in the axial scan mode with the following protocols for the Somatom Definition Flash: 80 kV, 180 mAs, collimation = 16 × 1.2 mm, no gantry tilt, tube rotation time = 0.3 seconds, toggle technique (5 blocks × 0.3 = 1.5 seconds), z-coverage = 100 mm, section thickness = 5 mm, increment = 5 mm, 27 series every 1.5 seconds = 42-second scan time (effective temporal resolution in the middle of the plane = 1.5 seconds, in the periphery = 2.6 seconds) after intravenous (16–18 G) injection of 40 mL of Imeron 400 (iopamidol; Bracco) + 30 mL of sodium chloride (NaCl) at a flow rate of 6 mL/s.

The parameters for Somatom Definition 64 AS were the following: 80 kV, 180 mAs, collimation = 32 × 1.2 mm, no gantry tilt, tube rotation time = 0.3 seconds, toggle technique (5 blocks × 0.3 = 1.5 seconds), z-coverage = 90 mm, section thickness = 5 mm, increment = 3 mm, 30 series every 1.5 seconds = 45-second scan time (effective temporal resolution in the middle of the plane = 1.5 seconds, in the periphery = 2.6 seconds) after intravenous (16–18 G) injection of 40 mL of Imeron 400 + 30 mL of NaCl at a flow rate of 6 mL/s.

CTA scans were obtained in the spiral mode covering the aortic arch to the vertex, with bolus tracking in the ascending aorta (threshold, 100 HU): The Somatom Definition Flash parameters were the following: reference tube voltage = 100 kV, reference tube current = 155 mAs, CARE Dose4D, collimation = 128 × 0.6 mm, tube rotation time = 0.28 seconds, no gantry tilt, after a cubital (16–18 G) injection of 70 mL of contrast agent (Imeron 400) + 60 mL of NaCl at a flow rate of 5 mL/s.

Somatom Definition 64 AS parameters were the following: reference tube voltage = 120 kV, reference tube current = 110 mAs, CARE Dose4D, collimation = 64 × 0.6 mm, tube rotation time = 0.3 seconds, no gantry tilt, after a cubital (16–18 G) injection of 70 mL of contrast agent (Imeron) + 60 mL of NaCl at a flow rate of 5 mL/s.

Comparison Strategy

We compared group A (CTP before CTA), group B (CTA before CTP), and the subgroup with bolus interference, focusing on multiple technical and clinical levels and metrics: a) Hounsfield unit–time-density and concentration curves, b) histograms of baseline images, c) perfusion parameter maps CBF, CBV, and time-to-maximum (Tmax), and finally, d) hypoperfusion and core volumes.

Analogous to best practices in quality-management systems, this multilevel approach can be regarded as a combination of unit tests and full-system tests: Items (a–c) address several technical levels of the processing pipeline, and item (d) covers potential clinical implications. Unit tests should be performed under well-defined and consistent conditions, so for technical items (a–c), statistical analysis has been restricted to the healthy, nonstroke brain hemisphere. Item (d) finally covers the clinically relevant stroke side. Group comparisons were performed by visual assessment of histograms, median values, and statistical tests.

Data Analysis

Perfusion scans were processed using the software package VEOcore (VEObrain; www.veobrain.com) and customized scripts for detailed data analysis, consisting of the following steps:

Motion Correction. The perfusion time-series were corrected by rigid-body realignment of all scan volumes to the first volume using the cross-correlation between volumes as the objective function.

Baseline Calculation. The baseline is defined as the set of volumes before bolus arrival. Bolus arrival was determined as the time point before the Hounsfield unit value exceeded 10% of the bolus peak Hounsfield unit value for the first time (typically 3–7 time points). An average baseline image was calculated as the mean during this time range before bolus arrival.

Coregistration of CTP and NCCT. NCCT was aligned to the CTP series by a rigid-body coregistration using the mutual information of the images as the objective function.

Brain Masking. A brain mask was created on the basis of the NCCT image as the Hounsfield unit range (20–100). Morphologic operations (image dilation/erosion, and connected component removal of speckle noise) were applied to clean the masks.

Arterial/Venous Masking. Arterial/venous masking has not been performed.

Contrast Agent Concentration Calculation. Relative contrast agent concentration curves were calculated by subtracting the baseline volume (assuming a linear relationship between signal attenuation and contrast agent concentration,[8] see also "CTA Flank Effect" below).

Arterial Input Function Selection. The arterial input function was automatically selected using a cluster analysis approach, dividing all curves into arterial and venous pools and selecting a collection of curves with the earliest bolus arrival of the arterial pool.

Deconvolution. Deconvolution of tissue curves and the arterial input function was performed by directly solving the inverse model with additional L1 Tikhonov regularization (the strength of the regularization parameter was 5). CBF and Tmax were calculated as the maximum and argmax of the thus obtained residue function. CBV was calculated by numeric integration of the concentration curves.

Selection of the Healthy Hemisphere. The midline of the brain was automatically determined by coregistration of the baseline image with its mirror image. The healthy hemisphere was selected as the one with smaller median Tmax value.

Normalization to Healthy Hemisphere. CBV and CBF were normalized to the median value of the healthy hemisphere = 100%.

Infarction Volumetry. Hypoperfusion and core volumes were calculated on the basis of the generally accepted Endovascular Therapy Following Imaging Evaluation for Ischemic Stroke 3 (DEFUSE-3) thresholds (Tmax > 6 seconds for hypoperfusion and CBF <30% for the infarction core).[2]

Time-Curve Alignment for Group Comparisons. The Hounsfield unit density and concentration time curves can differ significantly between patients due to contrast injection timing, cardiac output, and other factors. To make them comparable to some degree, we shifted the curves for each individual patient and resampled them so that the bolus maximum was commonly centered in an acquisition time window between 0 and 70 seconds.

Calculation of Histograms. Histograms of CBF, CBV, Tmax, and baseline images were calculated for the healthy hemisphere for each patient. We chose to select the healthy hemisphere because it allows testing for effects on the basis of well-defined and consistent conditions for each patient.

Statistical Analysis

Time curves, histograms of baseline image, CBF, CBV, and Tmax were averaged within the groups for visual comparison. The distribution of median values of the histograms serves as a benchmark. For comparison of infarction volumes, we wanted to address the clinical relevance; therefore, only cases with

Table 1: Statistics for perfusion metrics for all groups[a]

	CTP before CTA	CTA before CTP	CTP with Flank Effect	CTP with Corrected Flank Effect
Group	A	B	B flank	B corrected flank
No. of cases	368	1563	49	49
Age (yr)	78.0 (18.5)	76.0 (19.0)	83.0 (10.2)	83.0 (10.2)
	$P = .06$		$P < .001$	$P < .001$
Time gap (sec)	91.5 (89.5)	62.0 (67.0)	18.0 (4.0)	18.0 (4.0)
	$P < .001$		$P < .001$	$P < .001$
Baseline (HU)	41.0 (3.0)	45.0 (4.0)	48.0 (6.2)	45.0 (5.0)
	$P < .001$		$P < .001$	$P = .71$
Tmax (sec)	1.7 (0.6)	1.7 (0.5)	2.1 (1.2)	2.0 (0.9)
	$P = .054$		$P < .001$	$P < .001$
Core volume (mL)	18.5 (18.0)	18.0 (16.0)	26.0 (38.5)	23.0 (31.0)
	$P = .43$		$P = .027$	$P = .078$
Hypoperfusion volume (mL)	44.5 (96.0)	41.0 (76.0)	67.0 (81.8)	53.0 (85.0)
	$P = .2$		$P < .001$	$P < .001$

[a] Median values, interquartile range (parentheses), and *P* values from a Wilcoxon rank sum test are reported.

hypoperfusion values of >3 mL have been included in the comparison because cases with lower values are typically just segmentation noise and might introduce a bias. A 2-sided Wilcoxon rank sum test was performed to accept or reject the null hypothesis that the groups had the same median values. Test results with *P* values < .05 were considered statistically significant.

The CTA Flank Effect and Proposal of a Correction Method

Iodinated contrast agent results in increased x-ray absorption by an approximately linear relationship to its concentration.[8] The image intensity curve of a bolus-tracking measurement in a single voxel can be described by $I(t) = b + f \times c(t)$, where t denotes the dynamic scan time, b denotes the image intensity before contrast (baseline), f denotes a constant absorption factor and $c(t)$ denotes the contrast agent concentration at a given time. From this, $c(t)$ can be calculated as $c(t) = [I(t) - b] / f$. The baseline b is typically averaged over the images before arrival of the contrast bolus to improve robustness against noise.

From a theoretic point of view, a "contamination" of the vasculature and parenchyma by contrast agent from a previous injection, for example from a preceding CTA measurement, should not alter this relationship, as long as the concentration is constant across time because it then is absorbed into the constant b. This can be assumed after a sufficiently long interval on the order of 30–60 seconds between CTA and CTP. After such a time interval the CTA bolus is in a well-mixed, constant state in the vascular system (Fig 1).

However, if CTP is performed very shortly after CTA, the trailing flank of the first-pass CTA bolus may extend into the baseline of the CTP measurement in a nonconstant fashion, leading to imprecise results when using the standard perfusion calculation procedures.

A detailed and precise correction approach might involve voxelwise modeling of the curves as a superposition of the CTA and CTP boluses. However, this comes with a number of challenges, especially due to the low signal-to-noise ratios of the voxelwise curves and would involve many parameters.

Therefore, as a pragmatic approach, we propose the following: First, the presence of the artifact should be detected. Only if it exceeds a certain threshold should the curves be corrected. The correction will be performed by removing the first n time points, which are biased by the descending CTA flank. The scheme is illustrated in Fig 1*C*. The most prominent feature of the artifact is the presence of a negative lobe before the actual bolus peak. This feature can be detected by measuring and comparing the areas under the curve A1 and A2. These areas can be calculated from the mean brain curve over all voxels via the landmarks x and p, where p denotes the maximum of the curve, and *x*, the turning point where the curve integral/stem function starts to increase, ie, has a curve minimum. If a significant negative lobe is detected (ie, A1/A2 is above a certain threshold), all timepoints before point b, the minimum of the curve, will be dropped, and the value of b will be taken as the only point for baseline calculation, instead of the mean over a certain range. This result comes at the cost of a reduced SNR due to missing averaging of the baseline, but it is reproducible and robust.

Selection of the Optimal Threshold for the Correction Method

For determination of the optimal threshold for artifact detection based on the ratio A1/A2, we performed a receiver operating characteristic analysis on A1/A2 versus the manual reading of the flank effect, and we selected a threshold of A1/A2 = 0.25 as the optimal threshold with an equal number of false-positives ($n = 5$) and false-negatives ($n = 5$) in the 49 cases manually rated as the ground truth.

Analysis Software

All scripts and analysis have been performed using the Matlab R2018b software package (MathWorks).

RESULTS

We included 1980 patients in the final analysis. Numbers and histograms of the study population and groups are summarized in Table 1 and Fig 2. CTP was performed before CTA in 368 patients (group A) and after CTA in 1563 patients (group B), and bolus interference occurred in 49 cases (group B1). Age and sex were equally distributed between groups A and B; group B1 had a slight trend toward more elderly patients. Group B1 also showed a noticeably shorter time interval between the end of the CTA and start of the CTP scan (18 versus 62 seconds).

With respect to our comparison strategy, we obtained the following results (Table 1 and Figs 3–5):

1) Hounsfield unit–time-density and concentration curves: The baseline density values were significantly higher when CTP was performed after CTA (4 HU on average, Table 1 and Fig 3).
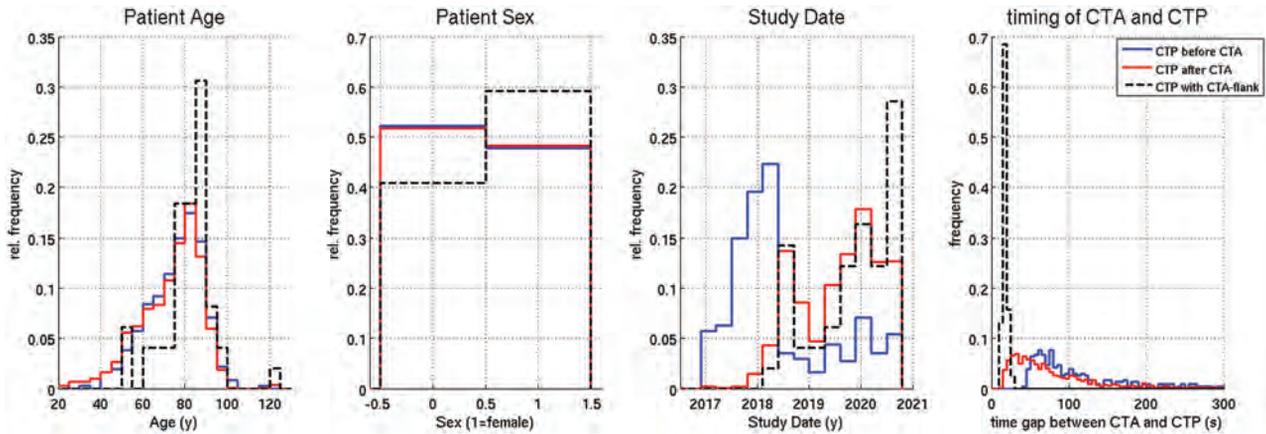
FIG 2. Study population statistics. Distribution of age, sex, and study date. There is no significant difference in sex and age between the groups CTP before versus after CTA. The group with a CTA flank effect comprised more elderly women. Furthermore, the time interval between the end of CTA and the start of CTP was significantly shorter. Histograms have been normalized to common areas. rel indicates relative.
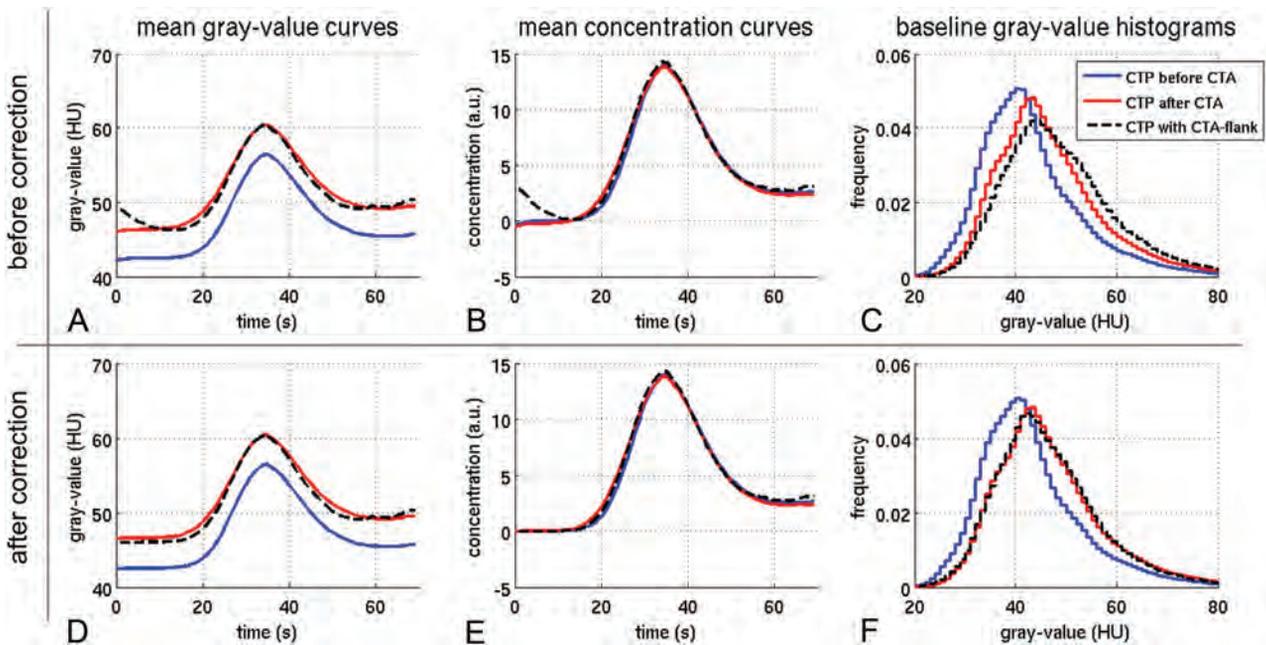
FIG 3. When CTP is performed after CTA, residual contrast agent from CTA causes a global Hounsfield unit shift in CTP (blue versus red). This can be seen in the time curves (*A*) and the gray-value histograms of the baseline image (*C*). When one converts the curves to contrast agent concentrations, the baseline is subtracted, and the effect cancels out (*B*). However, this is not the case when the falling CTA bolus extends into the CTP measurement (*A–C, black dotted lines*). The curve has a different shape, and the baseline histogram is shifted to even higher values. *D–F,* There is no significant effect in overall CBF, CBF, and Tmax for CTP before or after CTA (blue versus red). Only when the falling CTA bolus extends into CTP can a significant underestimation of CBV occur (*D–F, black dotted line*). a.u. indicates arbitrary unit.

This difference vanished in the concentration curves due to the baseline subtraction involved in this process. The subgroup with flank effect showed a relevant bolus interference at the beginning of the measurement. This interference could be corrected using the proposed correction algorithm (Fig 3*D*, -*F*).

2) Histograms of baseline images: The difference in the baseline can also be observed in the gray-value histograms in Fig 3*C*.

3) Perfusion parameter maps: CBF, CBV, and Tmax median values and histograms did not significantly differ between groups A and B, as indicated in Table 1 and Fig 4. For the group with the flank effect, different CBV and Tmax histograms were observed.

This difference vanished after application of the correction algorithm (Fig 4*B*). The correction also improved the visual appearance of the concentration curves and CBV maps, as exemplary shown in Fig 5 for 1 case.

4) Hypoperfusion and core volumes: No significant differences in hypoperfusion and core volumes were observed in groups A and B. Group B1 (flank effect) showed higher volumes. These were reduced after flank correction but were still significant for the hypoperfusion volume.

For confounding factors such as technical failures (strong patient movements, very low SNR) and higher-grade stenosis, the
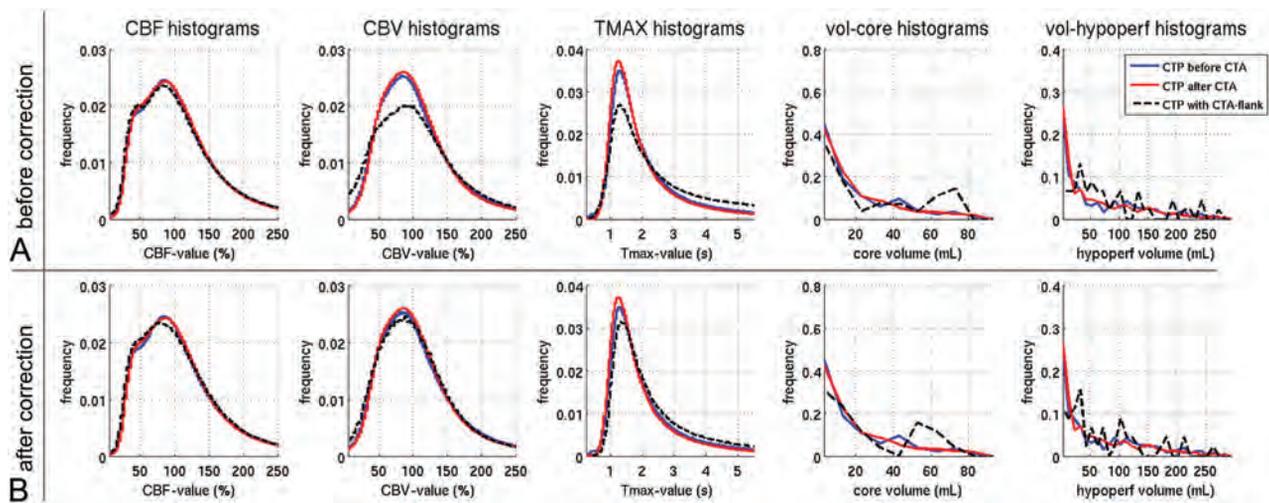
FIG 4. Histograms of perfusion parameters and infarction volumes. Upper row (*A*) and lower row (*B*) show the values before and after correction, respectively. Between the groups CTA before CTP and CTP before CTA, there is no relevant net effect on the perfusion metrics (*red versus blue lines*). Only in case of bolus interference (*black dotted lines*) do CBV and Tmax values tend to be misestimated (*A*). After application of the correction method, these differences are only marginal (*B*). hypoperf indicates hypoperfused; vol, volume.
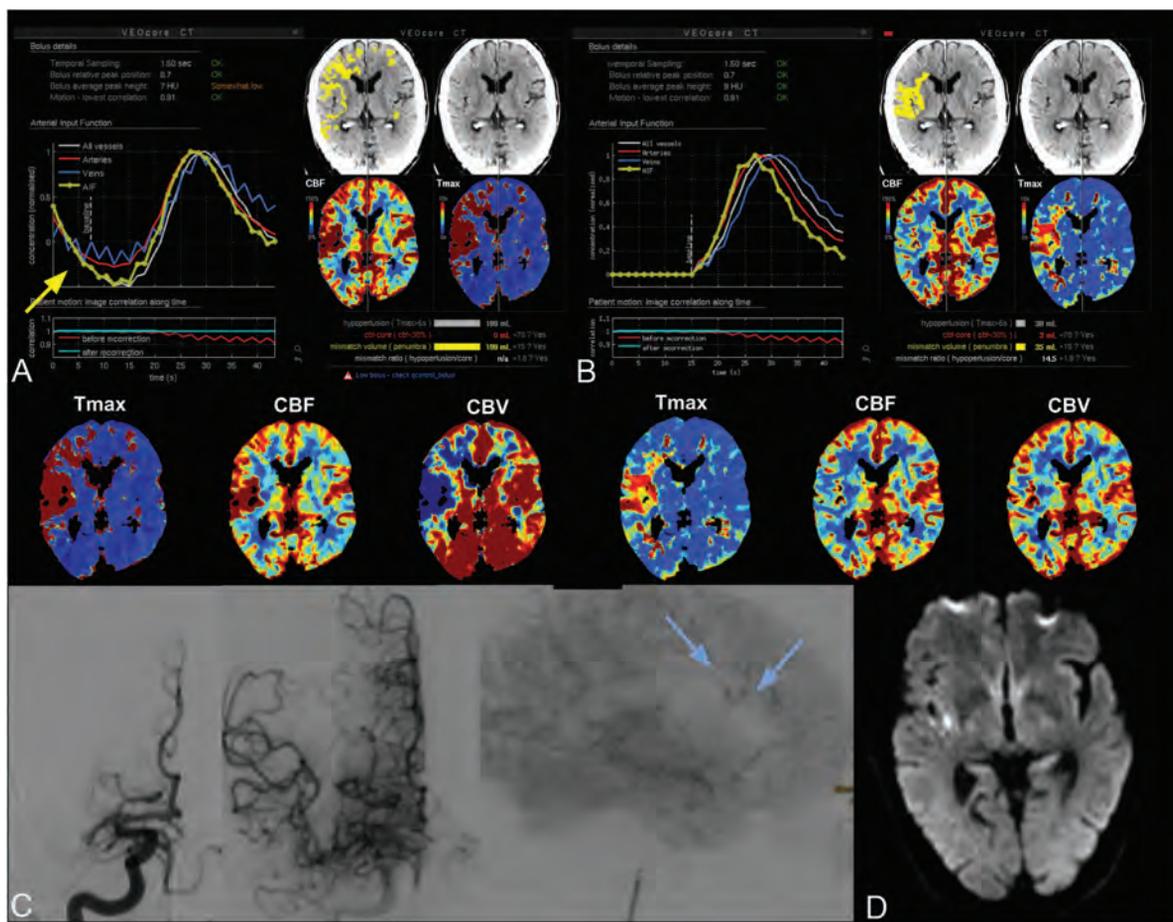
FIG 5. A 61-year-old man with left-sided hemiparesis and dysarthria due to atrial fibrillation and a cardiac ejection fraction of 20%. *A,* The time-density curves of CTP after CTA show the descending flank of the CTA bolus extending into the beginning of the CTP measurement (*A, yellow arrow*). Calculated CBF and CBV values are low (*blue arrows*). *B,* After mathematic correction, there is an appropriate time-density curve with a circumscript hypoperfusion area. *C,* Following TICI 2b recanalization of a M1 segment occlusion (*blue arrows),* follow-up MR imaging shows only tiny insular (*D*) and peripheral infarcts (not shown).

Table 2: Counts of cases excluded due to technical failures and the presence of ipsilateral extracranial carotid artery stenosis, derived from DSA data[a]

No. of Cases	Total	CTP before CTA	CTA before CTP	CTP with Flank Effect
Total	2074	393	1628	53
Excluded due to technical issues	94 (4.5%)	25 (6.3%)	65 (4.0%)	4 (7.5%)
Finally used in evaluation	1980	368	1563	49
With subsequent DSA	387 (20%)	76 (21%)	300 (19%)	11 (22%)
With ICA stenosis identified in DSA	86 (22%)	21 (28%)	62 (21%)	3 (27%)

[a] Data are the number of cases (percentage). Relative counts of technical failures and the presence of higher-grade carotid stenosis are similar in all groups

relative counts were similar in all groups, as summarized in Table 2 (approximately 5% technical failures, and 21%–28% higher-grade stenosis).

DISCUSSION

In this study, we analyzed the time-density curves and perfusion metrics of a large number of patients with stroke for CTP performed before or after CTA. In the raw concentration curves, the mean baseline values were, on average, 4 HU higher when CTP was performed after CTA, which is in line with the theoretic expectation due to venous filling with contrast agent from the preceding CTA bolus. However, as expected, the difference cancels out when calculating the tracer concentration curves. Consequently, no relevant net effect on the perfusion metrics and infarct core and hypoperfusion volumes was observed. These findings are in line with the results from the study by Dorn et al.[5]

However, due to the CTA flank effect, which occurred in around 3% of patients in this study, the interference of CTA and CTP boluses can suggest erroneously low CBV values when not corrected. Regarding the hypoperfusion and core volumes, the group with the flank effect showed significantly larger values, which were present even after application of the correction method. The flank group comprised more elderly patients, who are more likely to have larger infarctions, which might partly explain this difference. Furthermore, although not addressed in this study, low cardiac output has already been identified as a factor influencing the nature of the contrast bolus.[9]

Additionally, higher-grade carotid artery stenosis might be another factor encouraging a bolus interference. However, the average relative stenosis counts were similar in all groups (21%–28%, Table 2). Therefore, it seems that the dominant factor with respect to the prevalence of the flank effect is the time gap between CTA and CTP measurement, as also indicated in Table 1 and Fig 2.

According to our data, the median and maximum time intervals in the flank group were 18 and 25 seconds, respectively. These indicate that a pause of 30 seconds from the end of CTA and the start of CTP should be sufficient from a practical point of view.

Another strategy could be to consider the duration of the bolus-tracking interval—that is the time interval from start of the contrast injection for CTA and the time point when the density values in the ascending aorta reach the predefined threshold of 100 HU.

Copelan et al[10] described an underestimation of the infarct core when secondarily assigned patients had previously received IV contrast media. However, the delay between CTA and CTP was >30 minutes, indicating that this underestimation was due either to their software-specific postprocessing, possibly the arterial input function selection used, or other factors in the infarct dynamics themselves such as ischemic barrier disruption.

A potential technical issue not relevant to our study is the impact on masking the perfusion maps, ie, removal of CSF voxels. If such a masking is performed on the basis of the baseline images of CTP, masking might be insufficient due to venous filling if performed after CTA. The VEOcore software used in this study used NCCT for brain masking and was thus unaffected; however, that might be different for implementations of other vendors.

Another limitation is that we could not validate the accuracy of CTP maps with respect to the final outcome because only a minority of patients with TICI 2b, 2c, or 3 recanalization had follow-up MR imaging allowing calculation of Dice indices.

CONCLUSIONS

CTP can be performed after CTA; however, interferences from a preceding CTA bolus must be avoided to prevent miscalculations. This result can be achieved by scheduling a fixed delay on the order of 30 seconds or by applying a correction method.

REFERENCES

1. Nogueira RG, Jadhav AP, Haussen DC, et al. DAWN Trial Investigators. **Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct.** *N Engl J Med* 2017;378:11–21 CrossRef Medline
2. Albers GW, Marks MP, Kemp S, et al. **Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging.** *N Engl J Med* 2018;378:708–18 CrossRef Medline
3. Powers WJ, Rabinstein AA, Ackerson T, et al. American Heart Association Stroke Council. **Guidelines for the Early Management of Patients with Acute Ischemic Stroke: 2019 Update to the 2018 Guidelines for the Early Management of Acute Ischemic Stroke—A Guideline for Healthcare Professionals from the American Heart Association/American Stroke Association.** *Stroke* 2019;50:e344–418 CrossRef Medline
4. Turc G, Bhogal P, Fischer U, et al. **European Stroke Organisation (ESO): European Society for Minimally Invasive Neurological**

Therapy (ESMINT) Guidelines on Mechanical Thrombectomy in Acute Ischaemic Stroke Endorsed by Stroke Alliance for Europe (SAFE). *Eur Stroke J* 2019;4:6–12 CrossRef Medline

5. Dorn F, Liebig T, Muenzel D, et al. **Order of CT stroke protocol (CTA before or after CTP): impact on image quality.** *Neuroradiology* 2012;54:105–12 CrossRef Medline

6. Morhard D, Wirth CD, Reiser MF, et al. **Optimal sequence timing of CT angiography and perfusion CT in patients with stroke.** *Eur J Radiol* 2013;82:e286–89 CrossRef Medline

7. de Lucas EM, Sánchez E, Gutiérrez A, et al. **CT protocol for acute stroke: tips and tricks for general radiologists.** *Radiographics* 2008;28:1673–87 CrossRef Medline

8. Miles KA, Young H, Chica SL, et al. **Quantitative contrast-enhanced computed tomography: is there a need for system calibration?** *Eur Radiol* 2007;17:919–26 CrossRef Medline

9. Konno M, Hosokai Y, Usui A, et al. **Cardiac output obtained from test bolus injections as a factor in contrast injection rate revision of following coronary CT angiography.** *Acta Radiol* 2012;53:1107–11 CrossRef Medline

10. Copelan AZ, Smith ER, Drocton GT, et al. **Recent administration of iodinated contrast renders core infarct estimation inaccurate using RAPID software.** *AJNR Am J Neuroradiol* 2020;41:2235–42 CrossRef Medline

Transmantle Pressure Computed from MR Imaging Measurements of Aqueduct Flow and Dimensions

S.J. Sincomb, W. Coenen, E. Criado-Hidalgo, K. Wei, K. King, M. Borzage, V. Haughton, A.L. Sánchez, and J.C. Lasheras

ABSTRACT

BACKGROUND AND PURPOSE: Measuring transmantle pressure, the instantaneous pressure difference between the lateral ventricles and the cranial subarachnoid space, by intracranial pressure sensors has limitations. The aim of this study was to compute transmantle pressure noninvasively with a novel nondimensional fluid mechanics model in volunteers and to identify differences related to age and aqueductal dimensions.

MATERIALS AND METHODS: Brain MR images including cardiac-gated 2D phase-contrast MR imaging and fast-spoiled gradient recalled imaging were obtained in 77 volunteers ranging in age from 25–92 years of age. Transmantle pressure was computed during the cardiac cycle with a fluid mechanics model from the measured aqueductal flow rate, stroke volume, aqueductal length and cross-sectional area, and heart rate. Peak pressures during caudal and rostral aqueductal flow were tabulated. The computed transmantle pressure, aqueductal dimensions, and stroke volume were estimated, and the differences due to sex and age were calculated and tested for significance.

RESULTS: Peak transmantle pressure was calculated with the nondimensional averaged 14.4 (SD, 6.5) Pa during caudal flow and 6.9 (SD, 2.8) Pa during rostral flow. It did not differ significantly between men and women or correlate significantly with heart rate. Peak transmantle pressure increased with age and correlated with aqueductal dimensions and stroke volume.

CONCLUSIONS: The nondimensional fluid mechanics model for computing transmantle pressure detected changes in pressure related to age and aqueductal dimensions. This novel methodology can be easily used to investigate the clinical relevance of the transmantle pressure in normal pressure hydrocephalus, pediatric communicating hydrocephalus, and other CSF disorders.

ABBREVIATIONS: FSPGR = fast-spoiled gradient recalled; PC = phase-contrast; VENC = velocity encoding

Transmantle pressure, the instantaneous pressure difference between the lateral ventricles and the cranial subarachnoid space, drives cyclic CSF flow in and out of the ventricles through the cerebral aqueduct. The aqueductal flow velocities driven by transmantle pressure have clinical potential value in triaging patients with normal pressure hydrocephalus,[1,2] investigating the pathogenesis of hydrocephalus,[3] and understanding the effects of craniotomy (trephination)[4,5] or lumbar puncture,[6] and in other

indications. Transmantle pressure can be measured using simultaneous readings from dual pressure sensors placed surgically in the cranial vault. However, there is a need to develop an accurate methodology to calculate it through nonintrusive methods. This represents a challenge because the transmantle pressure is very small. Results of direct measurements in both humans and animals typically show small differences or differences within the experimental error of measurement.[7-10] Therefore, some prior studies have argued that these small changes of transmantle pressure could not be responsible for the development of hydrocephalus.[9,10]

Pressure drop is approximately proportional to the fluid acceleration, which increases in narrow channels such as the cerebral aqueduct. The pressure drop between the lateral ventricles and the

subarachnoid space is then mainly due to that occurring along the aqueduct, which is the lengthiest narrow passage in the ventricular system.[11] Therefore, noninvasive calculations of transmantle pressure typically focus on measuring the flow acceleration along the cerebral aqueduct.[11-14] These noninvasive computational studies estimate small transmantle pressure values on the order of a few Pascals, consistent with those measured invasively. However, previously reported calculations of transmantle pressure were performed on a limited number of subjects, neglected relevant flow effects like convective acceleration, and/or relied on full computational fluid dynamics simulations, which are difficult to implement clinically.

The goal of this study was to compute transmantle pressure from noninvasive aqueductal flow measurements using a simplified nondimensional fluid mechanics model previously published by Sincomb et al,[15] which accounts for the specific anatomic features of the aqueduct and relevant effects of local fluid acceleration, convective acceleration, and viscous forces. We applied this novel methodology to a group of volunteers to investigate differences related to age, sex, flow, and variations such as stroke volume, aqueductal length, and cross-sectional area on the transmantle pressure.

MATERIALS AND METHODS

Study Subjects

Subjects for this study were participants in a brain aging study funded by the LK Whittier Foundation ($n = 48$) and patients ($n = 34$) in a brain health and memory loss clinic and healthy student volunteers ($n = 5$). All subjects signed consent for enrollment in this institutional review board–approved study through Huntington Medical Research Institutes. Individuals with known neurologic or neurosurgical disorders were excluded. In the memory loss clinic group, individuals with a Clinical Dementia Rating of >0.5 of 5 were excluded.[16] Written informed consent for inclusion in the study was obtained for all participants. All images were reviewed by a neuroradiologist to identify and exclude cases of acute or obstructive hydrocephalus and other abnormalities. Blood pressure and arterial pulse pressure were not recorded.

MR Imaging Acquisition

Imaging was performed with a 3T Signa scanner (software Version HD23; GE Healthcare) with an 8-channel head coil. The MR imaging included the routine localizer, axial, coronal, and sagittal T1- and T2-weighted images. Sequences obtained included a sagittal 3D inversion recovery fast-spoiled gradient recalled (FSPGR) sequence with the following parameters: TR/TE = 6.9/2.5 ms, TI = 600 ms, flip angle = 8°, acquisition voxels = 0.9 × 0.9 × 1.2 mm; and a single-shot gradient recalled-echo 2D phase-contrast (PC) sequence acquired perpendicular to the cerebral aqueduct, with the following parameters: TR/TE = 20/10 ms, velocity encoding (VENC) = 5–20 cm/s, flip angle = 20°, peripheral pulse gating with 30 retrospective cardiac phases, voxel size = 0.2 × 0.2 × 5.0 mm to 0.47 × 0.47 × 5.0 mm.

Flow Analysis

PC MR images were processed in Matlab (Version R2019a; MathWorks). The investigator placed a cursor on the aqueduct in the average-magnitude image. The program generated an optimal rectangular ROI around the cursor for the Otsu threshold algorithm and converted pixels in the ROI to a binary scale
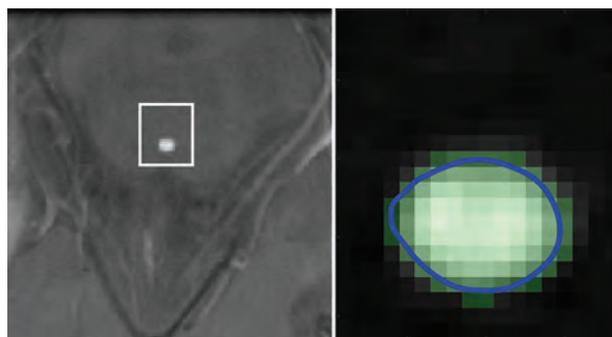
FIG 1. Axial PC MR imaging magnitude image (*left*) illustrating the aqueduct and optimal ROI window (*white line*) and corresponding binary image (*right*) from the Otsu algorithm (light green) displaying the boundary (*blue line*) automatically overlaid onto the magnitude image.
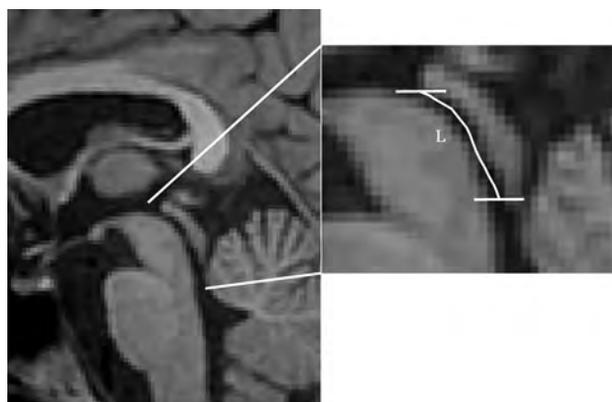
FIG 2. Sagittal FSPGR MR image (*left*) showing the aqueduct, third and fourth ventricles, and a collimated image (*right*). Enlarged MR image (*right*) shows the length of the aqueduct by means of a line connecting its centroid at each section (*white line*). The *horizontal lines* show the junction of the aqueduct with the third and fourth ventricles.

(graythresh). A boundary was computed separating the black and white pixels, and the cross-sectional area of the aqueduct was calculated automatically as the area within the boundary (bwboundaries) (Fig 1). Flow in each white pixel of the aqueductal area was calculated by linear magnitude–weighted conversion of phase to velocity and background correction.[17,18] If the temporal velocity difference in a voxel was larger than 1.1 times the velocity encoding, antialiasing was performed.[19] The flow rate $Q(t)$, was computed by averaging the velocity magnitudes of all pixels within the aqueductal ROI and multiplying by the cross-sectional area. Caudal flow has a positive sign, and rostral flow, a negative sign.[20] The stroke volume was computed as half the integral of the flow rate.[2] Subjects with evident motion-related artifacts were excluded. In 1 subject imaged twice, inter-examination repeatability of the flow rate was calculated.

FSPGR images were used to segment the aqueduct semiautomatically using ITK-SNAP (Version 3.6.0; www.itksnap.org).[21] The rostral and caudal ends of the aqueduct were identified manually as the point where the fluid lumen increased by 50%. The resulting 3D binary image stack was subsequently processed in Matlab. The geometric centroids of the aqueductal cross-sectional area were automatically identified in the axial plane of the 3D image stack, which was used along with the section spacing to compute the length of the aqueduct (Fig 2).
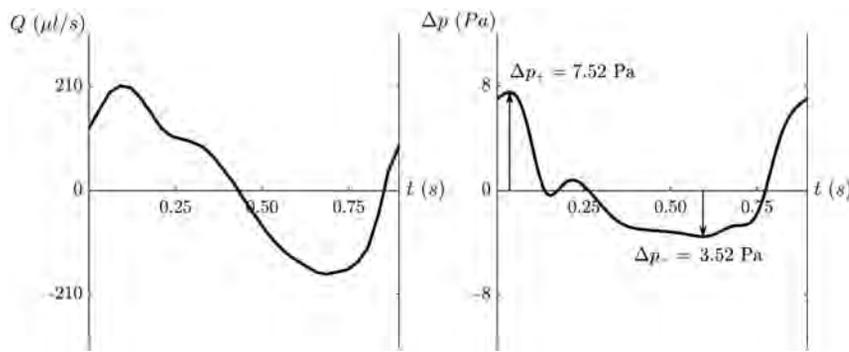
FIG 3. Plot of aqueductal flow (*left*) in a healthy 37-year-old man with an average heart rate of 67 beats per minute. Plot of the transmantle pressure (*right*) after entering variables into the nondimensional model[15] and converting back to dimensional pressure (in Pascals). On the plot of transmantle pressure, the peak positive (ΔP_+) and negative (Δp_-) transmantle pressures are demonstrated by *arrows* representing peak pressures during caudal and rostral aqueductal flow, respectively.

Table 1: Age, heart rate, aqueductal flow, and anatomic features and peak transmantle pressures[a]

	Total (*n* = 77)	Female (*n* = 49)	Male (*n* = 28)	Sex (*P* Value)
Age (yr)	69 (SD, 14)	72 (SD, 13)	65 (SD, 16)	.114
Heart rate (BPM)	64 (SD, 9)	63 (SD, 7)	66 (SD, 11)	.305
Flow and anatomic features				
Area (mm^2)	3.15 (SD, 1.03)	2.90 (SD, 0.88)	3.58 (SD, 1.13)	.016
Stroke volume (μl)	40 (SD, 20)	40 (SD, 20)	42 (SD, 20)	.571
Length (mm)	17.3 (SD, 2.8)	16.9 (SD, 2.8)	18.0 (SD, 2.5)	.065
Peak transmantle pressure				
Pressure+ (Pa)	14.44 (SD, 6.49)	14.54 (SD, 6.13)	14.25 (SD, 7.18)	.564
Pressure− (Pa)	6.94 (SD, 2.82)	7.02 (SD, 2.69)	6.79 (SD, 3.07)	.586

Note:—BPM indicates beats per minute; +, peak positive pressure during caudal flow; −, peak negative pressure during rostral flow.
[a] Data are means unless otherwise indicated.

Transmantle Pressure Computation

The pressure difference between the third and fourth ventricles was calculated using the fully nondimensional computational fluid dynamics method previously described in Sincomb et al.[15] The reader is referred to this publication for the details of the computational fluid dynamics model. The corresponding computer codes to calculate the transmantle pressure will be made available by the authors on request. One advantage of nondimensionalization is that it reduces the computational cost while retaining all important physical phenomena affecting the pressure variation including the local acceleration, convective acceleration, and viscous forces. In this respect, the model is superior in accuracy to 1-term formulas directly applicable to the PC MR image.[22] The main input to the model included the aqueductal length, aqueductal cross-sectional area, flow rate $Q(t)$, stroke volume, and heart rate. The cross-sectional area of the aqueduct was assumed to be constant along the length of the aqueduct for this computation. Each variable was scaled with its corresponding characteristic value. The density and kinematic viscosity of CSF were assumed as $\rho \simeq 10^3 \, \text{kg/m}^3$ and $\nu \simeq 0.71 \times 10^{-6}$ m^2/s, respectively. As mentioned previously, the resulting transmantle pressure is calculated as the pressure difference between

the third and fourth ventricles, given as $\Delta p = p_3(t) - p_4(t)$, because the drop in pressure across the foramen of Monro and the median aperture is an order of magnitude smaller. This estimate of the transmantle pressure also neglects spatial pressure variations in the subarachnoid space because there the pressure is nearly uniform at any instant in time; thus, the spatial pressure differences never exceed a fraction of a Pascal as revealed in detailed computational fluid dynamics simulations of the entire cranial cavity.[11] We, therefore, excluded any subject with MR imaging evidence of subarachnoid space obstruction.

For all subjects, we computed the variation of transmantle pressure over a cardiac cycle (Fig 3). The variation of the transmantle pressure was then used to determine the peak positive transmantle pressure difference and negative transmantle pressure difference during caudal and rostral flow, respectively.

Statistical Analysis

Differences in flow and anatomic and transmantle pressure values between men and women were tested by means of the Mann-Whitney *U* test. The Spearman rank correlation was used to assess the relationship between peak transmantle pressure values and aqueductal flow, age, heart rate, aqueductal length, and aqueductal area. These relationships were illustrated with scatterplots. Statistical significance was set as a $P < .05$.

RESULTS

Study Subjects

Of the 87 subjects, 9 subjects were excluded due to imaging artifacts such as motion. One subject with stroke volume >3 times the average was excluded as an outlier, which did not affect significant *P* values. The subjects for this study included 77 individuals, 28 men and 49 women, 25–92 years of age (average 69 [SD, 14] years) (Table 1). No subjects were excluded on the basis of review of MR images. Male and female subjects, on average, had similar ages and heart rates (Table 1).

Flow Analysis

Eight subjects had high velocities that exceeded our VENC and thus benefited from antialiasing. The subject scanned twice for interexamination repeatability showed a difference in the estimated cross-sectional area of 5.6% and peak flow rates of 3.2% (Online Supplemental Data). In the 77 subjects, the cross-sectional area of the aqueduct averaged 3.15 (SD, 1.03) mm^2 and

Table 2: Correlation of peak positive and negative transmantle pressure with age, heart rate, aqueductal area, stroke volume, and aqueductal length

	Pressure+ (Pa)		Pressure− (Pa)	
	Correlation Coefficient	**P**	**Correlation Coefficient**	**P**
Age (yr)	0.33	.004	0.18	.107
Heart rate (BPM)	−0.12	.284	0.11	.343
Area (mm²)	−0.19	.093	−0.28	.013
Stroke volume (µl)	0.51	<.001	0.28	.014
Length (mm)	0.33	.004	0.37	.001

Note:—BPM indicates beats per minute; +, peak positive pressure during caudal flow; −, peak negative pressure during rostral flow.
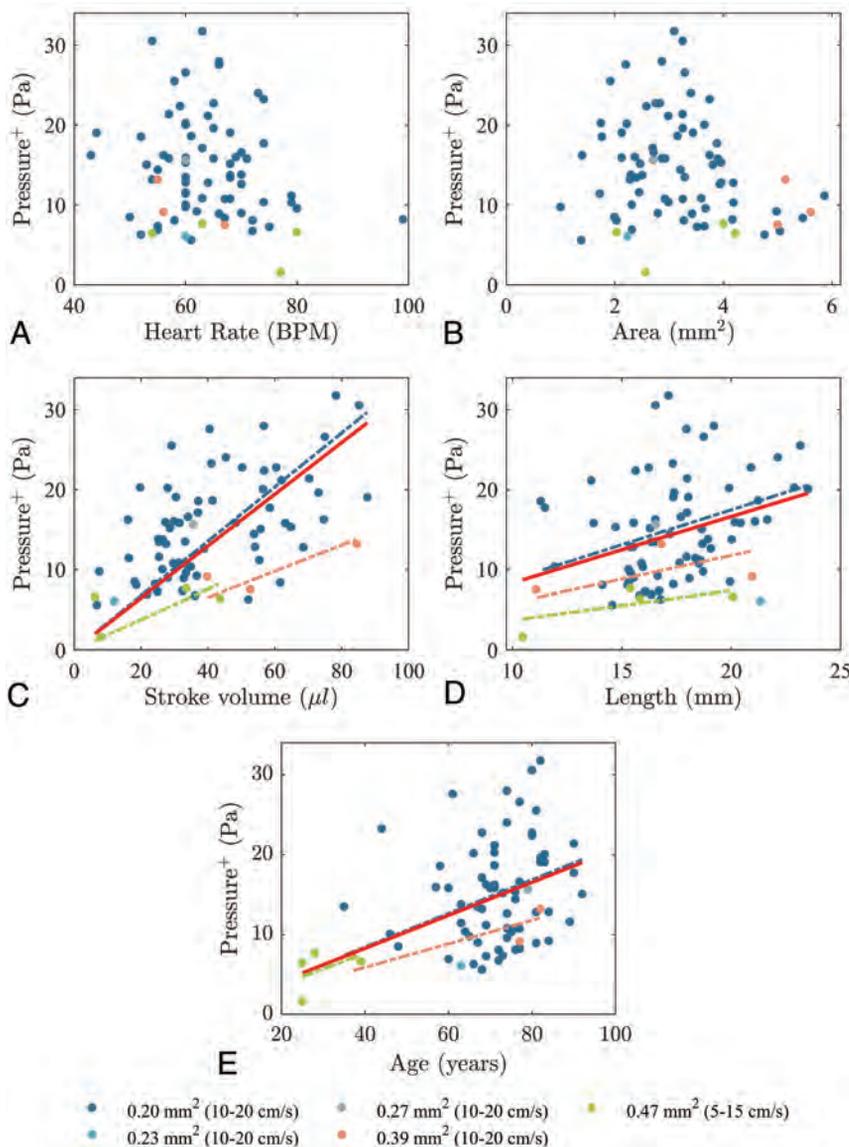
FIG 4. Peak transmantle pressure for all subjects and for subgroups of different voxel sizes (and VENC values) during caudal flow plotted against heart rate (*A*), aqueductal cross-sectional area (*B*), stroke volume (*C*), aqueductal length (*D*), and age (*E*), with trendlines including all subjects (*red solid line*) and subgroups of different voxel sizes with >1 subject. Transmantle pressure increases with stroke volume ($\rho = 0.51$, $P < .001$), aqueductal length ($\rho = 0.33$, $P = .004$), and age ($\rho = 0.33$, $P = .004$) for all subjects. Transmantle pressure also increases in each subgroup, but not significantly. BPM indicates beats per minute.

differed significantly ($P = .016$) between women and men. The aqueductal length averaged 17.3 (SD, 2.8) mm, and the stroke volume averaged 40 (SD, 20) µl (Table 1).

Transmantle Pressure

During caudal aqueductal CSF flow, the peak positive transmantle pressure (higher pressure in the third ventricle than in fourth ventricle) in our participants averaged 14.44 (SD, 6.49) Pa and ranged from 1.63 to 31.77 Pa. During rostral aqueductal CSF flow, the negative transmantle pressure (higher pressure in the fourth ventricle than in the third) averaged 6.94 (SD, 2.82) Pa and ranged from 1.80 to 14.01 Pa. Differences between women and men in transmantle pressures were not significant (Table 1).

Statistical Analysis

Peak transmantle pressure during caudal flow correlated significantly with age ($\rho = 0.33$, $P = .004$), aqueductal length ($\rho = 0.33$, $P = .004$), and stroke volume ($\rho = 0.51$, $P < .001$) (Table 2). It did not correlate with heart rate or aqueductal cross-sectional area, though the *P* value was near significance. Peak transmantle pressure during rostral flow correlated with the aqueductal cross-sectional area ($\rho = -0.28$, $P = .013$), aqueductal length ($\rho = 0.37$, $P = .001$), and stroke volume ($\rho = 0.28$, $P = .014$). It did not correlate with heart rate or age (Table 2). Scatterplots of peak positive and negative transmantle pressure with age, heart rate, aqueductal cross-sectional area, aqueductal length, and stroke volume are shown in Figs 4 and 5.

DISCUSSION

We found that peak transmantle pressure during caudal flow increased with age from 1.62 Pa in one 25-year-old subject to 15.05 Pa in the 92-year-old subject. This finding indicates the relevance of age in the estimation of transmantle pressure. Aqueduct length and stroke volume correlated significantly with transmantle pressures in both the caudal and rostral direction. Peak transmantle pressure was inversely correlated with aqueductal cross-sectional
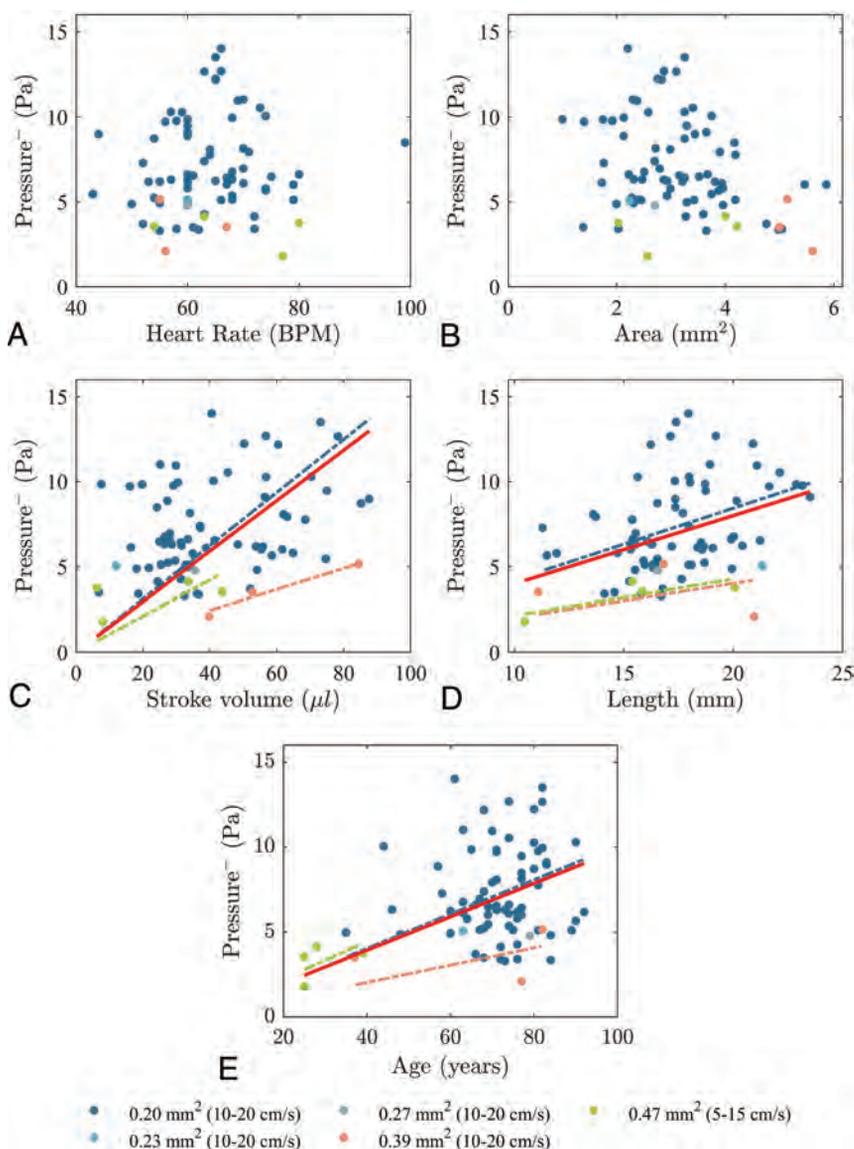
FIG 5. Peak transmantle pressure for all subjects and for subgroups of different voxel sizes (and VENC values) during rostral flow plotted against heart rate (*A*), aqueductal cross-sectional area (*B*), stroke volume (*C*), aqueductal length (*D*), and age (*E*), with trendlines including all subjects (*red solid line*) and subgroups of different voxel sizes with >1 subject. Transmantle pressure increases with stroke volume ($\rho = 0.28$, $P = .014$) and aqueductal length ($\rho = 0.37$, $P = .001$) for all subjects. The transmantle pressure tends to increase with age for the total group but does not reach significance ($\rho = 0.18$, $P = .107$). Transmantle pressure also increases in each subgroup, but not significantly. BPM indicates beats per minute.

transmantle pressures were immeasurably small, the pressure sensors placed in the cranial vault had a resolution of 66 Pa, which is greater than the typical transmantle pressure in our study.[8] To our knowledge, the peak pressures during both caudal and rostral flow have not been previously reported separately as we have done. This issue is important because the dynamics for caudal and rostral CSF flow (during systole and diastole, respectively) are distinct and may reflect influences of different system properties. The stroke volumes we measured in our volunteers also had a magnitude similar to those in previous studies including healthy younger and elderly subjects.[18,26-28]

The study reported here is important because it shows that computational fluid mechanics methods based on a nondimensional formulation have sufficient precision to detect differences in transmantle pressure due to age or aqueductal dimensions, 2 important issues that have not been sufficiently evaluated to date. Our approach differs from that in previous studies related to transmantle pressure computation. For example, Sartoretti et al[22] used the Bernoulli equation written in the form $\Delta p = 4V_{max}^2$ to determine the pressure difference. This approach is only valid for steady flow, a condition that is not realized for the flow in the aqueduct. Linninger et al[25] implemented a full computational fluid dynamics simulation that was computationally expensive. The computation used here can be performed with a readily available straightforward computer program in a few minutes and may be easily applied in the clinical setting.

Computation of transmantle pressure may be preferable to measurement of aqueductal velocity or stroke volume as an index of abnormal CSF dynamics. Computational studies of transmantle pressures are warranted in normal pressure hydrocephalus, pediatric communicating hydrocephalus, and other conditions with altered CSF dynamics. Owing to the viscoelastic nature of the brain parenchyma, even very small transmantle pressure may result in a gradual enlargement of the ventricles through the accumulation of small residual strain during months or years (tens of millions of cardiac cycles), while the mean intracranial pressure remains constant.[29] This methodology to estimate transmantle pressure may have a role in studies of factors affecting ventricular morphology.[30]

area during caudal flow in the aqueduct, but not significantly during the typically less rapid rostral flow, suggesting that the viscous boundary layer effects must be taken into consideration when computing the transmantle pressure.

Peak transmantle pressure computed with the nondimensional computational model varied during the cardiac cycle from an average of 14.44 (SD, 6.49) Pa during caudal flow in the aqueduct to 6.94 (SD, 2.82) Pa during rostral flow. The computed peak transmantle pressure during caudal and rostral flow in this study is in general agreement with that in other computational studies[23-25] and also with measurements made with invasive methods in humans[9,10] and in animals.[7] In reports in which

This study has limitations. The health of the volunteers was not rigorously tested. However, acute or obstructive hydrocephalus and other abnormal brain conditions detectable by MR imaging that affect CSF dynamics were excluded by the neuroradiologist's review of MR images in each case. For PC MR imaging acquisition, variation of the VENC and in-plane resolution may have introduced some bias through velocity noise and partial volume effects respectively.[31,32] However, groups with the same in-plane resolution showed the same trend for age and peak transmantle pressure. Neglecting longitudinal changes in the cross-sectional area of the aqueduct, for example at the ampulla, may have diminished the accuracy of the pressure computations. We did not study the effect of respiration on transmantle pressure, though it may be relevant.[33] The test-retest precision of the pressure computation was not quantified, and it requires further study. Validation of the method also requires additional in vivo and in vitro experimentation.

CONCLUSIONS

Transmantle pressure computed with a recently described nondimensional fluid mechanics model revealed an increase of transmantle pressure with age. Aqueduct flow and anatomic dimensions affected the pressure computed. This computational method which incorporates relevant flow and anatomic features of the aqueduct may contribute to future studies of NPH, pediatric communicating hydrocephalus, and other CSF disorders.

ACKNOWLEDGMENTS

We wish to dedicate this work to Professor Juan C. Lasheras who passed away on February 1, 2021, after a short battle with cancer. Juan was the guide in our quest to bridge the fields of engineering and medicine. S.S. would like to acknowledge support from Bill & Melinda Gates Foundation. E.C.-H. would like to acknowledge Fundación Bancaria "la Caixa".

REFERENCES

1. Greitz D. **Reprint of: Radiological assessment of hydrocephalus: new theories and implications for therapy.** *Neuroradiol J* 2006;19:475–95 CrossRef Medline
2. Bradley WG Jr, Scalzo D, Queralt J, et al. **Normal-pressure hydrocephalus: evaluation with cerebrospinal fluid flow measurements at MR imaging.** *Radiology* 1996;198:523–29 CrossRef Medline
3. Bradley WG Jr, Bahl G, Alksne JF. **Idiopathic normal pressure hydrocephalus may be a "two hit" disease: benign external hydrocephalus in infancy followed by deep white matter ischemia in late adulthood.** *J Magn Reson Imaging* 2006;24:747–55 CrossRef Medline
4. Wagshul ME, Rashid S, McAllister JP, et al. **The effect of decompressive craniectomy on CSF pulsatility in experimental communicating hydrocephalus.** *Cerebrospinal Fuild Res* 2010;7(Suppl 1):S35 CrossRef
5. Scollato A, Gallina P, Bahl G, et al. **Decompressive craniectomy arrests pulsatile aqueductal CSF flux: an in vivo demonstration using phase-contrast MRI—case report.** *Br J Neurosurg* 2015;29:440–42 CrossRef Medline
6. Blitz AM, Shin J, Baledent O, et al. **Does phase-contrast imaging through the cerebral aqueduct predict the outcome of lumbar CSF drainage or shunt surgery in patients with suspected adult hydrocephalus?** *AJNR Am J Neuroradiol* 2018;39:2224–30 CrossRef Medline
7. Conner ES, Foley L, Black PM. **Experimental normal-pressure hydrocephalus is accompanied by increased transmantle pressure.** *J Neurosurg* 1984;61:322–27 CrossRef Medline
8. Penn RD, Lee MC, Linninger AA, et al. **Pressure gradients in the brain in an experimental model of hydrocephalus.** *J Neurosurg* 2005;102:1069–75 CrossRef Medline
9. Stephensen H, Tisell M, Wikkelsö C. **There is no transmantle pressure gradient in communicating or noncommunicating hydrocephalus.** *Neurosurgery* 2002;50:763–73 CrossRef Medline
10. Eide PK, Saehle T. **Is ventriculomegaly in idiopathic normal pressure hydrocephalus associated with a transmantle gradient in pulsatile intracranial pressure?** *Acta Neurochir (Wien)* 2010;152:989–95 CrossRef Medline
11. Sweetman B, Xenos M, Zitella L, et al. **Three-dimensional computational prediction of cerebrospinal fluid flow in the human brain.** *Comput Biol Med* 2011;41:67–75 CrossRef Medline
12. Jacobson EE, Fletcher DF, Morgan MK, et al. **Fluid dynamics of the cerebral aqueduct.** *Pediatr Neurosurg* 1996;24:229–36 CrossRef Medline
13. Jacobson EE, Fletcher DF, Morgan MK, et al. **Computer modelling of the cerebrospinal fluid flow dynamics of aqueduct stenosis.** *Med Biol Eng Comput* 1999;37:59–63 CrossRef Medline
14. Bardan G, Plouraboué F, Zagzoule M, et al. **Simple patient-based transmantle pressure and shear estimate from cine phase-contrast MRI in cerebral aqueduct.** *IEEE Trans Biomed Eng* 2012;59:2874–83 CrossRef Medline
15. Sincomb S, Coenen W, Sánchez AL, et al. **A model for the oscillatory flow in the cerebral aqueduct.** *J Fluid Mech* 2020;899 CrossRef
16. Hughes CP, Berg L, Danziger WL, et al. **New clinical scale for the staging of dementia.** *Br J Psychiatry* 1982;140:566–72 CrossRef Medline
17. Simoncini C. **Patient specific numerical modeling for the optimization of HCC selective internal radiation therapy an image-based approach (in French).** 2017. http://www.theses.fr/2017REN1S020. Accessed May 19, 2020
18. Enzmann DR, Pelc NJ. **Cerebrospinal fluid flow measured by phase-contrast cine MR.** *AJNR Am J Neuroradiol* 1993;14:1301–07 Medline
19. Lindstrøm EK, Ringstad G, Mardal KA, et al. **Cerebrospinal fluid volumetric net flow rate and direction in idiopathic normal pressure hydrocephalus.** *Neuroimage Clin* 2018;20:731–41 CrossRef Medline
20. Puy V, Zmudka-Attier J, Capel C, et al. **Interactions between flow oscillations and biochemical parameters in the cerebrospinal fluid.** *Front Aging Neurosci* 2016;8:1–8 CrossRef Medline
21. Ambarki K, Lindqvist T, Wåhlin A, et al. **Evaluation of automatic measurement of the intracranial volume based on quantitative MR imaging.** *AJNR Am J Neuroradiol* 2012;33:1951–56 CrossRef Medline
22. Sartoretti T, Wyss M, Sartoretti E, et al. **Sex and age dependencies of aqueductal cerebrospinal fluid dynamics parameters in healthy subjects.** *Front Aging Neurosci* 2019;11:199 CrossRef Medline
23. Levine DN. **The pathogenesis of normal pressure hydrocephalus: a theoretical analysis.** *Bull Math Biol* 1999;61:875–916 CrossRef Medline
24. Levine DN. **Intracranial pressure and ventricular expansion in hydrocephalus: Have we been asking the wrong question?** *J Neurol Sci* 2008;269:1–11 CrossRef Medline

25. Linninger AA, Xenos M, Zhu DC, et al. **Cerebrospinal fluid flow in the normal and hydrocephalic human brain.** *IEEE Trans Biomed Eng* 2007;54:291–302 CrossRef Medline

26. Balédent O, Gondry-Jouet C, Meyer ME, et al. **Relationship between cerebrospinal fluid and blood dynamics in healthy volunteers and patients with communicating hydrocephalus.** *Invest Radiol* 2004;39:45–55 CrossRef Medline

27. Stoquart-El Sankari S, Lehmann P, Villette A, et al. **A phase-contrast MRI study of physiologic cerebral venous flow.** *J Cereb Blood Flow Metab* 2009;29:1208–15 CrossRef Medline

28. Lokossou A, Metanbou S, Gondry-Jouet C, et al. **Extracranial versus intracranial hydro-hemodynamics during aging: a PC-MRI pilot cross-sectional study.** *Fluids Barriers CNS* 2020;17:1 CrossRef Medline

29. Sincomb S, Haughton V, Sanchez A, et al. **Strain accumulation visco-elastic ventriculomegaly hypothesis for the onset of idiopathic normal pressure hydrocephalus (iNPH).** *Biophys J* 2020;118:452a CrossRef

30. Chiang WW, Takoudis CG, Lee SH, et al. **Relationship between ventricular morphology and aqueductal cerebrospinal fluid flow in healthy and communicating hydrocephalus.** *Invest Radiol* 2009;44:192–99 CrossRef Medline

31. Tang C, Blatter DD, Parker DL. **Accuracy of phase-contrast flow measurements in the presence of partial-volume effects.** *J Magn Reson Imaging* 1993;3:377–85 CrossRef Medline

32. Lotz J, Meier C, Leppert A, et al. **Cardiovascular flow measurement with phase-contrast MR imaging: basic facts and implementation.** *Radiographics* 2002;22:651–71 CrossRef Medline

33. Vinje V, Ringstad G, Lindstrøm EK, et al. **Respiratory influence on cerebrospinal fluid flow: a computational study based on long-term intracranial pressure measurements.** *Sci Rep* 2019;9:9732 CrossRef Medline

Brazilian FRED Registry: A Prospective Multicenter Study for Brain Aneurysm Treatment—The BRED Study

L.B. Manzato, R.B. Santos, P.M.M. Filho, G. Miotto, A.M. Bastos, and J.R. Vanzin

ABSTRACT

BACKGROUND AND PURPOSE: The development of flow diverters has changed the endovascular approach to intracranial aneurysms. On the basis of good results, the indications for flow diverters have expanded to include aneurysms of different shapes, locations, and sizes. The objective of the study was to report on the performance of the Flow Re-Direction Endoluminal Device (FRED) in intracranial aneurysm treatment at early and medium-term follow-up.

MATERIALS AND METHODS: This single-arm, multicentric, prospective, observational study assessed aneurysm treatment with the FRED. The primary outcome was complete aneurysm occlusion at 6 and 12 months, and the secondary outcome was to evaluate the safety of the FRED with respect to stroke and death rates.

RESULTS: Between June 2016 and August 2018, a total of 100 consecutive patients with 131 aneurysms were treated in 107 procedures. Total occlusion rates were 91% and 95% at 6 and 12 months. There was 1 death, and the total final morbidity rate was 1.8%. The complication rate was 4.6%.

CONCLUSIONS: As reported previously, the FRED has proved to be a safe and effective tool, with high occlusion rates. The design of the stent makes it more difficult to perform balloon angioplasty compared with similar devices. A branch arising from the aneurysm sac was found to be a predictor of nonocclusion at 12 months, though larger series are needed to estimate the magnitude of the association.

ABBREVIATIONS: FD = flow diverter; FRED = Flow Re-Direction Endoluminal Device

Endovascular coiling is currently the treatment of choice for intracranial aneurysms.[1,2] The use of flow diverters has changed the concept of intracranial aneurysm treatment by shifting the focus from the sac to the parent vessel, where the aneurysm is actually located.[3,4] Large studies have demonstrated the efficacy and safety of flow-diverter stents,[5,6] and the range of indications has grown considerably, including aneurysms at the level of the circle of Willis and beyond.[7,8] Several stent-design enhancements have been incorporated in recent years, such as the development of low-profile devices compatible with small-lumen microcatheters, improvement of fluoroscopic visibility, surface modifications to reduce thrombus formation, and implementation of resheathing mechanisms to allow accurate deployment and repositioning of the stent.[9,10] The Flow Re-Direction Endoluminal Device (FRED;

MicroVention) is a braided stent-in-stent device composed of an inner closed-cell stent with low porosity and an outer mesh with high porosity. It has been assessed in only prospective multicenter studies, mainly with medium- and long-term follow-up.[11-15] In this registry, we report the performance of FRED as used in the treatment of intracranial aneurysms at early and medium-term follow-up. In this multicenter series, we prospectively evaluated the safety and efficacy of FRED and identified predictors of occlusion and complications.

MATERIALS AND METHODS

Study Design

We conducted a single-arm, prospective, observational study to assess intracranial aneurysms treated with the FRED at 3 interventional neuroradiology centers in Brazil (Interventional Neuroradiology Department, Dr. Astrogildo Hospital, Santa Maria, Rio Grande do Sul, Brazil). The study was approved by the research ethics committees of the participating centers, and written informed consent was obtained from each study participant.

All surgeons were certified by the Brazilian Society of Neuroradiology. Centers were eligible for participation in this

study if they had treated >100 aneurysms in each of the preceding 3 years with a minimum of 50 cases treated with flow-diverter stents and had a neurosurgical team participating in all treatment decisions.

The inclusion criteria were patients of all ages with intracranial aneurysms, regardless of size, topography, location, and ruptured or unruptured status, and untreated or recurrent aneurysms after previous surgical or endovascular treatment. Coils were allowed during treatment. Patients were excluded if they had a history of a coagulation disorder, evidence of active infection, or unfavorable arterial anatomy for stent deployment (an acute angle formed by the side branch and parent vessel, a parent artery of <2.0 mm, or severe stenosis of the target vessel).

Primary End Point
The primary outcome was complete aneurysm occlusion as assessed by DSA at 6 and 12 months. Complete occlusion was defined as complete obliteration of the aneurysm sac, including the neck.

Secondary End Point
As a secondary outcome, we evaluated the safety of the FRED and FRED Jr (MicroVention) regarding stroke and death rates. Procedure-related complications were also reported, regardless of clinical impact, as well as stent underexpansion, thromboembolic events, and artery dissection or perforation.

Endovascular Procedure
All procedures were performed with the patient under general anesthesia. After puncture of the common femoral artery, a 5000-IU bolus of heparin was administered. A standard triaxial apparatus was used in all cases to navigate through the stent. A delivery microcatheter (Headway 0.027 or 0.021; MicroVention) and microguidewire (Traxcess 0.014; MicroVention) were used to advance the device to the aneurysm neck under roadmap guidance.

For antiplatelet therapy, patients were given either aspirin (100 mg) + prasugrel (10 mg) or aspirin (100 mg) + clopidogrel (75 mg). The patients were started on medications at least 5 days before the elective procedure. Antiplatelet aggregation tests were not performed in any of the patients due to their unavailability at the participating centers. Patients with ruptured aneurysms were given a bolus of 500 mg of aspirin + 600 mg of clopidogrel or 60 mg of prasugrel 5 hours before embolization. Dual-antiplatelet therapy was continued for 6 months until the first follow-up angiogram. Clopidogrel or prasugrel was then discontinued if there was no evidence of in-stent stenosis or hyperplasia. Aspirin was not discontinued unless hemorrhagic complications occurred.

Data Collection
All data were collected prospectively using a predefined data base query for demographic information, such as neurologic status (assessed by the mRS) and aneurysm characteristics (location, topography, rupture or unruptured status, morphology, diameter, neck size, dome-to-neck ratio, and previous treatment). Perioperative data included treated aneurysms, flow-diverter size, balloon angioplasty, technical complications, and thromboembolic

complications. Postoperative follow-up included neurologic and systemic complications, stent migration, occlusion rate, and neurologic status.

Patients were evaluated clinically, including the mRS score, before treatment, at hospital discharge, and at 30 days, 6 months, and 12 months. Vascular imaging was performed at 6 and 12 months, and data were collected from cerebral angiograms. Noninvasive imaging was not acceptable for follow-up purposes.

Data Analysis
Statistical analysis was performed using SPSS, Version 26 for Windows (IBM). Categoric variables were expressed as absolute and relative frequencies, and numeric variables were expressed as mean (SD). Potential predictors of nonocclusion at 6 and 12 months were tested by univariate binary logistic regression analysis. The probability (P) values were estimated using the likelihood ratio test. A P value < .05 was considered statistically significant.

RESULTS
Between June 2016 and August 2018, one hundred consecutive patients with 131 intracranial aneurysms were treated at the 3 participating centers for a total of 107 procedures. During enrollment, no flow diverters other than the FRED were used to treat intracranial aneurysms in the participating centers.

Patient and Aneurysm Characteristics
Of the 100 patients, 89 were women. The mean patient age was 52.2 (SD, 12.3) years. Ninety-five patients had incidental aneurysms; 1 patient presented with subarachnoid hemorrhage; 1, with transient ischemic attack; and 3, with mass effect symptoms. Twenty-four patients had multiple aneurysms. Two aneurysms had been previously treated with an operation, and 2 had been treated with other flow diverters, all of which failed to achieve complete closure. All other patients were previously untreated. The preoperative mRS score was 0 in 100% of patients.

Of the 131 intracranial aneurysms, 124 were located in the ICA; 3, in the vertebrobasilar system; and 4, beyond the circle of Willis (3 in the anterior communicating artery and 1 in the middle cerebral artery). The most common location was the C6 segment of the ICA (88 aneurysms). Except for 1 dissecting aneurysm, all other aneurysms were saccular. One hundred eighteen aneurysms were small (< 10 mm) and 13 were large (range, 10–25 mm); there were no giant aneurysms in this series. A large neck, defined as ≥4 mm and/or a dome-to-neck ratio of <2, was observed in 93% of cases (Table 1).

Treatment
Treatment was successful in all cases. All procedures were performed using only 1 FRED for each patient. In several patients, a single FRED was used to treat multiple aneurysms: 2, were used in 18 patients; 3, in 5 patients; and 4, in 1 patient. Twelve aneurysms were treated with stent and coil placement. Balloon angioplasty was attempted in 5 cases, but it failed in 2 because the balloon could not be navigated through the entire length of the stent.

Table 1: Aneurysm characteristics ($n = 131$)[a]

Variable	Frequency
Location	
Trunk	125 (95.4%)
Bifurcation	6 (4.6%)
Branch arising from the sac	49 (37.4%)
Topography	
Cavernous ICA	13 (9.9%)
Ophthalmic ICA	89 (67,9%)
Communicating ICA	20 (15,2%)
Carotid bifurcation	2 (1.4%)
AcomA	3 (2.2%)
MCA	1 (0.7%)
Basilar artery	1 (0.7%)
Vertebral artery	2 (1.4%)
Size	
Small	118 (90.1%)
Large	13 (9.9%)
Large neck	
No	9 (6.9%)
Yes	122 (93.1%)
Rupture etiology	1 (0.7%)
Saccular	130 (99.3%)
Dissecting	1 (0.7%)
Prior treatment	4 (3%)

Note:—AcomA indicates anterior communicating artery; ICA, internal carotid artery; MCA, middle cerebral artery.
[a] Values expressed as absolute and relative frequency.

Table 2: Treatment ($n = 131$)[a]

Variable	Frequency
Use of coils	12 (9.2%)
Angioplasty	5 (4.6%)
Antiplatelet (aspirin +)	
Clopidogrel	11 (8.4%)
Prasugrel	89 (67.9%)
Complications	5 (4.6%)
Occlusion at 6 mo[b]	116 (91.3%)
Occlusion at 12 mo[c]	117 (95.9%)

[a] Values expressed as absolute and relative frequencies.
[b] $n = 127$.
[c] $n = 122$.

Primary End Point

Follow-up angiography was performed at 6 months in 116 aneurysms and at 12 months in 117 aneurysms. Complete occlusion rates were 91% and 95%, respectively. There were no cases of recanalization after complete occlusion. The rate of residual aneurysm filling did not change from 6 to 12 months. There was 1 case of asymptomatic carotid stenosis at 6 months. Table 2 shows treatment-related characteristics.

At 6-month follow-up, the likelihood of nonocclusion increased by 161% for every 10-year increase in patient age. Male sex, bifurcation aneurysm, branch involvement, large size, and use of clopidogrel (rather than prasugrel) were associated with higher odds of nonocclusion at 6 months, but these associations were not statistically significant. Branch involvement was positively associated with lower odds of occlusion at 12 months; however, due to the small sample size, the logistic model could not accurately estimate the magnitude of the association (Table 3).

Secondary End Point

Of the 100 patients, 96 had an uneventful postoperative course. One patient who presented with retroperitoneal hemorrhage due to a femoral artery pseudoaneurysm was successfully treated with a covered stent but died of myocardial infarction during hospitalization. In 1 case acute carotid occlusion during stent deployment was successfully treated with tirofiban and balloon angioplasty, but despite complete arterial reperfusion, the patient had a stroke, with an mRS score of 3 at discharge (left hemiparesis). One case of traumatic carotid cavernous fistula occurred when we attempted to cross the aneurysm at the cavernous segment of the internal carotid artery. We decided to deploy the stent as planned and monitor the patient. Because the patient developed chemosis, proptosis, and ophthalmoplegia at early follow-up, a second procedure was performed using a venous approach. The fistula was completely closed, and the patient was asymptomatic. In 1 case, the stent did not open completely, and balloon angioplasty was not feasible. The stent was not fully attached to the vessel wall, and the patient had a minor stroke and developed right-arm paresthesia.

Periprocedural complications occurred in 5 cases. In 1 patient, a misplaced stent was removed and reinserted correctly. Two stents did not open completely at the end of the procedure. One carotid artery occluded and one ruptured, both causing a carotid cavernous fistula.

Asymptomatic parent artery thrombosis was observed in 1 patient at 6-month follow-up. The subsequent course was uneventful, and the patient was compliant with dual-antiplatelet therapy of aspirin and prasugrel. No retreatment was performed.

Most cases in this sample were small aneurysms located at the carotid artery ($n = 112$). The analysis of this subgroup showed 35% of the cases with branches coming out of the aneurysm, 92% with a wide neck, use of coils in 5% of the cases, angioplasty in 3% of the procedures, and occlusion rates at 6 and 12 months of 89%. Age was associated with a higher probability of nonocclusion in 6 months (OR = 2.74 for each 10-year increase in age), and male patients were more likely to not be occluded at 12 months (OR = 8.9 compared with women). Still, there was no statistically significant association among bifurcation site, secondary branch involvement, use of coils, wide neck, adjuvant angioplasty, and antiaggregation scheme in relation to the aneurysm occlusion rate at 6 and 12 months.

DISCUSSION

Flow-diverter stents are a large step forward in the endovascular approach to intracranial aneurysms, shifting the focus from occluding the aneurysm dome to repairing the parent vessel wall. Incomplete coiling and reperfusion are still major limitations that prevent long-term stability of aneurysm occlusion. Recanalization or neck remnants or both may occur despite technologic refinements (such as coated platinum coils) and procedural modifications (such as the balloon-remodeling technique and stent-assisted coil embolization). The recurrence rate can reach up to 28.6%, especially in large and giant aneurysms.[16] Studies have demonstrated the safety and efficacy of flow diverters in the treatment of large and giant sidewall saccular aneurysms of the ICA, with low morbidity and mortality rates and occlusion rates of 80%–90% at

Table 3: Univariate predictors of nonocclusion at 6 months (n = 127) and 12 months (n = 122)

	6 Months		12 Months	
	OR (95% CI)	P	OR (95% CI)	P
Male sex	2.36 (0.5–12.5)	.345	7.13 (1.1–47.8)	.065
Age[a]	2.61 (1.2–5.7)	.004	1.47 (0.6–53.4)	.340
Bifurcation	6.22 (1.0–538.7)	.078	7.06 (0.6–578.4)	.172
Branch vessel aneurysm	3.20 (0.9–511.6)	.069	–	.002
Coil	–	.130	–	.304
Large aneurysm	2.12 (0.4–511.1)	.402	–	.304
Large neck	–	.139	–	.437
Postangioplasty restenosis	–	.389	–	.558
Clopidogrel[b]	2.01 (0.6–57.0)	.282	0.52 (0.1–5 4.8)	.542

Note:—Data express odds ratio (95% confidence interval). P, probability value; –, analysis not performed due to insufficient number of events.
[a] For every 10 years.
[b] Versus prasugrel.

12 months. On the basis of these good results, the indications for flow diverters have been expanded to include aneurysms of different shapes, locations, and sizes.[7,8,17]

New flow diverters made available in recent years have advantages over first-generation devices. Their smaller size is compatible with smaller delivery systems, such as a 0.021-inch microcatheter in the FRED Jr and a 0.017-inch microcatheter in the Silk Vista Baby (Balt Extrusion). The Pipeline Flex Embolization Device with Shield Technology (Medtronic), for example, underwent a surface modification to reduce thromboembolic complications. FRED and FRED Jr are dual-layer flow-diverter devices that have a stentlike outer layer and a flow diverter inside the stent. This design enhances navigation and facilitates deployment. Retrospective case series have demonstrated the safety and efficacy of FRED and FRED Jr, though there are few prospective reports in the literature.

The present series demonstrates the safety and efficacy of FRED and FRED Jr, with success, morbidity, and mortality rates similar to those of previous studies.[12-14] There was only 1 death in our series (0.9%), and the morbidity rate was 1.8% (1 patient with left hemiparesis and another with right-arm paresthesia). All cases were successfully treated. Technical complications occurred in 4.6%, and complete occlusion, in 95% of patients at 12 months. Thromboembolic complications were observed in 2 cases. In the first case, a patient previously treated with a Silk device had an acute carotid occlusion during FRED deployment. Although we do not have a definite explanation for this event, certain factors might be important in this case: The patient may have been clopidogrel-resistant; the patient already had a flow diverter in the target vessel, which may have hindered visualization, placement, and opening of the FRED stent; and the stent failed to open completely. In the second case, the FRED stent was not fully attached to the ICA. Balloon angioplasty was attempted but was not feasible, and the patient had a minor stroke.

The SAFE study reported rates of 2.9% for morbidity, 1.9% for mortality, and 6.8% for thromboembolic complications, with a 1-year complete occlusion rate of 73%.[13] The Italian FRED study, which achieved the highest occlusion rate (97% at 1 year), reported a slightly higher rate of complications: 4.3% for mortality and 7.3% for morbidity.[18] Guimaraens et al,[19] in a prospective study

of 185 intracranial aneurysms treated with the FRED, reported a major complication rate of 6.5%, though only 0.5% were clinically relevant. The complete occlusion rate at 12 months was 84%. In the present series, we found slightly higher complete occlusion rates: 91% at 6 months and 95% at 12 months. Of the 11 aneurysms that were not completely occluded at 6 months, 5 were occluded at 12 months. Of the remaining cases, a branch originated in 4 (66.6%). Although branch involvement was positively associated with lower odds of occlusion at 12 months, the logistic model could not accurately estimate the magnitude of the association because of the small sample size. This was also demonstrated by Trivelato et al,[20] who found that the presence of a branch arising from the sac predicted persistent filling of the aneurysm at 6- and 12-month follow-up. Another factor associated with nonocclusion at 6 months in the present study was that for every 10-year increase in patient age, the likelihood of nonocclusion increased by 161%.

A problematic issue associated with the FRED is the difficulty of crossing the stent with a balloon for angioplasty, which we reported in our initial experience with this stent.[14] Because the FRED has a stent inside a stent, the inner mesh may hinder (and in some cases even preclude) the navigation of the balloon through the entire length of the stent. In the present series, balloon angioplasty was attempted in 5 cases, but it failed in 2, resulting in a clinical deficit in 1 of them. Surprisingly, this problem has not been reported in previously published series, and it is important for physicians to be aware of this. The device should preferably be resheathed to ensure that it will open completely before redeployment because it may not be possible to cross it with a balloon afterward to perform angioplasty. This issue may be an important risk factor for technical failure and clinical complications.

Our results are similar to those of studies using other flow diverters. The International Retrospective Study of the Pipeline Embolization Device reported a morbidity and mortality rate of 8.4% after treatment, highest in the posterior circulation group.[21] Pumar et al,[4] using Silk devices, reported morbidity and mortality rates of 9.6% and 3.2%, respectively, with complete occlusion in 78% at 12 months. The Brazilian Registry of Aneurysms Assigned to Intervention with the Derivo Embolization Device (BRAIDED) study achieved a 1-year complete occlusion rate of 89%, and there were no serious adverse events during follow-up in 94.5% of patients.[20] Trivelato et al,[22] in a series of 182 aneurysms treated with the Pipeline Flex Embolization Device with Shield Technology, reported a periprocedural complication rate of 7.3% and complete occlusion rates of 79.7% at 6 months and 85.3% at 12 months.

CONCLUSIONS

As reported previously, the FRED has proved to be a safe and effective tool for the treatment of intracranial aneurysms, with high occlusion rates at 6 and 12 months. The design of the stent

makes it more difficult to perform balloon angioplasty compared with similar devices, and this feature may lead to clinical complications. A branch arising from the aneurysm sac was found to be a predictor of nonocclusion at 12 months, though larger series are needed to estimate the magnitude of the association.

REFERENCES

1. Molyneux A, Kerr R, Stratton I, et al. **International Subarachnoid Aneurysm Trial (ISAT) Collaborative Group: International Subarachnoid Aneurysm Trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms—a randomised trial.** *Lancet* 2002;360:1267–74 CrossRef Medline

2. Brinjikji W, Rabinstein AA, Nasr DM, et al. **Better outcomes with treatment by coiling relative to clipping of unruptured intracranial aneurysms in the United States, 2001–2008.** *AJNR Am J Neuroradiol* 2011;32:1071–75 CrossRef Medline

3. Fiorella D, Lylyk P, Szikora I, et al. **Curative cerebrovascular reconstruction with the Pipeline Embolization Device: the emergence of definitive endovascular therapy for intracranial aneurysms.** *J Neurointerv Surg* 2009;1:56–65 CrossRef Medline

4. Pumar JM, Banguero A, Cuellar H, et al. **Treatment of intracranial aneurysms with the SILK embolization device in a multicenter study: a retrospective data analysis.** *Neurosurgery* 2017;81:595–601 CrossRef Medline

5. Lylyk P, Miranda C, Ceratto R, et al. **Curative endovascular reconstruction of cerebral aneurysms with the Pipeline embolization device: the Buenos Aires experience.** *Neurosurgery* 2009;64:632–42; discussion 642 CrossRef Medline

6. Becske T, Kallmes DF, Saatci I, et al. **Pipeline for uncoilable or failed aneurysms: results from a multicenter clinical trial.** *Radiology* 2013;267:858–68 CrossRef Medline

7. Pistocchi S, Blanc R, Bartolini B, et al. **Flow diverters at and beyond the level of the circle of Willis for the treatment of intracranial aneurysms.** *Stroke* 2012;43:1032–38 CrossRef Medline

8. Martínez-Galdámez M, Romance A, Vega P, et al. **Pipeline endovascular device for the treatment of intracranial aneurysms at the level of the circle of Willis and beyond: multicenter experience.** *J Neurointerv Surg* 2015;7:816–23 CrossRefMedline

9. Martínez-Galdámez M, Lamin SM, Lagios KG, et al. **Treatment of intracranial aneurysms using the Pipeline Flex Embolization Device with Shield Technology: angiographic and safety outcomes at 1-year follow-up.** *J NeuroInterv Surg* 2019;11:396–99 CrossRef Medline

10. Rautio R, Rahi M, Katila A, et al. **Single-center experience with six-month follow-up of FRED Jr® flow diverters for intracranial aneurysms in small arteries.** *Acta Radiol* 2019;60:917–24 CrossRef Medline

11. Luecking H, Engelhorn T, Lang S, et al. **FRED flow diverter: a study on safety and efficacy in a consecutive group of 50 patients.** *AJNR Am J Neuroradiol* 2017;38:596–602 CrossRef Medline

12. Möhlenbruch MA, Kizilkilic O, Killer-Oberpfalzer M, et al. **Multicenter experience with FRED Jr Flow Re-Direction Endoluminal Device for intracranial aneurysms in small arteries.** *AJNR Am J Neuroradiol* 2017;38:1959–65 CrossRef Medline

13. Pierot L, Spelle L, Berge J, et al. **Feasibility, complications, morbidity, and mortality results at 6 months for aneurysm treatment with the Flow Re-Direction Endoluminal Device: report of SAFE study.** *J Neurointerv Surg* 2018;10:765–70 CrossRef Medline

14. Manzato LB, Santos RB, Teixeira DO, et al. **Initial experience with a Flow Redirection Endoluminal Device stent: a Brazilian multicenter study.** *J Stroke Cerebrovasc Dis* 2018;27:e158–64 CrossRef Medline

15. Dinc H, Saatci I, Oguz S, et al. **Long-term clinical and angiographic follow-up results of the dual-layer flow diverter device (FRED) for the treatment of intracranial aneurysms in a multicenter study.** *Neuroradiology* 2021;63:943–52 CrossRef Medline

16. Piotin M, Spelle L, Mounayer C, et al. **Intracranial aneurysms: treatment with bare platinum coils–aneurysm packing, complex coils, and angiographic recurrence.** *Radiology* 2007;243:500–08 CrossRef Medline

17. Becske T, Brinjikji W, Potts MB, et al. **Long-term clinical and angiographic outcomes following Pipeline embolization device treatment of complex internal carotid artery aneurysms: five-year results of the Pipeline for Uncoilable or Failed Aneurysms Trial.** *Neurosurgery* 2017;80:40–48 CrossRef Medline

18. Piano M, Valvassori L, Lozupone E, et al., FRED Italian Registry Group. **FRED Italian registry: a multicenter experience with the Flow Re-Direction Endoluminal Device for intracranial aneurysms.** *J Neurosurg* 2019 May 10. [Epub ahead of print] CrossRef Medline

19. Guimaraens L, Vivas E, SaldañaJ, et al. **Efficacy and safety of the dual-layer flow-diverting stent (FRED) for the treatment of intracranial aneurysms.** *J Neurointerv Surg* 2020;12:521–25 CrossRef Medline

20. Trivelato FP, Abud DG, Ulhôa AC, et al. **Derivo Embolization Device for the treatment of intracranial aneurysms.** *Stroke* 2019;50:2351–58 CrossRef Medline

21. Kallmes DF, Hanel R, Lopes D, et al. **International retrospective study of the Pipeline embolization device: a multicenter aneurysm treatment study.** *AJNR Am J Neuroradiol* 2015;36:108–15 CrossRef Medline

22. Trivelato FP, Wajnberg E, Rezende MTS, et al. **Safety and effectiveness of the Pipeline Flex Embolization Device with Shield Technology for the treatment of intracranial aneurysms: midterm results from a multicenter study.** *Neurosurgery* 2020;87:104–11 CrossRef Medline

Flow Diversion of Posterior Circulation Aneurysms: Systematic Review of Disaggregated Individual Patient Data

A. Alwakeal, N.A. Shlobin, P. Golnari, W. Metcalf-Doetsch, P. Nazari, S.A. Ansari, M.C. Hurley, D.R. Cantrell, A. Shaibani, B.S. Jahromi, and M.B. Potts

ABSTRACT

BACKGROUND: Experience with endoluminal flow diversion for the treatment of posterior circulation aneurysms is limited.

PURPOSE: We sought to investigate factors associated with the safety and efficacy of this treatment by collecting disaggregated patient-level data from the literature.

DATA SOURCES: PubMed, EMBASE, and Ovid were searched up through 2019 for articles reporting flow diversion of posterior circulation aneurysms.

STUDY SELECTION: Eighty-four studies reported disaggregated data for 301 separate posterior circulation aneurysms.

DATA ANALYSIS: Patient, aneurysm, and treatment factors were collected for each patient. Outcomes included the occurrence of major complications, angiographic occlusion, and functional outcomes based on the mRS.

DATA SYNTHESIS: Significant differences in aneurysm and treatment characteristics were seen among different locations. Major complications occurred in 22%, angiographic occlusion was reported in 65% (11.3 months of mean follow-up), and good functional outcomes (mRS 0–2) were achieved in 67% (13.3 months of mean follow-up). Multivariate analysis identified age, number of flow diverters used, size, and prior treatment to be associated with outcome measures. Meta-analysis combining the current study with prior large nondisaggregated series of posterior circulation aneurysms treated with flow diversion found a pooled incidence of 20% ($n = 712$ patients) major complications and 75% ($n = 581$ patients) angiographic occlusions.

LIMITATIONS: This study design is susceptible to publication bias. Use of antiplatelet therapy was not uniformly reported.

CONCLUSIONS: Endoluminal flow diversion is an important tool in the treatment of posterior circulation aneurysms. Patient age, aneurysm size, prior treatment, and the number of flow diverters used are important factors associated with complications and outcomes.

ABBREVIATIONS: aOR = adjusted OR; FD = flow diverter; PCA = posterior cerebral artery; VBJ = vertebrobasilar junction

Endoluminal flow diversion is a well-established treatment for cerebral aneurysms, but most flow diverters (FDs) have been approved for the treatment of anterior circulation aneurysms.[1] Posterior circulation aneurysms comprise ~10%–15% of all cerebral aneurysms,[2] and some of the first cases of flow diversion in the cerebral circulation involved posterior circulation aneurysms.[3,4] However, the early flow-diversion experience was notable for several reports of complications associated with the treatment of posterior circulation aneurysms,[5,6] resulting in severe morbidity or death.[7] Although less common than anterior circulation aneurysms, aneurysms of the posterior circulation have an increased risk of rupture with respect to size,[8] and the treatment risks are typically higher than in anterior circulation aneurysms regardless of treatment technique.[8,9] Although there have been several single- and multicenter series, reviews, and meta-analyses[10] of flow diversion for posterior circulation aneurysms, such aneurysms comprise a diverse set of morphologies and anatomic configurations that would presumably affect the technical outcomes and safety of flow diversion. Given the relative rarity of posterior circulation aneurysms, we sought to perform a systematic review of disaggregated

individual case data of flow diversion for posterior circulation aneurysms to obtain a large series to investigate specific factors associated with outcome.

MATERIALS AND METHODS

Literature Search

We performed a systematic review in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) statement.[11] We searched PubMed MEDLINE (National Library of Science), EMBASE (Elsevier), and Ovid (Wolters Kluwer) databases in December 2019 using Boolean combinations of search terms associated with the use of flow diverters for cerebral aneurysms and the names of specific endoluminal flow-diverter devices, including the Pipeline Embolization Device (PED; Medtronic), Silk (Balt Extrusion), Flow-Redirection Endoluminal Device (FRED; MicroVention), Surpass Streamline Flow Diverter (Stryker Neurovascular), the p64 Flow-Modulation Device (phenox), Derivo embolization device (Acandis), and Tubridge flow diverter (MicroPort NeuroTech). Full search terms are listed in the Online Supplemental Data. All data bases were searched back to 2008, and the search was limited to articles published in or translated into English. The protocol for this systematic review was not registered.

Full text articles were included in this study if they reported the use of endoluminal flow diversion for the treatment of cerebral aneurysms of the posterior circulation and included disaggregated individual patient-level data for all patients in the series. Posterior circulation aneurysms were defined as involving the vertebral arteries, PICA, basilar artery, and major branches of the basilar artery, including the AICA, superior cerebellar artery (SCA), and posterior cerebral artery (PCA). A distinction was made between extradural and intradural vertebral artery aneurysms. Similarly, distal PICA aneurysms in which the FD was placed within the PICA were distinguished from PICA-origin aneurysms in which the FD was placed in the parent vertebral artery. Basilar perforator artery aneurysms arising distal to the perforator origin were specifically excluded from this analysis because such aneurysms are not amenable to direct flow diversion (ie, placement of an FD within a basilar perforator artery). Additional articles were identified through review of citations in included articles.

Data Collection and Analysis

Data were collected for each individual patient in each included article and included patient demographics (sex, age), presentation (ruptured versus unruptured), aneurysm characteristics (location, size, type), treatment details (prior treatment, type of flow diverter, number of FDs, adjuvant use of coils), occurrence of major complications, and radiographic and functional outcomes. Functional outcomes were recorded using the mRS. When mRS outcomes were not reported but a description of the patient's neurologic function was provided, this was translated into an mRS score. If no functional outcome was provided but the patient underwent imaging follow-up, the patient was noted to be "alive." Major complications were defined as any hemorrhage, infarct, or other complication that resulted in a functionally significant neurologic deficit or death. Unexplained neurologic worsening of ≥2 points on the mRS was also considered a major complication.

Individual cases were excluded if they lacked both angiographic and functional outcome data. Specific aneurysm locations were excluded if we found <5 cases in the literature. In addition, we differentiated proximal-versus-distal PICA aneurysm locations but excluded cases if it was not clear where the flow diverter was placed (ie, within the PICA itself or within the vertebral artery across the origin of the PICA). After full data collection, individual cases were again reviewed to remove any potential duplicates.

Descriptive statistics were performed using R statistical and computing software (http://CRAN.R-project.org). Major complication, angiographic occlusion, and poor neurologic outcome (mRS >2) were each considered as a separate outcome. Comparisons of baseline variables among aneurysm locations were performed in R using analysis of variance for continuous variables and the Fisher exact test for categoric variables. To determine factors associated with these outcomes, we performed uni- and multivariate binary logistic regression analyses with SPSS (IBM) for each outcome using patient (sex, age), aneurysm (presentation, location, type, size), and treatment characteristics (prior treatment, adjunct coiling, number of FDs) to obtain unadjusted ORs and adjusted ORs (aORs), respectively. Meta-analyses of pooled proportions from trials reporting aggregated data of ≥10 patients with posterior circulation aneurysms treated with flow diversion were performed with the metaphor package in R (http://CRAN.R-project.org/package=metafor) using a DerSimonian and Laird random-effects model. Statistical significance was set at $P < .05$.

RESULTS

A systematic review (Online Supplemental Data) of PubMed, EMBASE, and Ovid returned 84 articles reporting disaggregated patient-level data for the use of endoluminal FDs for the treatment of posterior circulation aneurysms meeting the inclusion and exclusion criteria in 301 separate aneurysms (Online Supplemental Data). Fewer than 5 cases of superior cerebellar artery and AICA aneurysms were identified, so the final locations used in this analysis included the extradural vertebral artery ($n = 8$, 2.7%), intradural vertebral artery ($n = 116$, 38.5%), proximal PICA ($n = 20$, 6.6%), distal PICA ($n = 9$, 3.0%), vertebrobasilar junction (VBJ, $n = 53$, 17.6%), basilar trunk ($n = 51$, 16.9%), basilar apex ($n = 17$, 5.6%), and PCA ($n = 27$, 5.6%; Fig 1). Baseline demographic data for the final population and by aneurysm location are presented in the Online Supplemental Data. Significant differences were seen among the various aneurysm locations in nearly every baseline variable measured (Online Supplemental Data). Similarly, rates of major complications and angiographic occlusion as well as functional outcomes differed among aneurysm locations, with the VBJ and basilar trunk aneurysms associated with the highest rates of complications and poor outcomes, while the VBJ, basilar trunk, and proximal PICA aneurysms showed the lowest occlusion rates (Fig 1). Outcomes by demographic, aneurysm, and treatment factors are shown in the Online Supplemental Data.

Major complications were reported in 66 cases (21.9%) and included hemorrhagic (12 cases) and ischemic events (33 cases), symptomatic mass effect (5 cases), unspecified neurologic worsening (5 cases), and mortalities attributed to subarachnoid hemorrhage (7 cases), premorbid status (2 cases), and other medical comorbidities (2 cases, Online Supplemental Data). Of the 12
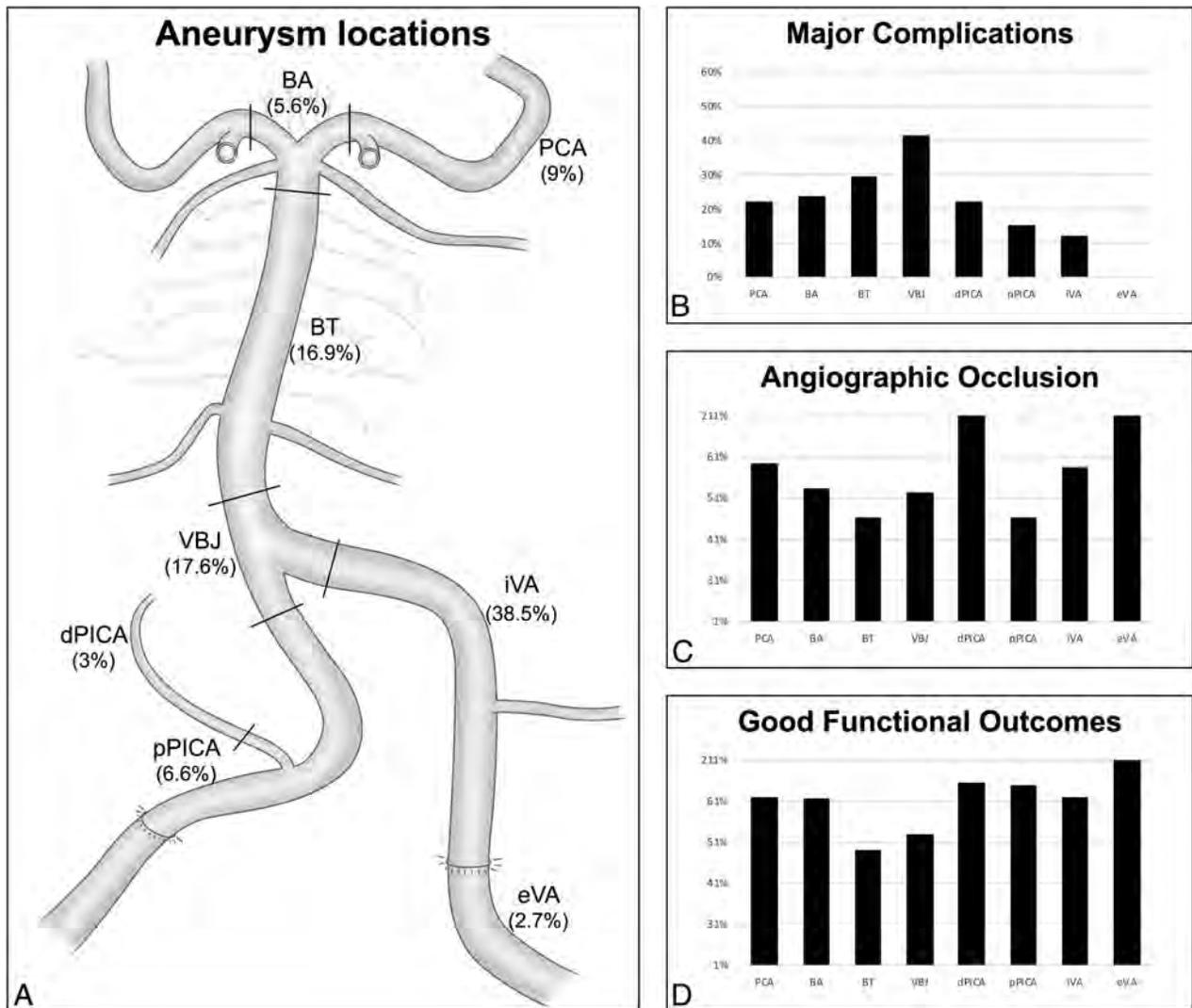
FIG 1. *A*, Final posterior circulation aneurysm locations included the extradural vertebral artery (eVA), intradural vertebral artery (iVA), proximal PICA (pPICA), distal PICA (dPICA), VBJ, basilar trunk (BT), basilar apex (BA), and the PCA. *B*, Proportion of major complications by aneurysm location. *C*, Proportion of angiographic occlusions by aneurysm location. *D*, Proportion of good functional outcomes (defined as mRS 0–2) by aneurysm location.

hemorrhagic complications, 2 were specifically attributed to aneurysm rerupture, one of which occurred after treatment of a giant fusiform basilar trunk aneurysm and the other after treatment of a large proximal PICA aneurysm. One hemorrhage occurred in a delayed fashion after treatment of a distal PICA aneurysm, while another reported hemorrhage was related to external ventricular drain placement in the setting of a ruptured aneurysm (Online Supplemental Data). Of the 33 ischemic complications, 12 were specifically related to in-stent thrombosis. Seven of these occurred intra- or periprocedurally, while the remainder occurred in a delayed fashion, ranging from 4.5 months to 4 years after flow-diversion treatment (Online Supplemental Data). In univariate analysis (Online Supplemental Data), age ($P = .01$), aneurysm location ($P = .009$), type ($P = .008$), size ($P = .006$), and the number of FDs 9 ($P < .001$) were all associated with the occurrence of major complications. Multivariate logistic regression (Online Supplemental Data) found only increasing age ($P = .006$) and the number of FDs ($P = .003$)

to be significantly associated with a major complication, with increasing age and increasing number of FDs associated with higher aORs (1.04 and 1.54, respectively). Of note, rupture status was not associated with the occurrence of major complications.

Angiographic follow-up was reported in 288 cases (95.7%), with a mean angiographic follow-up time of 11.3 months (range, 1 day to 65 months; length of angiographic follow-up was not reported in 16%). Overall, complete aneurysm occlusion was reported in 196 cases (65.1%). Univariate analysis (Online Supplemental Data) identified age ($P = .004$), location ($P = .08$), size ($P < .001$), the use of adjunct coils ($P = .02$), and the number of FDs ($P = .03$) to be associated with angiographic outcomes. Although the omnibus test results for aneurysm types were nonsignificant, dissecting aneurysms were also found to have a significantly increased likelihood of occlusion compared with saccular aneurysms (OR = 2.23, $P = .02$). Multivariate logistic regression (Online Supplemental Data) found that only age ($P = .001$) and size ($P = .02$) remained significant predictors of angiographic

modified Rankin Scale Score

☐ 0 ▨ 1 ▨ 2 ▨ 3 ▨ 4 ▨ 5 ■ 6 ▨ Alive*
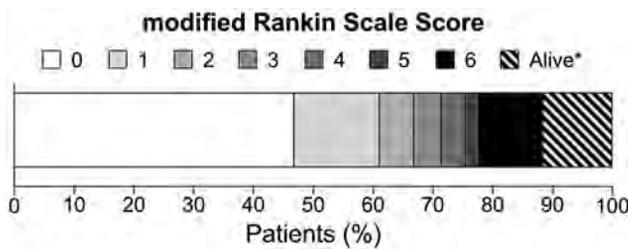
FIG 2. Overall long-term neurologic outcomes based on the mRS score. The *asterisk* indicates patients who did not have a reported neurologic outcome but had angiographic follow-up and were designated as alive.

occlusion in this cohort, with older age and giant aneurysms associated with decreased aneurysm occlusion (aOR = 0.96 and 0.19, respectively).

The mean time to clinical follow-up was 13.3 months (range, 1 day to 65 months; the length of the clinical follow-up time was not reported in 23.9%). Good functional outcomes (mRS 0–2) were reported in 201 cases (66.8%), and the overall mortality in this series was 10.6% (Fig 2). Thirty-five cases (11.6%) did not report on neurologic status but patients did have angiographic follow-up and were, therefore, designated as alive at the time of clinical follow-up and were not included in the outcome analysis. Univariate analysis (Online Supplemental Data) identified rupture status ($P = .03$); aneurysm location ($P = .02$); type ($P = .02$); size ($P = .001$); prior treatment ($P = .01$); and the number of FDs ($P < .001$) as factors associated with a dichotomized functional outcome (mRS 0–2 versus >2). Multivariate logistic regression (Online Supplemental Data) found that younger age ($P = .004$) and fewer FDs ($P < .001$) remained significantly associated with good functional outcome (aOR = 0.96 and 0.53, respectively). In addition, prior treatment significantly increased the likelihood of a good functional outcome (aOR = 3.7, $P = .04$). Rupture status was not significant in this analysis (aOR = 0.38, $P = .07$).

Meta-analyses were performed combining data from this study with other nonoverlapping studies reporting aggregated data of ≥10 cases of posterior circulation aneurysms treated with flow diversion.[12-18] Among 712 patients, the pooled incidence of major complications after treatment of posterior circulation aneurysms with flow diversion was 19.6% (95% CI, 15.3%–23.9%) with heterogeneity $I^2 = 42.9\%$ (Fig 3). A similar meta-analysis of angiographic outcomes in 581 patients found that complete aneurysm occlusion was achieved in 75.2% of cases (95% CI, 66.8%–83.6%) with heterogeneity $I^2 = 77.9\%$ (Fig 3).

DISCUSSION

The experience of flow diversion for the treatment of posterior circulation aneurysms is relatively limited compared with the anterior circulation. For example, the most recent published meta-analysis by Liang et al[10] combined 12 studies comprising 358 posterior circulation aneurysms treated with the PED to yield an 82% obliteration rate with an 18% complication rate. Posterior circulation aneurysms, however, represent an anatomically diverse group of aneurysms, with differing morphologic subtypes and technical challenges. To further investigate factors associated

with the safety and efficacy of this treatment in posterior circulation aneurysms, we performed a systematic review of the literature to collect disaggregated, patient-level data and identified 301 individual cases of posterior circulation aneurysms treated with endoluminal flow diversion. This is the largest series of case-reported data to date, permitting more granular analysis of variables associated with complications, occlusion rates, and outcomes. Our findings reveal complications in 22%, angiographic occlusion in 65%, and good outcomes in 67%. Age, aneurysm size, prior treatment, and the number of flow diverters were significantly and independently associated with clinical and radiographic outcomes. We then performed a meta-analysis of prior large (≥10 patients) series,[12-18] demonstrating that our results are in line with prior aggregated studies and revealing an overall complication rate of 19.6% and a 75.2% occlusion rate (Fig 2).

Prior studies have shown that flow diversion in the posterior circulation can be performed successfully, albeit with higher inherent risks than in the anterior circulation. Several factors associated with complications and outcomes have been identified, including size, morphology, and the involvement of critical perforators.[19] As part of their meta-analysis, Liang et al[10] performed a meta-regression analysis and found that age was the only identifiable factor associated with angiographic outcomes. They found no factors associated with complications. The largest study in that meta-analysis was by Griessenauer et al,[14] who reported a retrospective multicenter series of 129 patients with posterior circulation aneurysms treated with the PED. They analyzed aneurysms separately by morphology (saccular, fusiform, and dissecting) and found very few factors associated with angiographic outcomes or complications. In fact, the only significant variable found on multivariate analysis was clopidogrel responsiveness, which was associated with complications in saccular aneurysms.

More recently, Lopes et al[16] reported a subgroup analysis of posterior circulation aneurysms from the International Retrospective Study of Pipeline Embolization Device (IntrePED) trial, a multicenter retrospective PED registry. They identified 91 patients with a mean follow-up of 21 months. Multivariate analysis identified fusiform morphology and the use of ≥3 PEDs as factors associated with combined neurologic morbidity and mortality. Aneurysm size was separately associated with morbidity, while the use of ≥3 PEDs, rupture status, and age were associated with mortality. Our data are consistent with these prior studies in that multivariate analysis identified only a few factors associated with outcomes. Increasing age was uniformly associated with higher procedural complications and poorer angiographic and functional outcomes, and advanced age has been associated with increased complications with interventional procedures in general.[20] Similar to findings in the IntrePED subgroup analysis,[16] multiple FDs were associated with increased complications and poor functional outcomes in our systematic review. Deployment of multiple FDs increases the technical complexity of an intervention and has been associated with a higher risk of complications in prior studies,[21] though a more recent analysis of this topic found no difference in complications with increasing PEDs.[22] The important association of multiple flow diverters with both complication rates and outcomes on multivariable analysis may indicate that this factor is a more relevant marker of aneurysm complexity than size, location, or morphology alone. However, these

Major Complications

A

Aneurysm Occlusion

C

B

D
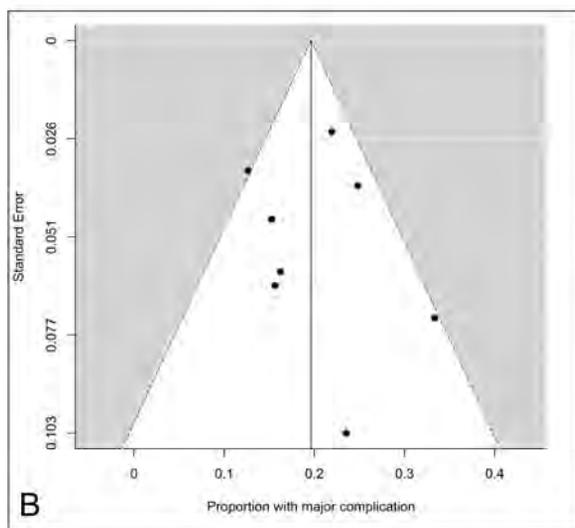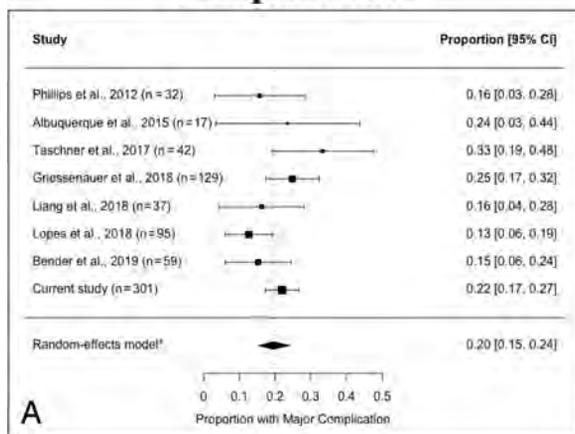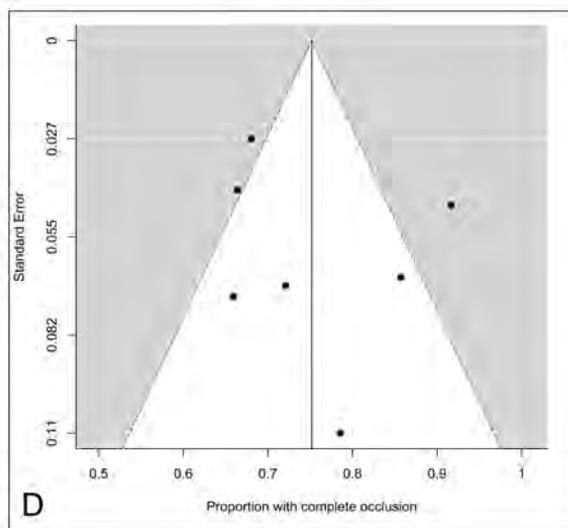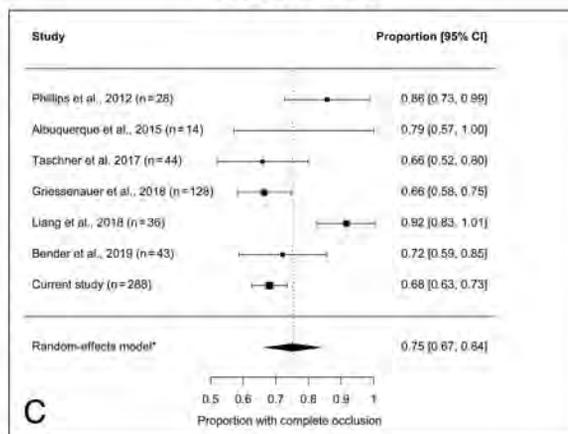
FIG 3. Meta-analysis of studies reporting flow diversion for posterior circulation aneurysms with ≥10 patients. *A*, Forest plot comparing the proportion of patients in each study with a reported major complication. The pooled incidence of major complications among 712 patients was 19.6%, with a heterogeneity $I^2 = 42.9\%$. *B*, Funnel plot of the included studies reporting major complications. *C*, Forest plot comparing angiographic aneurysm occlusion for the included studies. The pooled incidence of complete angiographic occlusion among 594 patients was 75.2%, with a heterogeneity $I^2 = 77.9\%$. *D*, Funnel plot of the included studies reporting angiographic outcomes.

findings should not discount the importance of size, location, and morphology, and additional studies will be required to further investigate these factors.

Most important, while we found no significant differences in outcomes among aneurysms of different locations on multivariable analysis, significant differences in complication rates and outcomes across aneurysm locations were found on univariate analysis (Online Supplemental Data and Fig 1). These univariate differences highlight the different pathologies, anatomy, and technical challenges of treatment with flow diversion predicted by each posterior circulation aneurysm subtype. For example, VBJ and basilar trunk aneurysms have poor natural histories[23,24] but also pose significant treatment challenges regardless of treatment technique.[24-26] In our series, these aneurysms had the highest proportions of giant and fusiform morphologies and were treated with the greatest number of FDs. They, similarly, had the poorest outcomes (Fig 1).

Conversely, vertebral artery aneurysms have a relatively benign natural history,[27] and endovascular treatment is known to be safe and effective.[28] Other unique challenges encountered in the posterior circulation include the bifurcation configuration of basilar apex aneurysms, which were the most likely to require adjunct coiling in our series, and distal locations of PICA and PCA aneurysms.

In the United States, the use of the PED for posterior circulation aneurysms remains off-label, though this device was used in most cases (85% in our series). Similarly, the FRED is approved in the United States for use with wide-neck and fusiform internal carotid artery aneurysms from the petrous segment to the terminus. The Surpass Streamline Flow Diverter has approval for the treatment of unruptured large and giant wide-neck intracranial aneurysms, including those in the posterior circulation. Other flow-diverting devices, including the p64, Derivo, and Tubridge are not yet approved for use in the United States. The decision to

use flow diversion to treat posterior circulation aneurysms must be made in the context of other possible treatments. Before the widespread use of flow diversion, several studies evaluated the safety of surgical management,[26,29,30] which includes direct clipping or reconstruction, as well as proximal occlusion, bypass occlusion, and wrapping. Reported favorable neurologic outcomes range from 55% to 90%, depending on location and series.[26] Before flow diversion, endovascular options included primary coil embolization, stent- or balloon-assisted coil embolization, and vessel sacrifice. Favorable outcomes for endovascular treatment similarly range from 67% to 100%, again depending on location.[26] The ultimate choice of treatment technique must be based on individual anatomic, patient, and treating physician factors. Off-label use of flow diversion should be discussed with patients, even if it is considered the optimal treatment method.

We acknowledge several limitations to this study design. First and foremost, this systematic review identified data from case reports or small series and is, therefore, susceptible to publication bias. Case reports are often published to describe a novel approach, an unusual disease presentation, a notable complication, or an exceptional treatment outcome. Given the relative rarity of flow diversion for posterior circulation aneurysms and the novelty and controversy of this treatment in this population, it is reasonable to think that a substantial proportion of such cases have led to publication. A compilation of published case reports and series of flow diversion for posterior circulation aneurysms may, therefore, be a reasonable reflection of the overall population. This is supported by our meta-analysis (Fig 3), which shows similar findings in our series compared with prior large series of a similar patient population. Second, FD use has evolved across time with different generations of devices and increasing experience and skill— factors that could not be analyzed with our study design. Finally, antiplatelet therapy is a critical component of flow diversion. Antiplatelet responsiveness has been associated with complications in flow diversion of cerebral aneurysms— nonresponsiveness is associated with thrombotic complications, while hyper-responsiveness is associated with hemorrhagic complications.[31] This association has also been specifically demonstrated with flow diversion of posterior circulation aneurysms.[14] Unfortunately, our study design did not allow uniform assessment of antiplatelet management. This critical factor, however, warrants further study.

CONCLUSIONS

Endoluminal flow diversion is an important tool in the treatment of posterior circulation aneurysms. Although such aneurysms comprise a diverse set of anatomic configurations, these data suggest that increasing age and the use of multiple FDs are critical factors associated with procedural complications and neurologic outcomes, outweighing aneurysm location on multivariate analysis. Continued evaluation of endoluminal flow diversion in the treatment of posterior circulation aneurysms is warranted.

REFERENCES

1. Chancellor B, Raz E, Shapiro M, et al. **Flow diversion for intracranial aneurysm treatment: trials involving flow diverters and long-term outcomes.** *Neurosurgery* 2020;86:S36–45 CrossRef Medline

2. International Study of Unruptured Intracranial Aneurysms Investigators. **Unruptured intracranial aneurysms: risk of rupture and risks of surgical intervention.** *N Engl J Med* 1998;339:1725–33 CrossRef Medline

3. Pumar JM, Garcia-Dorrego R, Nieto A, et al. **Vascular reconstruction of a fusiform basilar aneurysm with the Silk embolization system.** *J Neurointerv Surg* 2010;2:242–44 CrossRef Medline

4. Fiorella D, Kelly ME, Albuquerque FC, et al. **Curative reconstruction of a giant midbasilar trunk aneurysm with the Pipeline Embolization Device.** *Neurosurgery* 2009;64:212–17; discussion 217 CrossRef Medline

5. Klisch J, Turk A, Turner R, et al. **Very late thrombosis of flow-diverting constructs after the treatment of large fusiform posterior circulation aneurysms.** *AJNR Am J Neuroradiol* 2011;32:627–32 CrossRef Medline

6. Fiorella D, Hsu D, Woo HH, et al. **Very late thrombosis of a Pipeline Embolization Device construct: case report.** *Neurosurgery* 2010;67:onsE313–14; discussion onsE314 CrossRef Medline

7. Siddiqui AH, Abla AA, Kan P, et al. **Panacea or problem: flow diverters in the treatment of symptomatic large or giant fusiform vertebrobasilar aneurysms.** *J Neurosurg* 2012;116:1258–66 CrossRef Medline

8. Wiebers DO, Whisnant JP, Huston J 3rd et al. **Unruptured intracranial aneurysms: natural history, clinical outcome, and risks of surgical and endovascular treatment.** *Lancet* 2003;362:103–10 CrossRef Medline

9. Raaymakers TW, Rinkel GJ, Limburg M, et al. **Mortality and morbidity of surgery for unruptured intracranial aneurysms: a meta-analysis.** *Stroke* 1998;29:1531–38 CrossRef Medline

10. Liang F, Zhang Y, Yan P, et al. **Outcomes and complications after the use of the Pipeline Embolization Device in the treatment of intracranial aneurysms of the posterior circulation: a systematic review and meta-analysis.** *World Neurosurg* 2019;127:e888–95 CrossRef Medline

11. Moher D, Liberati A, Tetzlaff J, et al. The PRISMA Group. **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.** *PLoS Med* 2009;6:e1000097 CrossRef Medline

12. Albuquerque FC, Park MS, Abla AA, et al. **A reappraisal of the Pipeline Embolization Device for the treatment of posterior circulation aneurysms.** *J Neurointerv Surg* 2015;7:641–45 CrossRef Medline

13. Bender MT, Colby GP, Jiang B, et al. **Flow diversion of posterior circulation cerebral aneurysms: a single-institution series of 59 cases.** *Neurosurgery* 2019;84:206–16 CrossRef Medline

14. Griessenauer CJ, Ogilvy CS, Adeeb N, et al. **Pipeline embolization of posterior circulation aneurysms: a multicenter study of 131 aneurysms.** *J Neurosurg* 2018;130:923–35 CrossRef Medline

15. Liang F, Zhang Y, Guo F, et al. **Use of Pipeline Embolization Device for posterior circulation aneurysms: single-center experiences with comparison with anterior circulation aneurysms.** *World Neurosurg* 2018;112:e683–90 CrossRef Medline

16. Lopes DK, Jang DK, Cekirge S, et al. **Morbidity and mortality in patients with posterior circulation aneurysms treated with the Pipeline embolization device: a subgroup analysis of the international retrospective study of the Pipeline Embolization Device.** *Neurosurgery* 2018;83:488–500 CrossRef Medline

17. Phillips TJ, Wenderoth JD, Phatouros CC, et al. **Safety of the Pipeline Embolization Device in treatment of posterior circulation aneurysms.** *AJNR Am J Neuroradiol* 2012;33:1225–31 CrossRef Medline

18. Taschner CA, Vedantham S, de Vries J, et al. **Surpass flow diverter for treatment of posterior circulation aneurysms.** *AJNR Am J Neuroradiol* 2017;38:582–89 CrossRef Medline

19. Adeeb N, Ogilvy CS, Griessenauer CJ, et al. **Expanding the indications for flow diversion: treatment of posterior circulation aneurysms.** *Neurosurgery* 2020;86:S76–84 CrossRef Medline

20. Khosla A, Brinjikji W, Cloft H, et al. **Age-related complications following endovascular treatment of unruptured intracranial aneurysms.** *AJNR Am J Neuroradiol* 2012;33:953–57 CrossRef Medline

21. Chalouhi N, Tjoumakaris S, Phillips JL, et al. **A single Pipeline Embolization Device is sufficient for treatment of intracranial aneurysms.** *AJNR Am J Neuroradiol* 2014;35:1562–66 CrossRef Medline

22. Waqas M, Vakharia K, Gong AD, et al. **One and done? The effect of number of Pipeline embolization devices on aneurysm treatment outcomes.** *Interv Neuroradiol* 2020;26:147–55 CrossRef Medline

23. Kobayashi N, Murayama Y, Yuki I, et al. **Natural course of dissecting vertebrobasilar artery aneurysms without stroke.** *AJNR Am J Neuroradiol* 2014;35:1371–75 CrossRef Medline

24. Saliou G, Sacho RH, Power S, et al. **Natural history and management of basilar trunk artery aneurysms.** *Stroke* 2015;46:948–53 CrossRef Medline

25. Sonmez O, Brinjikji W, Murad MH, et al. **Deconstructive and reconstructive techniques in treatment of vertebrobasilar dissecting aneurysms: a systematic review and meta-analysis.** *AJNR Am J Neuroradiol* 2015;36:1293–98 CrossRef Medline

26. Sanai N, Tarapore P, Lee AC, et al. **The current role of microsurgery for posterior circulation aneurysms: a selective approach in the endovascular era.** *Neurosurgery* 2008;62:1236–49; discussion 1249–53 CrossRef Medline

27. Kim BM, Kim SH, Kim DI, et al. **Outcomes and prognostic factors of intracranial unruptured vertebrobasilar artery dissection.** *Neurology* 2011;76:1735–41 CrossRef Medline

28. Guan J, Li G, Kong X, et al. **Endovascular treatment for ruptured and unruptured vertebral artery dissecting aneurysms: a meta-analysis.** *J Neurointerv Surg* 2017;9:558–63 CrossRef Medline

29. Drake CG, Peerless SJ. **Giant fusiform intracranial aneurysms: review of 120 patients treated surgically from 1965 to 1992.** *J Neurosurg* 1997;87:141–62 CrossRef Medline

30. Coert BA, Chang SD, Do HM, et al. **Surgical and endovascular management of symptomatic posterior circulation fusiform aneurysms.** *J Neurosurg* 2007;106:855–65 CrossRef Medline

31. Tonetti DA, Jankowitz BT, Gross BA. **Antiplatelet therapy in flow diversion.** *Neurosurgery* 2020;86:S47–52 CrossRef Medline

Patient-Relevant Deficits Dictate Endovascular Thrombectomy Decision-Making in Patients with Low NIHSS Scores with Medium-Vessel Occlusion Stroke

R. McDonough, P. Cimflova, N. Kashani, J.M. Ospel, M. Kappelhof, N. Singh, A. Sehgal, N. Sakai, J. Fiehler, M. Chen, and M. Goyal

ABSTRACT

BACKGROUND AND PURPOSE: There is a paucity of evidence regarding the safety of endovascular treatment for patients with acute ischemic stroke due to primary medium-vessel occlusion. The aim of this study was to examine the willingness among stroke physicians to perform endovascular treatment in patients with mild-yet-disabling deficits due to medium-vessel occlusion.

MATERIALS AND METHODS: In an international cross-sectional survey consisting of 7 primary medium-vessel occlusion case scenarios, participants were asked whether the presence of personally disabling deficits would influence their decision-making for endovascular treatment despite the patients having low NIHSS scores (<6). Decision rates were calculated on the basis of physician characteristics. Univariable logistic regression clustered by respondent and scenario identity was performed.

RESULTS: Three hundred sixty-six participants from 44 countries provided 2562 answers to the 7 medium-vessel occlusion scenarios included in this study. In scenarios in which the deficit was relevant to the patient's profession, 56.9% of respondents opted to perform immediate endovascular treatment compared with 41.0% when no information regarding the patient's profession was provided (risk ratio = 1.39, $P < .001$). The largest effect sizes were seen for female participants (risk ratio = 1.68; 95% CI, 1.35–2.09), participants older than 60 years of age (risk ratio = 1.61; 95% CI, 1.23–2.10), those with more experience in neurointervention (risk ratio = 1.60; 95% CI, 1.24–2.06), and those who personally performed >100 endovascular treatments per year (risk ratio = 1.63; 95% CI, 1.22–2.17).

CONCLUSIONS: The presence of a patient-relevant deficit in low-NIHSS acute ischemic stroke due to medium-vessel occlusion is an important factor for endovascular treatment decision-making. This may have relevance for the conduct and interpretation of low-NIHSS endovascular treatment in randomized trials.

ABBREVIATIONS: AIS = acute ischemic stroke; EVT = endovascular therapy; RR = risk ratio; MeVO = medium-vessel occlusion

In 2015, five randomized controlled trials demonstrated the safety and efficacy of endovascular therapy (EVT) for acute ischemic stroke (AIS) due to large-vessel occlusion.[1] The observed superiority over best medical management, however, partially resulted from strict inclusion criteria, including restriction of randomization to patients with more severe strokes, as determined by the NIHSS score. As a result, there is no consensus on whether EVT should be offered to patients with mild stroke symptoms (NIHSS score <6). While clinical outcomes of conservatively treated patients with low-NIHSS AIS are more favorable compared with those with larger deficits, approximately 27%–35% of these patients are either functionally dependent at discharge or deceased,[2-4] partly due to the risk of early neurologic deterioration (12%–15%)[5,6] without EVT. Furthermore, not all low-NIHSS strokes are necessarily mild from the patient's perspective; for example, an NIHSS score of 2 due to complete hemianopia would be devastating to a truck driver. The current literature

presents conflicting results regarding the safety and efficacy of EVT in patients with low NIHSS scores with large-vessel occlusion, and multiple randomized trials are presently underway (Endovascular Therapy for Low NIHSS Ischemic Strokes [ENDOLOW], ClinicalTrials.gov identifier: NCT04167527; InExtremis/Minor Stroke Therapy Evaluation [MOSTE], ClinicalTrials.gov identifier: NCT03796468), which will hopefully provide a clearer picture of the optimal treatment strategy.[7,8]

At the same time, EVT for other underrepresented trial subgroups such as those with medium-vessel occlusions (MeVOs) is being discussed. MeVOs, defined as occlusions of the M2, M3, A2, A3, P2, or P3 segment with disabling deficits,[9,10] account for approximately 25%–40% of all AIS.[11] Current recanalization rates with intravenous alteplase range from 21% to 43%, with only approximately half of these patients achieving excellent outcome.[12] This result underscores the need for a more effective treatment strategy, with EVT being the most obvious choice.

In an effort to understand current clinician perspectives concerning the management of MeVO stroke, a case-based survey (MeVO–Finding Rationales and Objectifying New Targets for IntervEntional Revascularization in Stroke [MeVO-FRONTIERS]) was conducted. We sought to determine how the interplay of patient profession, physician and patient characteristics, and the specific nature of "minor" deficits influences physicians' EVT decision-making in MeVO stroke.

MATERIALS AND METHODS
Survey Design
An international online, cross-sectional, anonymous survey (MeVO FRONTIERS) using Qualtrics (www.qualtrics.com) was conducted among stroke physicians to understand their current treatment practice and endovascular decision-making in AIS caused by MeVO. The survey consisted of 7 cases with clinical descriptions and illustrative images. The cases were subdivided into 3–6 consecutive scenarios in which 1 key feature was changed. Respondents were asked whether and how they would treat the described patient. The survey took approximately 30 minutes to complete, and participants were required to answer each question before moving on ("forced response"). Response data were obtained from November 12, 2020 to December 31, 2020. Approval by the local research ethics board (The Conjoint Health Research Ethics Board [CHREB] of the University of Calgary) was obtained. Data used in the current study are available from the author on reasonable request.

Survey Participants
Approximately 1400 stroke physicians (neurologists, interventional neurologists and neuroradiologists, neurosurgeons, and other physicians directly involved in acute stroke care) from 44 countries were invited to participate in this Web-based survey. There were no restrictions with regard to case volume, experience, or academic-versus-nonacademic centers. Before accessing the case scenarios, physicians were asked to provide the following personal data: age, sex, years of stroke-treatment experience, annual personal and center stroke treatment volumes, geographic region, subspecialty, and hospital setting.

Clinical Case Scenarios
Overall, the survey comprised of 7 MeVO narrative cases (4 primary MeVOs and 3 secondary MeVOs) with 3–6 clinical case vignettes each. Noncontrast head CT and/or digital subtraction angiography images were provided for each case. The accompanying clinical case vignettes provided patient age, profession, medical history, stroke severity and imaging details (ASPECTS), location of the occlusion, and, if relevant, CT perfusion–based core and penumbral volumes. For each combination of images and clinical factors, participants were asked to state whether they would treat the patient with EVT. One key piece of information was then changed in each subsequent case vignette (eg, age, stroke severity, ischemic core volume), to assess changes in decision-making based on these items. The current study was restricted to cases in which the patients in question had low NIHSS scores (<6). For detailed descriptions of all case-scenarios that were included in this study see the Online Supplemental Data.

Statistical Analysis
Case scenarios were subdivided into cases in which a profession was listed (ie, in which a person-relevant deficit that would substantially impair the patient in his or her profession was present) and those without any information of this nature. Treatment decisions were binarized into immediate EVT yes/no, the latter including the options "no EVT" and "EVT only if the patient worsens." Participant baseline characteristics and response data were analyzed using descriptive statistics, and group differences in the willingness to endovascularly treat patients with mild MeVO strokes were assessed using the χ^2 test. Univariable regression of the effect of a person-relevant deficit on EVT decision-making was performed, stratified by participant baseline characteristics (respondent's age and sex, years of experience, personal and center stroke-treatment volume per year, and hospital setting and specialty) and clustered for participant and scenario identity. To control for multiple testing, we performed subgroup interaction analyses for profession and each of the respondent-variable subgroups. Risk ratios (RRs) derived from binary logistic regression are reported. P values $<.05$ were considered statistically significant. Only completed surveys were included in the analysis. Data analysis was performed in STATA 16.1 (StataCorp), and figures were created with Power BI Desktop 2016 and the Mapbox Visual plugin (Microsoft).

RESULTS
Respondent Demographics
A total of 366 physicians (84.2% male) from 44 countries completed the survey. Most (170, 46.5%) were interventional neuroradiologists, followed by neurologists (97, 26.5%), neurosurgeons (39, 10.7%), interventional neurologists (36, 9.8%), interventional radiologists (18, 4.9%), and other physicians involved in acute stroke care (6, 1.6%). A total of 2562 responses were obtained for the primary MeVO cases with low NIHSS scores. Detailed participant characteristics are listed in the Online Supplemental Data.

EVT in Low-NIHSS MeVO
Figure 1 shows the main results of the study. In scenarios in which the occlusion-related deficit was relevant to the patient's
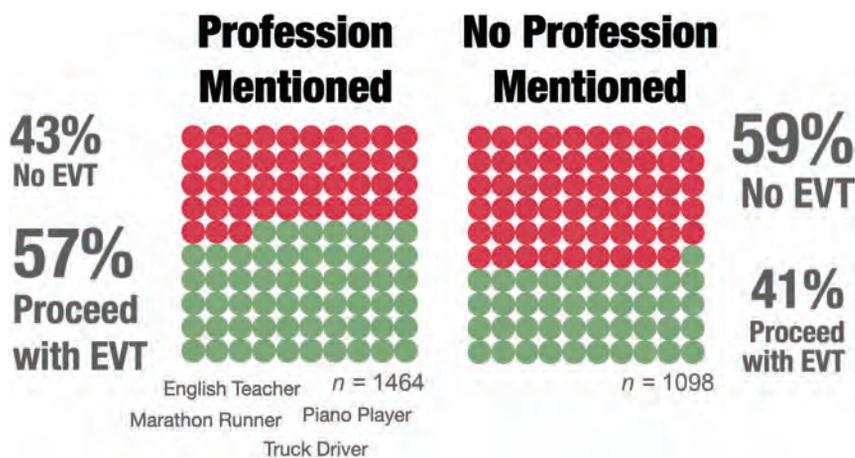
FIG 1. Willingness to perform immediate EVT in patients with NIHSS < 6, stratified by patients with mild-yet-disabling deficits with respect to their professions versus patients with mild deficits without a described profession.

profession, respondents were more likely to perform immediate EVT (56.9% in favor of EVT, 833/1464) compared with those case vignettes in which no information regarding the patient's profession was provided (41.0% in favor of EVT, 450/1098; RR = 1.39, $P < .001$.

Physician Factors Influencing EVT Decisions in Low-NIHSS MeVO

Subgroup analyses of each variable revealed no significant interactions between profession and each of the respondent-variable subgroups on EVT decision-making (Online Supplemental Data); the presented results are, therefore, exploratory in nature. In almost every subgroup of physician characteristics, the presence of a deficit relevant to the individual's livelihood led to a significantly increased chance of proceeding with immediate EVT (Online Supplemental Data). The most prominently influencing factors appeared to be sex, age, experience, and personal EVT treatment volume per year. Female participants were more likely to choose immediate EVT (RR = 1.68; 95% CI, 1.35–2.09; $P < .001$), as were those older than 60 years of age (RR = 1.61; 95% CI, 1.23–2.10; $P < .001$). Physicians at either end of the spectrum with respect to years of experience were the most likely to select immediate EVT for patients with a described profession (0–5 years of experience: RR = 1.49; 95% CI, 1.19–1.86; $P < .001$; >20 years of experience: RR = 1.60; 95% CI, 1.24–2.06; $P < .001$). Annual stroke-treatment volume also played a role, particularly for participants performing >100 procedures per year (RR = 1.63; 95% CI, 1.22–2.17; $P < .001$).

Within the no EVT responses, most physicians chose no EVT rather than performing EVT only if the patient worsens, irrespective of the presence of a patient-relevant deficit (54.6% versus 45.4%, $P = .57$).

Regional Factors Influencing EVT Decisions in Low-NIHSS MeVO

EVT decision rates differed significantly by the respondents' geographic location (Fig 2). In general, participants from Europe and

the rest of the world (comprising Africa, Asia and the Pacific, the Middle East, and South and Latin America) were more willing to treat patients with low NIHSS scores, regardless of the patient's profession (716/1253, 57.1%, and 304/644, 47.2%, respectively), compared with physicians working in North America (263/665, 39.6%). Indeed, the North American group was the only one still less likely to perform EVT in the presence of a patient-relevant deficit, with 44.5% (169/380) for EVT.

DISCUSSION

In this survey-based study, we found that physicians were approximately 50% more willing to treat mild-deficit MeVO strokes with EVT if the deficits were personally disabling with respect to a patient's profession (56.9 versus 41.0%). Our results demonstrate the importance of individual patient factors for EVT decision-making in AIS. The largest effect sizes were seen for female participants, as well as those who were older than 60 years of age, more experienced in neurointervention (>20 years), and having a higher annual personal EVT case volume (>100). The latter findings are in some ways consistent: Physicians who are older, more experienced, and regularly perform EVT are more familiar with novel devices and techniques and are more likely to have encountered cases with low-NIHSS MeVO in their clinical practice. As a result, they may be more confident in their skills with respect to such cases. The fact that women were more likely to opt for EVT in this particular population could be due to a number of reasons: Female respondents comprised only 15% of the participant population, possibly inflating the results, and most (72%) were from high-volume centers (>100 EVTs/year). Thus, the female participants may have also had more experience with such scenarios.

The results of this survey highlight the interplay of factors currently relevant to stroke-treatment decision-making. While a previous study demonstrated the willingness of physicians to pursue EVT in patients with low-NIHSS strokes due to large-vessel occlusion,[13] the consensus becomes more vague for MeVOs. Providing clarity for this patient subgroup is important for several reasons: AIS due to primary MeVO is commonly assumed to be milder than cases due to large-vessel occlusion.[12] Furthermore, MeVOs tend to result in a more heterogeneous clinical presentation.[14] At the same time, the current rates of excellent outcome after best medical management in AIS due to M2 and M3 occlusions are moderate, ranging from 21% to 55%.[15-17]

Finally, not all low NIHSS scores are equal, with varying outcomes observed for such AIS patient groups.[4] A study that looked at differences in outcome between 2 syndromic low-NIHSS patient collectives (level of consciousness and language versus right motor function) found that those in the latter category had lower morbidity and mortality rates, suggesting that identical
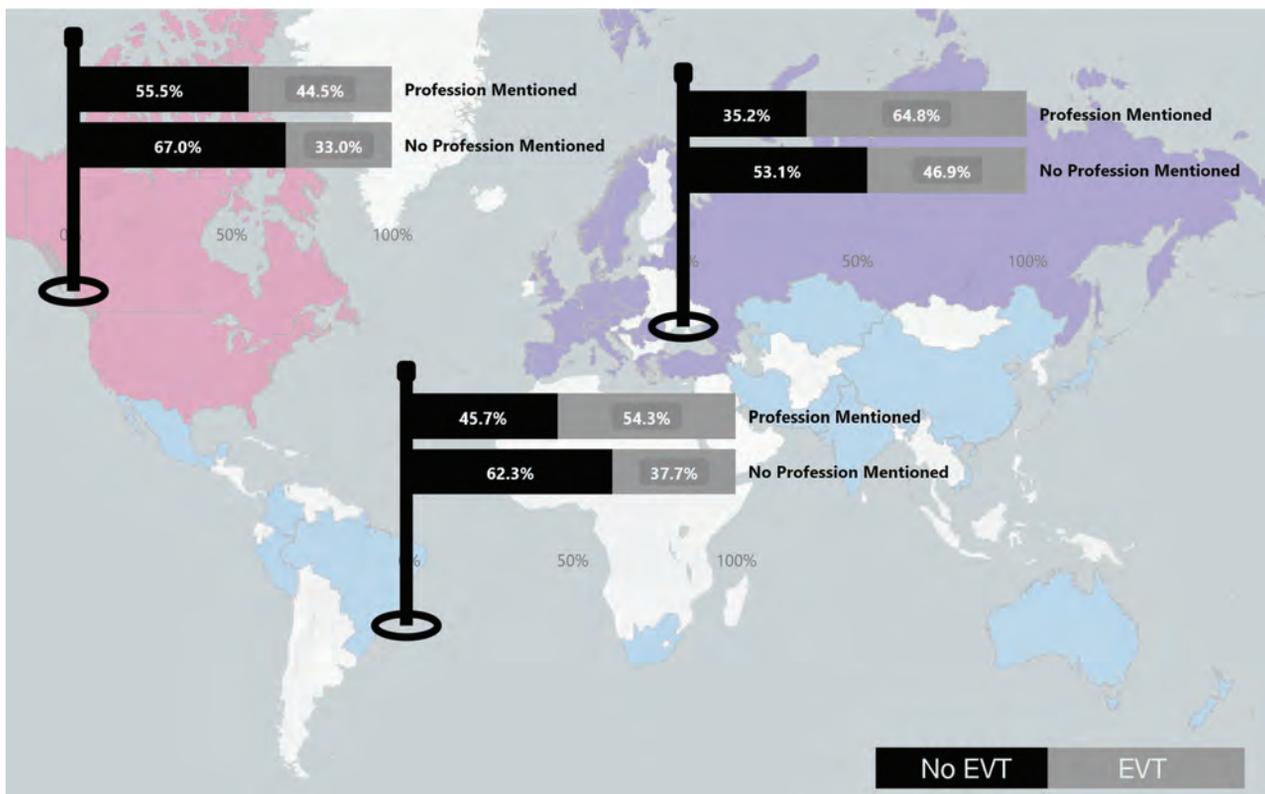
FIG 2. Geographic differences in the decision to proceed with immediate EVT, stratified by patients with mild-yet-disabling deficits with respect to their professions versus patients with mild deficits without a described profession. In North America, physicians would have treated patients with low NIHSS scores and personally disabling deficits in 44.5% (169/380) of cases, in contrast to 33% (94/285) of cases for which information regarding a profession was not provided. In Europe and the rest of the world (comprising Africa, Asia and the Pacific, the Middle East, and South and Latin America), the practitioners were in favor of immediate EVT in 64.8% (464/716) and 54.3% (200/368) in patients with a personally relevant deficit, respectively, compared with 46.9% (252/537) and 37.7% (104/276), respectively, of scenarios in which a personally disabling deficit was not mentioned.

NIHSS scores at presentation may still translate to very different patient prognoses.[18] These findings are important from a trial perspective because the inclusion of patients with low NIHSS scores has the potential for biased interpretation or "cherry-picking" of results.

Taken together, the results of the current survey and the above studies point toward a need for a case-by-case risk-benefit analysis for low-NIHSS MeVO cases.

Limitations
A major limitation of this study comes from the construction of the case scenarios. Their design to best reflect the real-world situation introduces some inherent confounding factors. For example, we could not examine the effects of patient-specific factors, such as age, on the decision to treat because all the patients with low NIHSS scores were younger (younger than 65 years of age). Furthermore, the occlusion sites were specifically selected to have an impact on the patient profession. Although the effect of scenario characteristics on EVT decision-making was not analyzed in this study, participants may have been more willing to perform EVT for occlusions of the M2 segment as opposed to those of the A3 segment, for example. We were also unable to perform a direct comparison of scenarios with and without information regarding profession because other features of the vignette were simultaneously changed.

Because there is no international registry of stroke physicians or interventionalists, participant enrollment was based on institutional networks and co-operations. The participant sample may therefore not be representative of the entire stroke community. In addition, while survey responses reflect participants' attitudes, they are not necessarily equivalent to the decision-making processes of routine clinical practice. Finally, participants were provided with radiologic images for each case, with anatomic details that are not generalizable to all occlusions of a particular vessel segment. Despite these limitations, we believe that this study provides important insight into physician decision-making for EVT in MeVO strokes with mild neurologic deficits.

CONCLUSIONS
In this survey, the presence of a patient-relevant deficit in low-NIHSS AIS due to MeVO had a significant impact on physicians' decision-making in favor of immediate EVT, highlighting the importance of not only guideline recommendations but also patient-specific factors for the current stroke community.

ACKNOWLEDGMENTS

REFERENCES

1. Goyal M, Menon BK, van Zwam WH, et al. **Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials.** *Lancet* 2016;387:1723– 31 CrossRef Medline

2. Barber PA, Zhang J, Demchuk AM, et al. **Why are stroke patients excluded from TPA therapy? An analysis of patient eligibility.** *Neurology* 2001;56:1015–20 CrossRef Medline

3. Goldhoorn RB, Mulder MJ, Jansen IG, et al. MR CLEAN Registry investigators. **Safety and outcome of endovascular treatment for minor ischemic stroke: results from the Multicenter Clinical Registry of Endovascular Treatment of Acute Ischemic Stroke in the Netherlands.** *J Stroke Cerebrovasc Dis* 2019;28:542–49 CrossRef Medline

4. Kenmuir CL, Hammer M, Jovin T, et al. **Predictors of outcome in patients presenting with acute ischemic stroke and mild stroke scale scores.** *J Stroke Cerebrovasc Dis* 2015;24:1685–89 CrossRef Medline

5. Seners P, Ben Hassen W, Lapergue B, et al. MINOR-STROKE Collaborators. **Prediction of early neurological deterioration in individuals with minor stroke and large-vessel occlusion intended for intravenous thrombolysis alone.** *JAMA Neurol* 2021;78:321–28 CrossRef Medline

6. Kim JT, Park MS, Chang J, et al. **Proximal arterial occlusion in acute ischemic stroke with low NIHSS scores should not be considered as mild stroke.** *PLoS One* 2013;8:e70996 CrossRef Medline

7. Turc G, Bhogal P, Fischer U, et al. **European Stroke Organisation (ESO): European Society for Minimally Invasive Neurological Therapy (ESMINT) Guidelines on Mechanical Thrombectomy in Acute Ischaemic Stroke Endorsed by Stroke Alliance for Europe (SAFE).** *Eur Stroke J* 2019;4:6–12 CrossRef Medline

8. Ospel JM, Kashani N, Turjman F, et al. **Discrepancies between current and ideal endovascular stroke treatment practice in Europe and North America: results from UNMASK EVT—a multidisciplinary survey.** *Interv Neuroradiol* 2020;26:420–24 CrossRef Medline

9. Saver JL, Chapot R, Agid R, et al. Distal Thrombectomy Summit Group. **Thrombectomy for distal, medium vessel occlusions: a consensus statement on present knowledge and promising directions.** *Stroke* 2020;51:2872–84 CrossRef Medline

10. Goyal M, Ospel JM, Menon BK, et al. **MeVO: the next frontier?** *J Neurointerv Surg* 2020;12:545–47 CrossRef Medline

11. Goyal M, Kappelhof M, McDonough R, et al. **Secondary medium vessel occlusions: when clots move north.** *Stroke* 2021;52:1147–53 CrossRef Medline

12. Ospel JM, Menon BK, Demchuk AM, et al. **Clinical course of acute ischemic stroke due to medium vessel occlusion with and without intravenous alteplase treatment.** *Stroke* 2020;51:3232–40 CrossRef Medline

13. Ospel JM, Kim B, Heo JH, et al. **Endovascular treatment decision-making in acute ischemic stroke patients with large vessel occlusion and low National Institutes of Health Stroke Scale: insights from UNMASK EVT, an international multidisciplinary survey.** *Neuroradiology* 2020;62:715–21 CrossRef Medline

14. Kunz WG, Almekhlafi MA, Goyal M. **Distal vessel occlusions: when to consider endovascular thrombectomy.** *Stroke* 2018;49:1581–83 CrossRef Medline

15. Lima FO, Furie KL, Silva GS, et al. **Prognosis of untreated strokes due to anterior circulation proximal intracranial arterial occlusions detected by use of computed tomography angiography.** *JAMA Neurol* 2014;71:151–57 CrossRef Medline

16. Menon BK, Hill MD, Davalos A, et al. **Efficacy of endovascular thrombectomy in patients with M2 segment middle cerebral artery occlusions: meta-analysis of data from the HERMES Collaboration.** *J Neurointerv Surg* 2019;11:1065–69 CrossRef Medline

17. Tian H, Parsons MW, Levi CR, et al. **Influence of occlusion site and baseline ischemic core on outcome in patients with ischemic stroke.** *Neurology* 2019;92:e2626–43 CrossRef Medline

18. Sucharew H, Khoury J, Moomaw CJ, et al. **Profiles of the National Institutes of Health Stroke Scale items as a predictor of patient outcome.** *Stroke* 2013;44:2182–87 CrossRef Medline

Diagnostic Role of Diffusion-Weighted and Dynamic Contrast-Enhanced Perfusion MR Imaging in Paragangliomas and Schwannomas in the Head and Neck

Y. Ota, E. Liao, A.A. Capizzano, R. Kurokawa, J.R. Bapuraj, F. Syed, A. Baba, T. Moritani, and A. Srinivasan

ABSTRACT

BACKGROUND AND PURPOSE: Distinguishing schwannomas from paragangliomas in the head and neck and determining succinate dehydrogenase (*SDH*) mutation status in paragangliomas are clinically important. We aimed to assess the clinical usefulness of DWI and dynamic contrast-enhanced MR imaging in differentiating these 2 types of tumors, as well as the *SDH* mutation status of paragangliomas.

MATERIALS AND METHODS: This retrospective study from June 2016 to June 2020 included 42 patients with 15 schwannomas and 27 paragangliomas (10 *SDH* mutation-positive and 17 *SDH* mutation-negative). ADC values, dynamic contrast-enhanced MRI parameters, and tumor imaging characteristics were compared between the 2 tumors and between the mutation statuses of paragangliomas as appropriate. Multivariate stepwise logistic regression analysis was performed to identify significant differences in these parameters.

RESULTS: Fractional plasma volume ($P \leq .001$), rate transfer constant ($P = .038$), time-to-maximum enhancement ($P < .001$), maximum signal-enhancement ratio ($P < .001$) and maximum concentration of contrast agent ($P < .001$), velocity of enhancement ($P = .002$), and tumor characteristics including the presence of flow voids ($P = .001$) and enhancement patterns ($P = .027$) showed significant differences between schwannomas and paragangliomas, though there was no significant difference in ADC values. In the multivariate logistic regression analysis, fractional plasma volume was identified as the most significant value for differentiation of the 2 tumor types ($P = .014$). ADC values were significantly higher in nonhereditary than in hereditary paragangliomas, while there was no difference in dynamic contrast-enhanced MR imaging parameters.

CONCLUSIONS: Dynamic contrast-enhanced MR imaging parameters show promise in differentiating head and neck schwannomas and paragangliomas, while DWI can be useful in detecting *SDH* mutation status in paragangliomas.

ABBREVIATIONS: AUC = area under the curve; DCE = dynamic contrast-enhanced; EES = extravascular extracellular space; Kep = rate transfer constant between EES and blood plasma per minute; K^{trans} = volume transfer constant between EES and blood plasma per minute; SDH = succinate dehydrogenase; SER = signal-enhancement ratio; TIC = time-intensity curve; TME = time-to-maximum enhancement; Ve = EES volume per unit tissue volume; Vp = blood plasma volume per unit tissue volume

Schwannomas are benign nerve sheath tumors arising from Schwann cells, and paragangliomas are neuroendocrine tumors arising from the autonomic system.[1,2] Both schwannomas and paragangliomas can occur in the head and neck region. On conventional MR imaging, schwannomas typically present as a homogeneously enhancing mass with cystic changes, whereas paragangliomas usually demonstrate heterogeneous enhancement with a "salt-and-pepper" appearance and necrotic or cystic changes.[3,4] When present, these classic imaging characteristics can help define these lesions, but a definitive diagnosis on imaging remains challenging, especially when these imaging characteristics are not present or overlap.

Moreover, conventional MR imaging and CT have been reported to be unable to identify the difference between nonhereditary and hereditary paragangliomas.[5] Hereditary paragangliomas are primarily related to mutations in the genes of succinate dehydrogenase (*SDH*), which are responsible for *SDH* subunits *A, B, C, D,* and *AF2* (*SDHA, SDHB, SDHC, SDHD, SDHAF2*) proteins and play an important role in the mitochondria for energy production. For example, *SDHD* and *SDHB* mutations are related to multiplicity and malignancy, respectively,[1] and these associations make the differentiation of nonhereditary and hereditary paragangliomas highly important.

Schwannomas are histologically characterized by regions of high cellularity and fewer cells with cystic or xanthomatous changes, whereas paragangliomas mainly demonstrate nests of tumor cells separated by peripheral capillaries.[6,7] ADC values may

FIG 1. A 41-year-old woman with a schwannoma in the left parotid space. *A*, T2-weighted image shows a hyperintense mass in the left parotid gland. *B* and *C*, Contrast-enhanced T1-weighted image shows homogeneous enhancement. *D*, Mean ADC and normalized mean ADC are 1.15 × 10^{-3} mm^2/s and 1.48, respectively. *E* and *F*, Vp and TME are demonstrated. *G*, The TIC demonstrates a low peak enhancement with a long time-to-peak.

in a single center. There were 97 pathologically proved schwannomas and 36 paragangliomas in the head and neck. DCE-MR imaging was performed when head and neck malignancy was suspected or the lesions required further characterization. Patients who did not have pretreatment conventional MR imaging or DCE-MR imaging ($n =$ 72); had been previously treated by surgery, embolization, or radiation therapy ($n = 15$); or did not have genetic testing for *SDH* mutations for paragangliomas ($n = 4$) were excluded. Genetic testing was performed by the PGLNext panel (Ambry Genetics), which requires collecting blood or saliva samples by an appropriate kit and analyzes 12 genes including *SDH* subunits *SDHA*, *SDHAF2*, *SDHB*, *SDHC,* and *SDHD*. This test was designed and validated to detect >99.9% of the gene mutations noted above.

In total, 42 patients (12 men, 30 women; 46.0 [SD, 16.5] years of age; range, 18–70 years of age) with 15 schwannomas and 27 paragangliomas (10 *SDH* mutation-negative, 17 *SDH* mutation-positive) were included in this study.

MR Imaging Acquisition
MR imaging examinations were performed using 1.5T ($n = 30$) and 3T ($n = 12$) scanners (Ingenia; Philips Healthcare). They were performed with a 16-channel Neurovascular coil (Stryker) with the patient in the supine position. Acquired sequences included axial T1WI and T2WI, axial and coronal contrast-enhanced fat-saturated T1WI, and DWI using echo-planar imaging with the following DWI parameters: TR range = 5000–10,000 ms; TE range = 58–106 ms; number of excitations = 1, 2; section thickness/gap = 3.5–4/0–1 mm; FOV = 220–260 mm; matrix size = 128–200 × 128–200; and 3 diffusion directions. Sensitizing diffusion gradients were applied sequentially with b-values of 0 and 1000 s/mm^2.

The DCE-MR imaging sequence was performed using 3D T1-weighted fast-field echo images, with the administration of gadobenate dimeglumine contrast (MultiHance; Bracco Diagnostics). An intravenous bolus of 20 mL of gadobenate dimeglumine was administered using a power injector with a flow rate of 5.0 mL/s through a peripheral arm vein, followed by a 20-mL saline flush. DCE-MR imaging was sequentially performed for 30 dynamic phases. These techniques were performed for all patients in a single center.

The parameters of 3D T1 fast-field echo were as follows: TR = 4.6 ms, TE = 1.86 ms, flip angle =30°, section thickness =

contribute to their imaging differentiation on the basis of these histopathologic differences. Also, dynamic contrast-enhanced (DCE)-MR imaging can assess tumor microvasculature and permeability[8,9] and has been increasingly used to assess head and neck lesions.[10-13] However, DWI and DCE-MR imaging analysis for these tumors and for the mutation status of paragangliomas has not been fully explored. This study was designed to test the hypothesis that ADC values and DCE-MR imaging could help differentiate schwannomas from paragangliomas and distinguish nonhereditary and hereditary paragangliomas.

MATERIALS AND METHODS
Study Population
Our institutional review board approved this retrospective single-center study and waived the requirement for informed consent. Data were acquired in compliance with all applicable Health Insurance Portability and Accountability Act regulations.

We retrospectively reviewed the medical records of 775 patients from June 2016 to June 2020 with suspected head and neck tumors

FIG 2. A 26-year-old man positive for the *SDHB* mutation with a paraganglioma in the left jugular foramen. *A*, T2-weighted image shows a heterogeneous irregular mass in the left jugular foramen. *B* and *C*, Contrast-enhanced T1-weighted images show a heterogeneous, enhancing mass with a vascular flow void. *D*, Mean ADC and normalized mean ADC are 1.12×10^{-3} mm^2/s and 1.48. *E* and *F*, Vp and TME are demonstrated. *G*, The TIC demonstrates a higher peak enhancement with a shorter time-to-peak compared with Fig 1.

5.0 mm, FOV = 240 × 240 mm^2, voxel size = 1.0 × 1.0 × 5.0 mm^3, number of excitations = 1, number of slices per dynamic scan = 48, temporal resolution = 8.4 seconds, and total acquisition time of 4 minutes 13 seconds using a 16-channel Neurovascular coil.

Data Analysis

Tumor Characteristics. All conventional MR images were reviewed independently by 2 board-certified radiologists with 7 and 10 years of experience in neuroradiology, respectively. They were blinded to clinical information, imaging results from other modalities, and histopathologic results. Both radiologists evaluated imaging characteristics using the following metrics:

1) Maximum axial diameter of the tumor was evaluated on postcontrast T1-weighted images

2) Presence of flow voids, cystic or necrotic changes, and enhancement pattern (homogeneous or heterogeneous pattern),

evaluated on T1-weighted, T2-weighted, and pre- and postcontrast fat-saturated T1-weighted images. These were recorded as binary variables. Cystic changes were defined as the following: nonenhancing, predominantly T1 hypointense and T2 hyperintense areas; necrotic changes were defined as nonenhancing, predominantly T1 hypointense and heterogeneously T2 hyperintense areas; and flow voids were defined as nonenhancing T1 hypointense, T2 hypointense vessel structures within the tumors (Figs 1*A*–*C* and 2*A*–*C*).

3) As for location (percentage of head lesion/head and neck lesion), jugular foramen and jugulotympanicum lesions were classified into the head location, and carotid space and parotid space lesions were classified into the neck location.

ADC Analysis. ADC maps were constructed by a monoexponential fitting model using commercially available software (Olea Sphere, Version 3.0; Olea Medical). The radiologist with 7 years'

experience carefully outlined the lesions using freehand ROIs (Figs 1D and 2D) and transposed the ROIs to the ADC map. When geometric distortion was observed, the location and size were adjusted on the ADC map so that the ROI would be included in the tumor. The ROIs encompassed predominantly solid, enhancing portions of tumors without cystic or necrotic areas on postcontrast T1-weighted images. Additionally, ROIs spared the peripheral 2 mm of lesions to avoid volume averaging.[14] As an internal standard, an ROI was placed within the cervical spinal cord at the level of the C2–C3 disc space, which was included in the FOV of every study.[15] A normalized ADC ratio was calculated by dividing each mean ADC value of the lesion by the mean ADC value of the cervical cord to adjust for variation of ADC values across MR imaging scanners, magnetic field strengths, and matrix sizes.

Quantitative DCE-MR Imaging Analysis. All quantitative analyses in DCE-MR imaging were performed using the Olea Sphere 3.0 software Permeability Module, which is based on the extended Tofts model, by which pixel-based parameter maps are calculated from time-intensity curves (TICs) (Figs 1 E–G and 2E–G). The radiologist with 7 years' experience placed the ROIs on the permeability maps, predominantly including the enhancing components of the tumors without cystic or necrotic areas. Manual ROIs spared the peripheral 2 mm of lesions to avoid volume averaging. An ROI was placed at the external carotid artery of the affected side for the arterial input function. The calculated quantitative parameters were blood plasma volume per unit tissue volume (Vp), extravascular extracellular space (EES) volume per unit tissue volume (Ve), volume transfer constant between EES and blood plasma per minute (K^{trans}), and rate transfer constant between EES and blood plasma per minute (Kep).

Semiquantitative DCE-MR Imaging Analysis. Semiquantitative analysis was performed using the same ROIs described above with the Olea Sphere 3.0 software Permeability Module. The average signal intensity within the ROI was plotted against time, and TICs were constructed. The following parameters were calculated on a pixel-by-pixel basis from the TICs: area under the curve (AUC, the relative quantity of contrast agent over time), maximum concentration of contrast agent (peak enhancement), velocity of enhancement (wash-in), velocity of enhancement loss (washout), maximum signal-enhancement ratio (SER), and time-to-maximum enhancement (TME).

Statistical Analysis

Patient demographics including age and sex; the number of lesions; tumor characteristics including the maximum diameter of tumors and the presence of vascular flow voids; the presence of cystic or necrotic change; enhancement patterns (homogeneous or heterogeneous pattern); and location (percentage of head lesions/head and neck lesions) were compared between the 2 types of tumors and between SDH mutation-positive and SDH mutation-negative paragangliomas.

Age was compared using a t test and described as mean (SD). The maximum diameter of the tumor was compared using the Mann-Whitney U test. The binary variables such as sex (ratio of male to female), the presence of vascular flow voids, the presence of cystic or necrotic change, enhancement patterns, and location (the percentage of head lesions/head and neck legions) were compared using the Fisher exact test.

For statistically significant tumor characteristics, AUCs were evaluated from the receiver operating characteristic analysis.

Mean ADC and normalized mean ADC (mean ADC divided by the ADC of the cervical cord) were compared between paragangliomas and schwannomas using a t test. Additionally, mean ADC and normalized mean ADC were also compared between SDH mutation-negative and SDH mutation-positive paragangliomas. Quantitative parameters and semiquantitative parameters were compared between schwannomas and paragangliomas and between SDH mutation-negative and SDH mutation-positive paragangliomas using the Mann-Whitney U test.

For values that showed statistically significant differences in ADC values, quantitative parameters, and semiquantitative parameters, the optimal cutoff values in receiver operating characteristic analysis were determined as a value to maximize the Youden index (sensitivity + specificity–1). On the basis of the cutoff values, the diagnostic performances (sensitivity, specificity, positive predictive value, negative predictive value, and accuracy) were calculated.

Multivariate stepwise logistic regression analysis was performed to identify the most significant parameter to distinguish schwannomas and paragangliomas using the forward stepwise selection method. For this method, the values with a P value of < .05 according to the univariate analysis were used.

Interreader agreement for tumor characteristics was assessed using the κ coefficient, which was interpreted as follows: <0.40, poor-to-fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.81–1.00, almost perfect agreement.[16]

All statistical calculations were conducted with JMP Pro, Version 15.0.0 (SAS Institute). Variables with a P value <.05 were considered statistically significant.

RESULTS

The patient demographic and tumor characteristics are shown in Table 1.

There was no significant difference in age or sex (P = .27 and .73) between those with schwannomas and those with paragangliomas. Regarding tumor characteristics, the presence of vascular flow voids and the enhancement pattern (homogeneous or heterogeneous pattern) revealed significant differences between paragangliomas and schwannomas (P = .001 and .027, respectively). The AUCs of the presence of vascular flow voids and enhancement patterns were 0.77 (95% CI, 0.64–0.89) and 0.69 (95% CI, 0.53–0.84), respectively. Schwannomas were located at the jugular foramen (n = 5), carotid space (n = 6), and parotid space (n = 4). Paragangliomas included the glomus jugulare (n = 8), glomus jugulotympanicum (n = 3), and carotid body paraganglioma (n = 16). There was no significant difference in location (head lesions/head and neck lesions) between the 2 tumors (schwannomas: 11/27 versus paragangliomas: 5/15; P = .75).

There were 10 SDH mutation-negative and 17 SDH mutation-positive paragangliomas (SDHA: 1, SDHB: 8, SDHC: 2, SDHD: 6).

Table 1: Patient demographics and tumor characteristics in schwannomas and paragangliomas[a]

	Schwannomas (n = 15)	Paragangliomas (n = 27)	P Value
No.	15	27	NA
Sex (male/female)	5:10	7:20	.73
Age (yr)	42.2 (SD, 15.1)	48.2 (SD, 17.2)	.27
Maximum diameter (mm)	31.4 (SD, 15.4)	32.3 (SD, 15.6)	.76
Presence of vascular flow voids (salt-and-pepper appearance)	2/15	18/27	.001[b]
Cystic or necrotic change	5/15	14/27	.34
Enhancement pattern (homogeneous enhancement/total)	10/15	8/27	.027[b]
Location (head lesion/total)	5/15	11/27	.75

Note:—NA indicates not applicable.
[a] Values are mean (SD).
[b] P < .05.

Table 2: ADC values and DCE-MR imaging dynamic parameters of schwannomas and paragangliomas[a]

	Schwannomas (n = 15)	Paragangliomas (n = 27)	P Value
Mean ADC ($\times 10^{-3}$ mm^2/s)	1.17 (SD, 0.31)	1.12 (SD, 0.26)	.56
Normalized mean ADC	1.54 (SD, 0.43)	1.45 (SD, 0.37)	.53
Quantitative values			
Vp	0.06 (0.03–0.12)	0.40 (0.34–0.54)	<.001[b]
Ve	0.42 (0.25–0.77)	0.28 (0.16–0.47)	.38
Kep (min^{-1})	0.58 (0.36–0.67)	1.02 (0.43–2.27)	.038[a]
K^{trans} (min^{-1})	0.15 (0.11–0.26)	0.31 (0.13–0.74)	.16
Semiquantitative values			
AUC (mmol.min/L)	5.6×10^4 (1.95×10^4 –1.55×10^5)	7.0×10^4 (3.8×10^4 –4.1×10^5)	.08
Peak enhancement	137 (75–220)	294 (266–300)	<.001[b]
Wash-in	1.4 (0.74–3.8)	7.26 (2.54–42.6)	.002[b]
Washout	3.4 (0.77–8.4)	1.0 (0.42–6.0)	.35
SER	53.9 (47–90)	146 (112–211)	<.001[b]
TME (sec)	160 (108–239)	36.2 (27.7–77.4)	<.001[b]

[a] The numbers in parentheses represent interquartile range.
[b] P < .05.

Between *SDH* mutation-negative and mutation-positive paragangliomas, age was higher in *SDH* mutation-negative paragangliomas than in *SDH* mutation-positive paragangliomas (56.8 [SD, 12.9] years versus 43.1 [SD, 17.7] years; P = .043). There were no significant differences in other demographics and tumor characteristics (male/female: 1:9 versus 6:1; P = .20; maximum diameter: 36.2 [SD, 21.5] mm versus 28.6 [SD, 10.2] mm; P = .22; presence of vascular flow voids [salt-and-pepper appearance]: 6/10 versus 12/17; P = .68; percentage of cystic or necrotic changes: 5/10 versus 9/17; P = 1.0; enhancement pattern [homogeneous enhancement/total]: 4/10 versus 4/17; P = .42; location [head lesion/total]: 5/10 versus 6/17; P = .69).

Interreader agreement for tumor characteristics showed almost perfect agreement (κ = 0.85–0.89).

For schwannomas and paragangliomas, Table 2 summarizes the results of ADCs and dynamic perfusion data. There were no significant differences in mean ADC and normalized mean ADC between schwannomas and paragangliomas (P = .56 and 0.46, respectively). The mean ADC value of the cervical cord, used as an internal standard, was 0.77 (SD, 0.05) $\times 10^{-3}$ mm^2/s. Regarding quantitative DCE parameters, Vp (schwannomas versus paragangliomas; median, 0.06, versus 0.40; P = <.001) and Kep (median, 0.58, versus 1.02; P = .038) showed significant differences. Arterial input function curves showed the pulsed input pattern in all patients. Among semiquantitative parameters, TME (median, 160, versus 36.2; P < .001), SER (median, 53.9, versus 146; P < .001), peak enhancement (median, 137, versus 294; P < .001), and wash-in (median, 1.40, versus 7.26; P = .002) showed significant differences between schwannomas and paragangliomas. Representative cases of DCE-MR imaging analysis are shown in Figs 3 and 4.

The diagnostic performance of Vp, Kep, TME, SER, peak enhancement, and wash-in are shown in Table 3, and receiver operating characteristic curves of Vp, TME, and peak enhancement, in Fig 5. The AUCs of Vp and TME offered the best performance of the quantitative and semiquantitative perfusion parameters, respectively. In the multivariate logistic regression analysis for differentiating schwannomas and paragangliomas from ADCs, DCE-MR imaging parameters, and tumor characteristics, Vp was identified as the most significant variable in the differentiation of these 2 tumors (P = .014).

As for *SDH* mutation-negative and *SDH* mutation-positive paragangliomas, mean ADC and normalized mean ADC were significantly higher in *SDH* mutation-negative paragangliomas than in *SDH* mutation-positive paragangliomas (mean ADC: 1.27 [SD, 0.18] versus 1.04 [SD, 0.26] $\times 10^{-3}$ mm^2/s; P = .023; normalized mean ADC: 1.73 [SD, 0.32] versus 1.31 [SD, 0.31]; P < .001, respectively). The diagnostic performances of mean ADC and normalized mean ADC demonstrated sensitivities of 0.82 and 0.88, specificities of 0.89 and 0.78, positive predictive values of 0.93 and 0.88, negative predictive values of 0.73 and 0.78, accuracy of 0.85 and 0.85, and AUCs of 0.83 and 0.87, with cutoff values of 1.14×10^{-3} mm^2/s and 1.52, respectively.

Regarding DCE-MR imaging analysis of *SDH* mutation-negative and *SDH* mutation-positive paragangliomas, there were no perfusion parameters that showed statistical significance (Vp: median, 0.36 versus 0.41; P = .68; Ve: median, 0.33 versus 0.28; P = .89; Kep: median, 1.2 versus 0.90 minute^{-1}; P = .53; K^{trans}: median, 0.28 versus 0.31 minute^{-1}; P = 1.0; AUC: median, 6.6×10^4 versus 2.3×10^5 mmol.min/L; P = .90; peak enhancement: median, 287 versus 297; P = .96; wash-in: median, 5.5 versus 9.6; P = .67; washout: median, 1.0 versus 1.2; P = .98; SER: median, 150 versus 142; P = .54; TME: median, 46.4 versus 36.1; P = .98).
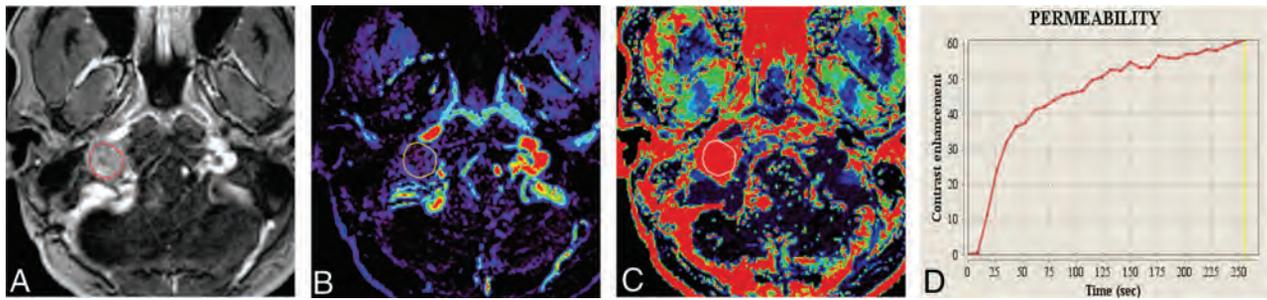
FIG 3. A 34-year-old woman with a schwannoma in the right jugular foramen. *A*, The permeability map shows a heterogeneously enhancing mass in the right jugular foramen. *B* and *C*, The values of Vp and TME are 0.03 and 219, respectively. *D*, The TIC demonstrates a low peak enhancement with a long time-to-peak compared with Fig 2.
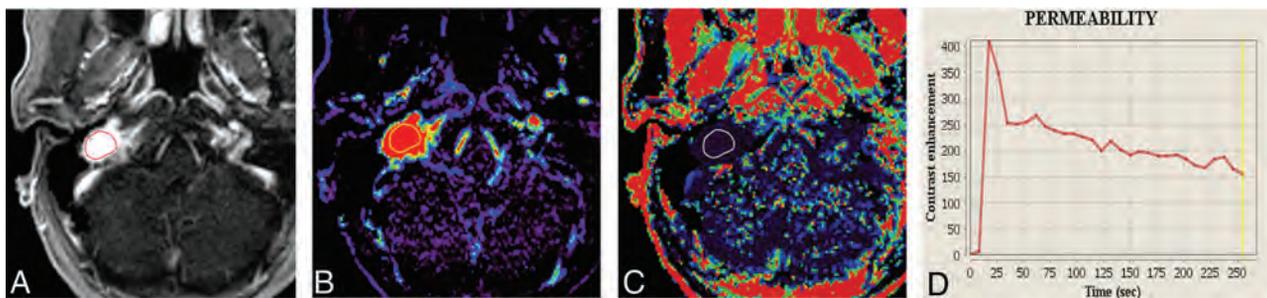
FIG 4. A 70-year-old man negative for the *SDH* mutation with a paraganglioma in the right jugular foramen. *A*, The permeability map shows a homogeneously enhancing mass in the right jugular foramen. *B* and *C*, The values of Vp and TME are 0.53 and 27.7, respectively. *D*, The TIC demonstrates a high peak enhancement with a short time-to-peak compared with Figs 1 and 3.

Table 3: Diagnostic performance of dynamic parameters in differentiating schwannomas and paragangliomas

	Vp	Kep	Peak Enhancement	Wash-In	SER	TME (Sec)
Cutoff	0.30	0.86	277	6.29	135	62.8
Sensitivity	0.91	0.67	0.74	0.63	0.89	0.72
Specificity	1.00	0.87	1.00	0.93	0.73	1.00
PPV	1.00	0.90	1.00	0.94	0.86	1.00
NPV	0.88	0.59	0.68	0.58	0.79	0.68
Accuracy	0.95	0.74	0.83	0.74	0.83	0.83
AUC	0.99	0.70	0.89	0.80	0.86	0.92

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

DISCUSSION

This retrospective study aimed to evaluate the clinical usefulness of DWI and DCE-MR imaging parameters to differentiate head and neck schwannomas and paragangliomas and distinguish *SDH* mutation status in paragangliomas. DCE-MR imaging parameters identified significant statistical differences between schwannomas and paragangliomas with AUCs of 0.70–0.99, though no significant differences in ADC values were identified. Vp was the most promising parameter to differentiate the 2 tumor types. Within the paraganglioma subgroup, *SDH* mutation-negative paragangliomas showed higher mean and normalized mean ADC values than *SDH* mutation-positive paragangliomas with AUCs of 0.88 and 0.92, while DCE-MR imaging parameters failed to show any significant differences. Regarding tumor characteristics, as previous studies have reported, the presence of flow voids and the enhancement

pattern showed significant differences between schwannomas and paragangliomas, while age was younger in *SDH* mutation-positive than *SDH* mutation-negative paragangliomas.[1,3,4]

Semiquantitative parameters are based on and calculated from TICs, which can differentiate divergent hemodynamic patterns. In our study, TME, SER, peak enhancement, and wash-in showed significant differences between the 2 tumors. Paragangliomas showed higher peak enhancement and SER, with a shorter TME and greater wash-in than schwannomas. TME showed the most significant difference between the 2 tumors with an AUC of 0.92. Pathologically, paragangliomas typically show chief cells forming variably sized clusters with a capillary network that primarily functions as an arteriovenous shunt,[17] while schwannomas show focal clusters of vascular proliferation with extensive hyalinization.[18] These histologic differences are suspected to underlie the divergent hemodynamic patterns between the 2 tumors.

Quantitative parameters were calculated on the basis of the extended Tofts model, which reflects the microcirculation within the lesion.[19] In our study, Vp and Kep were significantly higher in paragangliomas than in schwannomas, and Vp showed the most significant difference between the 2 tumors with an AUC of 0.99 at a cutoff of 0.30 seconds. Vp is defined as blood plasma
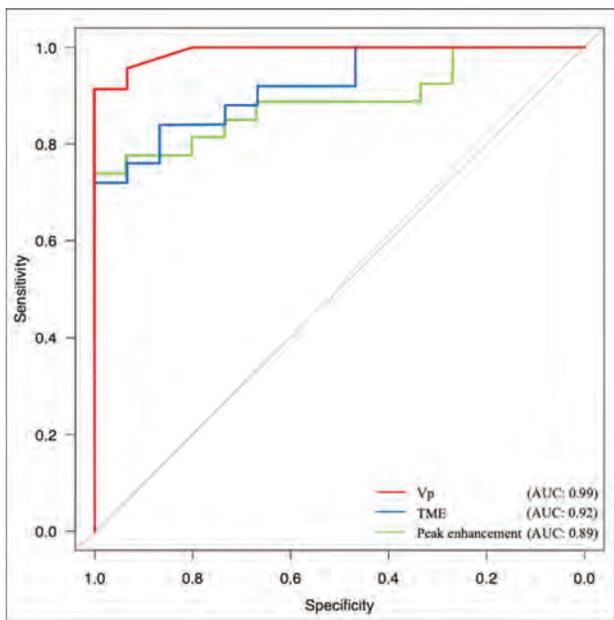
FIG 5. Receiver operating characteristic curves of Vp, TME, and peak enhancement. The AUCs of Vp, TME, and peak enhancement were 0.99, 0.92, and 0.89 with cutoff values of 0.30, 62.8, and 277 seconds, respectively.

volume per unit tissue volume, which may represent vascularity within the tumor, and Kep is defined as rate transfer of constant between EES and blood plasma per minute, which represents permeability.[20] Our results indicate that DCE-MR imaging can be used to differentiate schwannomas from paragangliomas on the basis of their disparate hemodynamic patterns and microcirculation, further refining our ability to differentiate these tumors on imaging.

There were no significant differences in DCE-MR imaging parameters when comparing *SDH* mutation-negative and mutation-positive paragangliomas in this study. Published literature evaluating potential differences in the vascularity or permeability of paragangliomas based on *SDH* mutation status is lacking. However, the results of this study indicate that there may be no significant difference in the microcirculation of paragangliomas based on *SDH* mutation status.

ADC analysis did not identify any significant differences between schwannomas and paragangliomas; however, it demonstrated significant differences between *SDH* mutation-negative and mutation-positive paragangliomas. Histologically, paragangliomas have been recognized as showing various tumor cell morphology and cellularity and a variety of histologic patterns,[21] including chief cells forming variable-sized clusters. A previous study suggested that the histopathologic background of paragangliomas may differ depending on *SDH* mutation status; this potential divergence in histology may account for the visualized differences in ADC values,[5] though correlative histologic studies are lacking. Schwannomas demonstrate a biphasic pattern of high cellularity, and fewer cells with cystic or xanthomatous changes.[7] Previously, it has been reported that schwannomas exhibit a wide range of ADC values (0.74 [SD, 0.08] to 2.08 [SD, 0.33] $\times 10^{-3}$ mm^2/s),[14] which could be due to the differences in

internal structures such as cystic or xanthomatous changes as mentioned above. Given that there were no differences in ADC values between the 2 tumor types, even despite avoiding cystic or necrotic components of the tumors when drawing ROIs, DWI may not reliably differentiate these tumors.

Our study had several limitations. First, this was a retrospective study and included a relatively small population from a single institution. Second, DCE-MR imaging was performed using both 1.5T and 3T scanners, which may add heterogeneity to the calculated perfusion parameters. Third, in DCE-MR imaging analysis, there could be a potential bias influenced by the previous knowledge of the morphologic findings of the tumors, even though the reader was blinded to the pathologic results. Finally, in our institution, genetic testing is recommended for all the patients who are clinically or radiologically suspected of having hereditary paragangliomas, so prior probability of genetic mutation in our study population may be higher than in the overall population.

CONCLUSIONS

DCE-MR imaging can provide promising noninvasive parameters that can be used to differentiate schwannomas and paragangliomas in the head and neck when the differential diagnosis is challenging. Vp is the most promising value to differentiate the 2 tumors. DWI can be useful to distinguish *SDH* mutation status in paragangliomas.

REFERENCES

1. Withey SJ, Perrio S, Christodoulou D, et al. **Imaging features of succinate dehydrogenase-deficient pheochromocytoma-paraganglioma syndromes.** *Radiographics* 2019;39:1393–1410 CrossRef Medline
2. Biswas D, Marnane CN, Mal R, et al. **Extracranial head and neck schwannomas–a 10-year review.** *Auris Nasus Larynx* 2007;34:353–59 CrossRef Medline
3. Woolen S, Gemmete JJ. **Paragangliomas of the head and neck.** *Neuroimaging Clin N Am* 2016;26:259–78 CrossRef Medline
4. Anil G, Tan TY. **Imaging characteristics of schwannoma of the cervical sympathetic chain: a review of 12 cases.** *AJNR Am J Neuroradiol* 2010;31:1408–12 CrossRef Medline
5. Ota Y, Naganawa S, Kurokawa R, et al. **Assessment of MR imaging and CT in differentiating hereditary and nonhereditary paragangliomas.** *AJNR Am J Neuroradiol* 2021;42:1320–26 CrossRef Medline
6. Williams MD. **Paragangliomas of the head and neck: an overview from diagnosis to genetics.** *Head Neck Pathol* 2017;11:278–87 Medline
7. Behuria S, Rout TK, Pattanayak S. **Diagnosis and management of schwannomas originating from the cervical vagus nerve.** *Ann R Coll Surg Engl* 2015;97:92–97 CrossRef Medline
8. Griffith B, Jain R. **Perfusion imaging in neuro-oncology: basic techniques and clinical applications.** *Radiol Clin North Am* 2015;53:497–511 CrossRef Medline
9. Cuenod CA, Balvay D. **Perfusion and vascular permeability: basic concepts and measurement in DCE-CT and DCE-MRI.** *Diagn Interv Imaging* 2013;94:1187–1204 CrossRef Medline
10. Gaddikeri S, Hippe DS, Anzai Y. **Dynamic contrast-enhanced MRI in the evaluation of carotid space paraganglioma versus schwannoma.** *J Neuroimaging* 2016;26:618–25 CrossRef Medline
11. Yuan Y, Shi H, Tao X. **Head and neck paragangliomas: diffusion weighted and dynamic contrast enhanced magnetic resonance imaging characteristics.** *BMC Med Imaging* 2016;16:12 CrossRef Medline

12. Gaddikeri S, Gaddikeri RS, Tailor T, et al. **Dynamic contrast-enhanced MR imaging in head and neck cancer: techniques and clinical applications.** *AJNR Am J Neuroradiol* 2016;37:588–95 CrossRef Medline

13. Ota Y, Liao E, Kurokawa R, et al. **Diffusion-weighted and dynamic contrast-enhanced MRI to assess radiation therapy response for head and neck paragangliomas.** *J Neuroimaging* 2021 May 8. [Epub ahead of print] CrossRef Medline

14. Srinivasan A, Dvorak R, Perni K, et al. **Differentiation of benign and malignant pathology in the head and neck using 3T apparent diffusion coefficient values: early experience.** *AJNR Am J Neuroradiol* 2008;29:40–44 CrossRef Medline

15. Koontz NA, Wiggins RH 3rd. **Differentiation of benign and malignant head and neck lesions with diffusion tensor imaging and DWI.** *AJR Am J Roentgenol* 2017;208:1110–15 CrossRef Medline

16. Landis JR, Koch GG. **The measurement of observer agreement for categorical data.** *Biometrics* 1977;33:159 CrossRef Medline

17. Offergeld C, Brase C, Yaremchuk S, et al. **Head and neck paragangliomas: clinical and molecular genetic classification.** *Clinics (Sao Paulo)* 2012;67:19–28 CrossRef Medline

18. Papiez J, Rojiani MV, Rojiani AM. **Vascular alterations in schwannoma.** *Int J Clin Exp Pathol* 2014;7:4032–38 Medline

19. Sourbron SP, Buckley DL. **On the scope and interpretation of the Tofts models for DCE-MRI.** *Magn Reson Med* 2011;66:735–45 CrossRef Medline

20. Zhao M, Guo LL, Huang N, et al. **Quantitative analysis of permeability for glioma grading using dynamic contrast-enhanced magnetic resonance imaging.** *Oncol Lett* 2017;14:5418–26 CrossRef Medline

21. Tischler AS, deKrijger RR. **15 years of paraganglioma: pathology of pheochromocytoma and paraganglioma.** *Endocr Relat Cancer* 2015;22:T123–33 CrossRef Medline

Differentiation of Jugular Foramen Paragangliomas versus Schwannomas Using Golden-Angle Radial Sparse Parallel Dynamic Contrast-Enhanced MRI

A. Pires, G. Nayak, E. Zan, M. Hagiwara, O. Gonen, and G. Fatterpekar

ABSTRACT

BACKGROUND AND PURPOSE: Accurate differentiation of paragangliomas and schwannomas in the jugular foramen has important clinical implications because treatment strategies may vary but differentiation is not always straightforward with conventional imaging. Our aim was to evaluate the accuracy of both qualitative and quantitative metrics derived from dynamic contrast-enhanced MR imaging using golden-angle radial sparse parallel MR imaging to differentiate paragangliomas and schwannomas in the jugular foramen.

MATERIALS AND METHODS: A retrospective study of imaging data was performed on patients ($n = 30$) undergoing MR imaging for jugular foramen masses with the golden-angle radial sparse parallel MR imaging technique. Imaging data were postprocessed to obtain time-intensity curves and quantitative parameters. Data were normalized to the dural venous sinus for relevant parameters and analyzed for statistical significance using a Student t test. A univariate logistic model was created with a binary output, paraganglioma or schwannoma, using a wash-in rate as a variable. Additionally, lesions were clustered on the basis of the wash-in rate and washout rate using a 3-nearest neighbors method.

RESULTS: There were 22 paragangliomas and 8 schwannomas. All paragangliomas demonstrated a type 3 time-intensity curve, and all schwannomas demonstrated a type 1 time-intensity curve. There was a statistically significant difference between paragangliomas and schwannomas when comparing their values for area under the curve, peak enhancement, wash-in rate, and washout rate. A univariate logistic model with a binary output (paraganglioma or schwannoma) using wash-in rate as a variable was able to correctly predict all observed lesions ($P < .001$). All 30 lesions were classified correctly by using a 3-nearest neighbors method.

CONCLUSIONS: Paragangliomas at the jugular foramen can be reliably differentiated from schwannomas using golden-angle radial sparse parallel MR imaging–dynamic contrast-enhanced imaging when imaging characteristics cannot suffice.

ABBREVIATIONS: AUC = area under the curve; DCE = dynamic contrast-enhanced; EES = extravascular extracellular space; GRASP = golden-angle radial sparse parallel; Kep = rate transfer constant; K^{trans} = volume transfer constant; SER = signal-enhancement ratio; 3-NN = 3-nearest neighbors; TIC = time-intensity curve; TME = time-to-maximum enhancement; Ve = extravascular extracellular space plasma volume fraction; Vp = plasma volume fraction

Lesions of the jugular foramen are relatively uncommon but can be encountered on routine imaging with some regularity.[1,2] Differential considerations for lesions intrinsic to this area can be extensive; however, schwannomas and paragangliomas are 2 of the most frequently encountered primary neoplasms in this location.[2,3] Conventional imaging can sometimes distinguish the 2 lesions because each can have characteristic features. Typical imaging characteristics of paragangliomas include permeative osseous erosion, intense contrast enhancement, as well as a "salt and pepper" appearance on T1-weighted images due to subacute hemorrhage visualized as "salt" and hypointense flow voids as "pepper."[2,3] Schwannomas typically demonstrate smooth osseous remodeling, may have more homogeneous signal characteristics, and also generally demonstrate avid enhancement.[2-4] However, many lesions do not demonstrate a "typical" appearance, particularly when smaller, when bony changes and signal characteristics may be more difficult to appreciate. Smaller paragangliomas do not always demonstrate a salt and pepper appearance, and schwannomas can also demonstrate heterogeneous T1 and T2 signal characteristics due to cystic or fatty degeneration.[5-7] Similarly, osseous changes may not yet be readily appreciable, particularly on MR imaging.

Accurate differentiation of the 2 lesions has important clinical implications because treatment strategies may vary.[1,8] While both lesions can, in some cases, be surgically resected, stereotactic radiosurgery plays an increasingly important role in the treatment strategy.[9] Additional treatment and clinical decisions such

as appropriate surveillance strategies, the need for adjunct endovascular embolization, genetic testing, and evaluation for additional lesions can all depend on the diagnosis. Tissue sampling in this region is challenging and carries its own set of risks due to the multiple neurovascular structures traversing this area, and histologic diagnosis may not be available in patients undergoing a nonoperative treatment strategy. Therefore, imaging plays an important role in diagnosis; however, conventional imaging can have limitations in distinguishing the 2 lesions.

Golden-angle radial sparse parallel (GRASP) MR imaging is a volumetric dynamic imaging technique that uses 3D gradient recalled-echo dynamic contrast-enhancement (DCE) imaging together with compressed sensing and parallel imaging to accelerate data acquisition.[10] GRASP MR imaging samples k-space radially in the xy-plane using a golden-angle scheme with Cartesian sampling in the z-plane.[10] Dynamic data are obtained in a single continuous acquisition after contrast injection and are reconstructed into sequential timeframes using an iterative method that permits excellent spatial and temporal resolution. GRASP MR imaging has been shown to have applications in abdominal, breast, and other head and neck lesions.[11-17] In particular, a recent study demonstrated the utility of GRASP MR imaging to differentiate parotid neoplasms, in particular with a high negative predictive value for malignant lesions.[18]

The purpose of this study was to evaluate the accuracy of both qualitative and quantitative metrics derived from DCE MR imaging using GRASP MR imaging to differentiate paragangliomas and schwannomas in the jugular foramen. Prior work has been done by Gaddikeri et al[19] to differentiate schwannomas and paragangliomas using DCE MR imaging, and Demerath et al[20] have applied GRASP MR imaging to the same problem, but our study focused on applying GRASP MR imaging to lesions strictly within the jugular foramen, a region not readily amenable to tissue sampling.

MATERIALS AND METHODS
Patient Selection
This retrospective study of patient data received approval from academic center's institutional review board, and the requirement for informed consent was waived. Health Insurance Portability and Accountability Act–compliant techniques for data management and analysis were used.

From August 2017 through April 2019, DCE MR imaging using the GRASP technique was performed on 35 consecutive patients with jugular foramen masses. The mean patient age was 49.5 years with a range of 26–79 years. There were 16 men and 19 women. Five patients were excluded due to the following: failed GRASP examination ($n = 2$), dural arteriovenous fistula ($n = 1$), meningioma ($n = 1$), and lymph node ($n = 1$). A total of 30 skull base lesions were ultimately included. None of the included lesions had undergone an operation at the time of imaging. The excluded dural arteriovenous fistula was not associated with a mass and was diagnosed with a conventional angiogram. The excluded lymph node was an incidental retropharyngeal lymph node, which was included in the results during the initial search for jugular foramen masses and did not have anything to do with metastatic paraganglioma.

GRASP DCE MR Imaging Technique
An in-depth description of the GRASP protocol for DCE MR imaging has been previously reported.[10] In summary, GRASP MR imaging acquires data continuously in the k-space in a radial fashion using a golden-angle scheme during contrast injection. The dynamic acquisition can be obtained to any desired length. Timeframes of any desired length can be obtained by grouping consecutive radial spokes into discrete temporal units. GRASP MR imaging can achieve high temporal resolution due to high undersampling, using both parallel imaging and compressed sensing in the reconstruction process.

Patients were scanned on both 1.5T and 3T MR imaging scanners. Each timeframe had 120 axial sections at a 2-mm section thickness. The FOV was 256×256 mm, and the matrix size was 256×256, resulting in a 1-mm in-plane resolution. Anatomic coverage comprised the whole neck from the superior orbit to the sternal notch. TR and TE were 3.65 and 1.75 ms, respectively, with a 12° flip angle. Temporal resolution was 5.5 s/frame for the first 67.3 seconds and 13.0 s/frame through the remainder of the examination for a total dynamic acquisition of 240 seconds. A single dose of gadobutrol (0.1 mL/kg, Gadavist; Bayer Schering Pharma) was administered intravenously at a rate of 3 mL/s via a power injector 20 seconds after the start of the GRASP sequence.

Qualitative and Quantitative Time-Intensity Curve Analysis
All DCE MR imaging data were postprocessed and analyzed using Olea Sphere 3.0 software (Olea Medical). The reader was blinded to the clinical history, pathologic diagnosis if available, and all prior and subsequent imaging studies. A single ROI was created within each lesion on axial images, taking care to inscribe only the enhancing component and to avoid nonenhancing portions. An additional ROI was created in a normal-appearing dominant ipsilateral or contralateral sigmoid sinus to serve as an internal reference standard, taking care to avoid mistaking enhancing tumor for the sigmoid sinus and vice versa, especially when the tumor was large. No invasion or thrombosis of the sigmoid sinus was seen. Time-intensity curves (TICs) were created for each ROI and qualitatively assigned 1 of 3 TIC classifications:[18] 1) continuously rising, 2) initial rapid wash-in with a subsequent plateau, and 3) initial rapid wash-in with subsequent washout.

Semiquantitative TIC metrics obtained for the lesion included peak signal intensity, wash-in rate, washout rate, and area under the curve (AUC), which were normalized to values obtained from the reference ROI in the sigmoid sinus. The AUC was calculated by Olea Sphere for the entire duration of the scan. A subset analysis of paragangliomas was later performed comparing those with prior radiation treatment with those with no prior radiation treatment.

Statistical Methods
Categoric data were summarized by frequencies and percentages, and continuous scaled data were summarized by the mean (SD) of the empiric distribution.

Raw data output from the Olea Sphere software included time-to-maximum enhancement (TME), AUC, wash-in rate (wash-in), washout rate (washout), peak signal intensity, extracellular extravascular space volume fraction (Ve), maximum signal-enhancement

Table 1: Quantitative summary statistics for paragangliomas and schwannomas[a]

	Paragangliomas	Schwannomas
Age (yr)	52.0 (SD, 14.8)	43.7 (SD, 16.9)
Sex (% female)	77.2%	33.3%
AUC	0.86 (SD, 0.19)	0.45 (SD, 0.15)
Peak intensity	0.86 (SD, 0.21)	0.43 (SD, 0.13)
TME	1.04 (SD, 0.24)	3.13 (SD, 0.64)
Wash-in	1.00 (SD, 0.12)	0.25 (SD, 0.13)
Washout	1.00 (SD, 0.14)	0.12 (SD, 0.18)
Ve	0.06 (SD, 0.17)	0.53 (SD, 0.25)
SER	148.1 (SD, 23.1)	61.5 (SD, 19.6)
Kep	0.46 (SD, 0.78)	0.55 (SD, 0.32)
Vp	0.47 (SD, 0.19)	0.087 (SD, 0.81)
K^{trans}	0.08 (SD, 0.22)	0.29 (SD, 0.28)

[a] Data are means unless otherwise indicated.
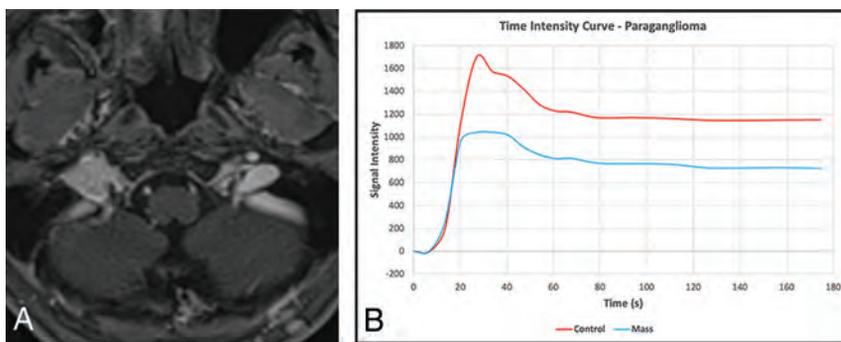
FIG 1. Example of a paraganglioma in the right jugular foramen on postcontrast imaging (*A*) with an associated time-intensity curve (*B*) demonstrating a type 3 TIC (blue) with wash-in and washout characteristics similar to those in the blood pool control obtained from the sigmoid sinus (red).
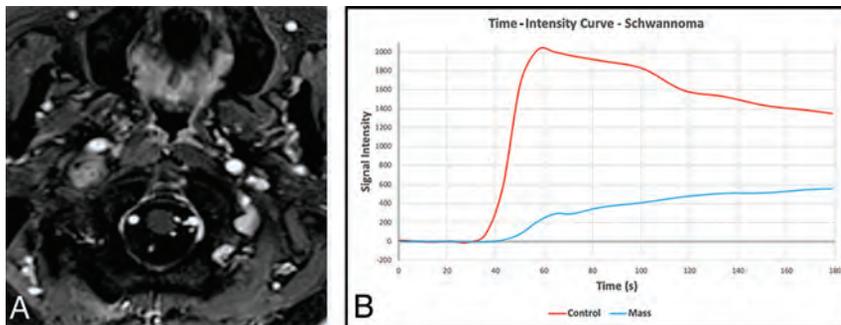
FIG 2. An example of a schwannoma in the right jugular foramen extending inferiorly into the carotid space on postcontrast imaging with an associated time-intensity curve demonstrating a type 1 TIC (blue) with gradual wash-in and a continuous gradual washout of contrast. The curve obtained from the blood pool within the sigmoid sinus (red) shows the expected rapid wash-in and washout of contrast, suggesting a good bolus of administered contrast.

ratio (SER), rate transfer constant (Kep), plasma volume fraction (Vp), and volume transfer constant (K^{trans}). AUC, TME, wash-in, washout, and peak intensity were all normalized to a control ROI placed in the sigmoid sinus.

A Student *t* test was used to determine whether the 2 sets of data containing paragangliomas and schwannomas were similar (null hypothesis) or different from each other. A 2-tailed *t* test

was used, assuming 2 sample sets of unequal variance for each parameter. Similar analysis was performed between subgroups of paragangliomas: those without prior radiation ($n = 13$) and those with prior radiation ($n = 9$).

Using SPSS (Version 16.0; IBM), we created a univariable logistic regression model with a binary output, paraganglioma or schwannoma, using wash-in as a single input variable. This univariate model was evaluated on the basis of a type III χ^2 test.

Additionally, by means of SPSS, lesions were classified on the basis of wash-in and washout using a 3-nearest neighbors (3-NN) method. This method plots all the data points according to input values, wash-in and washout in this case; then for each data point, it assigns a classification, either paraganglioma or schwannoma, based on a weighted score of its 3 nearest data points. Several studies in the field of radiology have demonstrated the utility of using classification schemes such as the nearest neighbors method to classify masses or patterns of abnormality in both the brain and lungs.[21-23]

RESULTS

There were 22 paragangliomas and 8 schwannomas. Pathologic diagnosis was available in 16 of 22 paragangliomas and 4 of 8 schwannomas. Diagnosis of the remaining lesions was made on the basis of conventional imaging features.

Table 1 demonstrates an overview of quantitative summary statistics. The mean patient age was 43.7 years for the schwannoma group and 52.0 years for the paraganglioma group without a significant difference. Thirty-three percent of the patients with schwannoma and 77% of the patients with paraganglioma were women. A χ^2 test demonstrated a significant difference between the 2 groups in terms of sex. While there is no documented sex predilection for schwannomas, there is a known female predilection for paragangliomas of the head and neck, which may explain the high proportion of women in our paraganglioma cohort.[24]

All paragangliomas demonstrated a type 3 TIC: initial rapid wash-in with subsequent washout (Fig 1), and all schwannomas demonstrated a type 1 TIC: continuously rising (Fig 2). There were no cases that demonstrated the type 2 TIC.

Results comparing the means (SDs) of quantitative DCE parameters for both paragangliomas and schwannomas are shown in Tables 2 and 3. There was a statistically significant difference between paragangliomas and schwannomas for AUC, peak intensity, TME, wash-in, washout, Ve, SER, and Vp.

Table 2: Semiquantitative perfusion metrics of jugular foramen tumors normalized to the sigmoid sinus[a]

	AUC	Peak	TME	Wash-in	Washout
Paraganglioma	0.86 (SD, 0.19)	0.86 (SD, 0.21)	1.04 (SD, 0.24)	1.00 (SD, 0.12)	1.00 (SD, 0.14)
Schwannoma	0.45 (SD, 0.15)	0.43 (SD, 0.13)	3.13 (SD, 0.64)	0.25 (SD, 0.13)	0.12 (SD, 0.18)
P value	<.001	<.001	<.001	<.001	<.001

[a] There is a statistically significant difference between paragangliomas and schwannomas when comparing their values for AUC, peak enhancement, wash-in, and washout. Of note, most paragangliomas had wash-in and washout values of ≥1.0, which means that these lesions demonstrated permeability characteristics similar to those of the blood pool because all values were normalized to the sigmoid sinus. Data are means unless otherwise indicated. (Tables 2 and 3 separated for formatting purposes).

Table 3: Kinetic permeability metrics of jugular foramen tumors[a]

	Ve	SER	Kep	Vp	K^{trans}
Paraganglioma	0.06 (SD, 0.17)	148.1 (SD, 23.1)	0.46 (SD, 0.78)	0.47 (SD, 0.19)	0.08 (SD, 0.22)
Schwannoma	0.53 (SD, 0.25)	61.5 (SD, 19.6)	0.55 (SD, 0.32)	0.087 (SD, 0.81)	0.29 (SD, 0.28)
P value	<.001	<.001	.77	<.001	.10

[a] There is a statistically significant difference between paragangliomas and schwannomas when comparing their values for VE, SER, and VP. Data are means unless otherwise indicated.
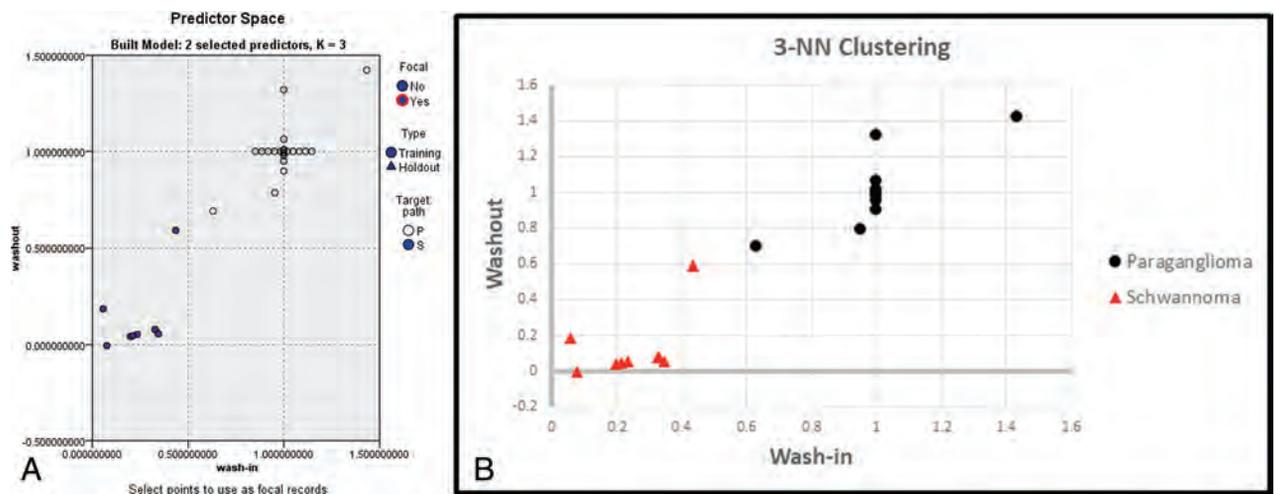
FIG 3. *A*, The predictor space model clearly shows differential group clustering for schwannomas and paragangliomas based on the wash-in and washout measures. *B*, The predictor space model clearly shows differential group clustering for schwannomas and paragangliomas based on the wash-in and washout measures. P indicates paraganglioma; S, schwannoma.

Table 4: Subgroup comparison of paragangliomas with prior radiation treatment (*n* = 9) and with no radiation treatment (*n* = 13) using normalized semiquantitative perfusion metrics[a]

	AUC	Peak	TME	Wash-in	Washout
Para-RT	0.80 (SD, 0.15)	0.84 (SD, 0.19)	1.04 (SD, 0.33)	0.96 (SD, 0.12)	1.00 (SD, 0.16)
Para-NT	0.90 (SD, 0.2)	0.87 (SD, 0.22)	1.04 (SD, 0.17)	1.03 (SD, 0.12)	1.01 (SD, 0.14)
P value	.22	.68	.97	.20	.86

Note:—NT indicates no radiation treatment; RT, radiation treatment, Para, paraganglioma.
[a] There was no statistical difference between these subcategories of paragangliomas in terms of semiquantitative perfusion parameters. Data are means unless otherwise indicated.

Table 5: Subgroup comparison of paragangliomas with prior radiation treatment (*n* = 9) and with no radiation treatment (*n* = 13) using kinetic permeability metrics[a]

	Ve	SER	Kep	Vp	K^{trans}
Para-RT	0.03 (SD, 0.07)	143.2 (SD, 14.4)	0.68 (SD, 1.03)	0.54 (SD, 0.19)	0.079 (SD, 0.14)
Para-NT	0.08 (SD, 0.21)	151.4 (SD, 27.7)	0.31 (SD, 0.54)	0.43 (SD, 0.18)	0.09 (SD, 0.27)
P value	.43	.38	.34	.21	.78

[a] There was no statistical difference between these subcategories of paragangliomas in terms of kinetic permeability parameters. Data are means unless otherwise indicated.

A univariable logistic regression model was created with a binary output (paraganglioma or schwannoma) using wash-in as a single input variable. This simple model was able to correctly predict all observed lesions ($P < .001$).

Lesions were clustered on the basis of wash-in and washout using a 3-NN method. All 30 lesions (22 paragangliomas and 8 schwannomas) were classified correctly using this method (Fig 3). Last, a subgroup comparison of paragangliomas was performed between those with prior radiation treatment and those without in Tables 4 and 5. There was no statistically significant difference between the 2 groups when comparing them across all parameters.

DISCUSSION

This study evaluated the utility of GRASP DCE to differentiate paragangliomas and schwannomas at the jugular foramen. Our work demonstrates that by means of DCE postprocessing software, there are specific TIC curves for each lesion, allowing reliable lesion differentiation based on imaging. This may, in the future, preclude the need for a biopsy in certain clinical situations.

Prior studies from Gaddikeri[19] et al and Demerath et al[20] have studied the use of DCE to differentiate these lesions, with the latter also using GRASP. However, we included a larger number of lesions compared with both of the above studies and focused on evaluating lesions only within the jugular foramen.

Our work, in concordance with the abovementioned studies, demonstrates that paragangliomas have a type 3 TIC: a sharp wash-in and peak followed by slow washout, similar to that of the venous blood pool. This is in accordance with the physiologic structure of paragangliomas, which have a rich network of highly vascularized fibrous septa and an attenuated vessel network, resulting in arteriovenous shunting.[25,26] This also explains why paragangliomas demonstrated earlier TME than schwannomas and appeared similar to the venous blood pool. Schwannomas, on the other hand, demonstrate type 1 TICs: a rapid initial wash-in phase followed by continued, albeit slower wash-in of contrast without washout. This accumulation of contrast in schwannomas has been attributed to the vascular proliferation of small, thin-walled vessels and larger, hyalinized vessels, which increases permeability, allowing contrast to progressively enter the extravascular extracellular space (EES).[27]

Our qualitative findings are corroborated by semiquantitative parameters, which were normalized to a control (dural venous sinus). AUC, peak enhancement, wash-in, and washout in paragangliomas were shown to be very similar to those in the venous blood pool, whereas in schwannomas, there was a much lower AUC and peak enhancement, as well as slower wash-in and washout velocities. Additionally, there was excellent classification of these lesions solely using wash-in and washout with a 3-NN method.

Kinetic parameters based on the Tofts model also provided insight into the different vascular structures of these 2 lesions, similar to those demonstrated by prior studies.[26] For example, the Ve, synonymous with the EES, was much lower for paragangliomas because most contrast volume in these lesions is expected to remain within the blood plasma with poor perfusion into the tumor tissue.[28] Given the high-flow, vascular nature of these lesions with arteriovenous shunting, there may not be enough time for contrast to enter the EES. Conversely, the

Vp was much lower for schwannomas because most contrast volume is located in the EES rather than in blood plasma due to a microvascular network, which more freely permits tumor tissue perfusion.

The SER was also significantly higher for paragangliomas compared with schwannomas, reflecting their marked relative contrast uptake and, therefore, larger difference between precontrast signal intensity and maximal signal intensity. The K^{trans} and Kep, the volume and rate transfer constants, respectively, between the EES and blood plasma, were not significantly different between the 2 lesions. This finding may be related to the faster wash-in and washout velocities seen with paragangliomas, which may not permit enough time for contrast to enter the EES from the blood plasma.

Comparison between paragangliomas that underwent prior radiation treatment versus those that did not yielded somewhat surprising results. There was no statistically significant difference between these 2 groups across all semiquantitative and kinetic parameters. This finding may be due to the mechanism of the action of radiation, which treats tumors by damaging DNA at the molecular level to induce cell death rather than damaging the macrostructure of a lesion, which is depicted by imaging.

One limitation of this subgroup analysis was the small sample size and follow-up. Treated paragangliomas with follow-up did not demonstrate interval growth, but only a small subset of these tumors had follow-up. Specifically, a small sample size limits how much we can confidently say about postradiation effects.

While our study demonstrates that GRASP DCE can differentiate paragangliomas and schwannomas in the jugular foramen when conventional imaging findings are not present, there are still some clear limitations. First, our sample size of schwannomas was relatively small compared with the number of paragangliomas. Additionally, histologic confirmation of pathology was not available in every case in part due to the inherent risks of an operation in this location. When pathology was not available, diagnosis was made on the basis of conventional imaging features.

CONCLUSIONS

Paragangliomas and schwannomas at the jugular foramen can frequently be distinguished on the basis of characteristic imaging features. However, occasionally lesions may be more difficult to differentiate, particularly when small, and histologic sampling in this location is particularly challenging given the critical neurovascular structures passing through this region. Our study adds further evidence that paragangliomas at the jugular foramen can be reliably differentiated from schwannomas using GRASP DCE imaging. This may, in the future, obviate the need for histologic sampling or may add increased confidence and reliability when diagnosis is made presumptively.

REFERENCES

1. Sanna M, De Donato G, Di Lella F, et al. **Nonvascular lesions of the jugular foramen: the gruppo otologico experience.** *Skull Base* 2009;19:57–57 CrossRef Medline

2. Vogl TJ, Bisdas S. **Differential diagnosis of jugular foramen lesions.** *Skull Base* 2009;19:3–16 CrossRef Medline

3. Ong CK, Fook-Hin Chong V. **Imaging of jugular foramen.** *Neuroimaging Clin N Am* 2009;19:469–82 CrossRef Medline

4. Eldevik OP, Gabrielsen TO, Jacobsen EA. **Imaging findings in schwannomas of the jugular foramen.** *AJNR Am J Neuroradiol* 2000;21:1139–44 Medline

5. Rao AB, Koeller KK, Adair CF. **From the archives of the AFIP: paragangliomas of the head and neck: radiologic-pathologic correlation—Armed Forces Institute of Pathology.** *Radiographics* 1999;19:1605–32 CrossRef Medline

6. Crist J, Hodge JR, Frick M, et al. **Magnetic resonance imaging appearance of schwannomas from head to toe: a pictorial review.** *J Clin Imaging Sci* 2017;7:38 CrossRef Medline

7. van Gils AP, van den Berg R, Falke TH, et al. **MR diagnosis of paraganglioma of the head and neck: value of contrast enhancement.** *AJR Am J Roentgenol* 1994;162:147–53 CrossRef Medline

8. Ramina R, Maniglia JJ, Fernandes YB, et al. **Jugular foramen tumors: diagnosis and treatment.** *Neurosurg Focus* 2004;17:E5 CrossRef Medline

9. Guss ZD, Batra S, Li G, et al. **Radiosurgery for glomus jugulare: history and recent progress.** *Neurosurg Focus* 2009;27:E5 CrossRef Medline

10. Feng L, Grimm R, Block KT, et al. **Golden-angle radial sparse parallel MRI: combination of compressed sensing, parallel imaging, and golden-angle radial sampling for fast and flexible dynamic volumetric MRI.** *Magn Reson Med* 2014;72:707–17 CrossRef Medline

11. Chandarana H, Feng L, Block TK, et al. **Free-breathing contrast-enhanced multiphase MRI of the liver using a combination of compressed sensing, parallel imaging, and golden-angle radial sampling.** *Invest Radiol* 2013;48:10–16 CrossRef Medline

12. Rosenkrantz AB, Geppert C, Grimm R, et al. **Dynamic contrast-enhanced MRI of the prostate with high spatiotemporal resolution using compressed sensing, parallel imaging, and continuous golden-angle radial sampling: preliminary experience.** *J Magn Reson Imaging* 2015;41:1365–73 CrossRef

13. Wu X, Raz E, Block TK, et al. **Contrast-enhanced radial 3D fat-suppressed T1-weighted gradient-recalled echo sequence versus conventional fat-suppressed contrast-enhanced T1-weighted studies of the head and neck.** *AJR Am J Roentgenol* 2014;203:883–89 CrossRef Medline

14. Bangiyev L, Raz E, Block TK, et al. **Evaluation of the orbit using contrast-enhanced radial 3D fat-suppressed T1 weighted gradient echo (Radial-VIBE) sequence.** *Br J Radiol* 2015;88:20140863 CrossRef Medline

15. Heacock L, Gao Y, Heller SL, et al. **Comparison of conventional DCE-MRI and a novel golden-angle radial multicoil compressed sensing method for the evaluation of breast lesion conspicuity.** *J Magn Reson Imaging* 2017;45:1746–52 CrossRef

16. Rossi Espagnet MC, Bangiyev L, Haber M, et al. **High-resolution DCE-MRI of the pituitary gland using radial k-space acquisition with compressed sensing reconstruction.** *AJNR Am J Neuroradiol* 2015;36:1444–49 CrossRef Medline

17. Sen R, Sen C, Pack J, et al. **Role of high-resolution dynamic contrast-enhanced MRI with golden-angle radial sparse parallel reconstruction to identify the normal pituitary gland in patients with macroadenomas.** *AJNR Am J Neuroradiol* 2017;38:1117–21 CrossRef Medline

18. Mogen JL, Block KT, Bansal NK, et al. **Dynamic contrast-enhanced MRI to differentiate parotid neoplasms using golden-angle radial sparse parallel imaging.** *AJNR Am J Neuroradiol* 2019;40:1029–36 CrossRef Medline

19. Gaddikeri S, Hippe DS, Anzai Y. **Dynamic contrast-enhanced MRI in the evaluation of carotid space paraganglioma versus schwannoma.** *J Neuroimaging* 2016;26: 618–25 CrossRef Medline

20. Demerath T, Blackham K, Anastasopoulos C, et al. **Golden-angle radial sparse parallel (GRASP) MRI differentiates head & neck paragangliomas from schwannomas.** *Magn Reson Imaging* 2020;70:73–80 CrossRef Medline

21. Cabon Y, Suehs C, Bommart S, et al. **k-nearest neighbor curves in imaging data classification.** *Front Appl Math Stat* 2019;5:1–22 CrossRef

22. Rajini N, Bhavani R. **Classification of MRI brain images using k-nearest neighbor and artificial neural network.** In: Proceedings of the International Conference on Recent Trends in Information Technology (ICRTIT), June 3-5, 2011, Chennai, India. 2011;563–68

23. Li C, Zhang S, Zhang H, et al. **Using the K-nearest neighbor algorithm for the classification of lymph node metastasis in gastric cancer.** *Comput Math Methods Med* 2012;2012:876545 CrossRef Medline

24. Kimura N, Chetty R, Capella C, et al. **Extra-adrenal paraganglioma: carotid body, jugulotympanic, vagal, laryngeal, aortico-pulmonary.** In: DeLellis RA, Lloyd RV, Heitz PU, et al, eds. *World Health Organization Classification of Tumours: Pathology and Genetics of Tumors of Endocrine Organs.* IARC Press 2004:159–61

25. Lee KY, Oh YW, Noh HJ, et al. **Extraadrenal paragangliomas of the body: imaging features.** *AJR Am J Roentgenol* 2006;187:492–504 CrossRef Medline

26. Kabadi SJ, Fatterpekar GM, Anzai Y, et al. **Dynamic contrast-enhanced MR imaging in head and neck cancer.** *Magn Reson Imaging Clin N Am* 2018;26:135–49 CrossRef Medline

27. Papiez J, Rojiani MV, Rojiani AM. **Vascular alterations in schwannoma.** *Int J Clin Exp Pathol* 2014;7:4032–38 Medline

28. Tofts PS, Brix G, Buckley DL, et al. **Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusible tracer: standardized quantities and symbols.** *J Magn Reson Imaging* 1999;10:223–32 CrossRef Medline

Correlation between Cranial Nerve Microstructural Characteristics and Vestibular Schwannoma Tumor Volume

A.M. Halawani, S. Tohyama, P.S.-P. Hung, B. Behan, M. Bernstein, S. Kalia, G. Zadeh, M. Cusimano, M. Schwartz, F. Gentili, D.J. Mikulis, N.J. Laperriere, and M. Hodaie

ABSTRACT

BACKGROUND AND PURPOSE: Vestibular schwannomas are common cerebellopontine angle tumors arising from the vestibulocochlear nerve and can result in cranial nerve dysfunction. Conventional MR imaging does not provide information that could correlate with cranial nerve compression symptoms of hearing loss or imbalance. We used multitensor tractography to evaluate the relationship between the WM microstructural properties of cranial nerves and tumor volume in a cohort of patients with vestibular schwannomas.

MATERIALS AND METHODS: A retrospective study was performed in 258 patients with vestibular schwannomas treated at the Gamma Knife clinic at Toronto Western Hospital between 2014 and 2018. 3T MR images were analyzed in 160 surgically naïve patients with unilateral vestibular schwannomas. Multitensor tractography was used to extract DTI-derived metrics (fractional anisotropy and radial, axial, and mean diffusivities of the bilateral facial and vestibulocochlear nerves [cranial nerves VII/VIII]). ROIs were placed in the transition between cisternal and intracanalicular segments, and images were analyzed using the eXtended Streamline Tractography reconstruction method. Diffusion metrics were correlated with 3D tumor volume derived from the Gamma Knife clinic.

RESULTS: DTI analyses revealed significantly higher fractional anisotropy values and a reduction in axial diffusivity, radial diffusivity, and mean diffusivity (all $P < .001$) within the affected cranial nerves VII and VIII compared with unaffected side. All specific diffusivities (axial, radial, and mean diffusivity) demonstrated an inverse correlation with tumor volume (axial, radial, and mean diffusivity, $P < .01$).

CONCLUSIONS: Multitensor tractography allows the quantification of cranial nerve VII and VIII WM microstructural alterations in patients with vestibular schwannomas. Our findings support the hypothesis that tumor volume may cause microstructural alterations of the affected cranial nerves VII and VIII. This type of advanced imaging may represent a possible avenue to correlate diffusivities with cranial nerve function.

ABBREVIATIONS: AD = axial diffusivity; CN = cranial nerve; FA = fractional anisotropy; IC = intracanalicular; MD = mean diffusivity; MTT = multitensor tractography; RD = radial diffusivity; VS = vestibular schwannoma

Vestibular schwannomas (VSs) are slow-growing tumors of myelin-forming cells that originate from the vestibulocochlear nerve. The tumor commonly originates near the vestibular ganglion at the junction of the central and peripheral myelin near the porus of the auditory meatus but can be found anywhere along the nerve from the internal auditory canal to the terminal ends of the vestibulocochlear nerve (cranial nerve [CN] VIII) within the vestibule, cochlea, or semicircular canals.[1-4] VSs typically grow within a capsule and remain peripherally attached to the parent vestibulocochlear nerve, which is splayed over the surface of the tumor.[5,6] As these tumors grow, they can affect hearing status and balance, resulting in asymmetric hearing loss, tinnitus, and dizziness. Further growth can result in mass effect on the facial nerve, trigeminal nerve, or, occasionally, the brain stem, resulting in variable symptoms.

DWI is a neuroimaging technique focused on the measurement of tissue water diffusion by applying special diffusion-encoding gradients.[7] In white matter tracts and CNs, the architecture of the axonal bundles with their myelin sheath facilitates water molecule motion along rather than transverse to the direction of the bundle. When multiple diffusion gradients are applied, the diffusion tensor within each voxel can be calculated, allowing visualization of CNs including CNs VII and VIII.[8] CNs VII and VIII will be, hereafter, grouped together as 1 entity—the facial/vestibular complex (CN VII/VIII)—because both fibers tend to be closely aligned with one another in diffusion-weighted scans.[9] DTI has been investigated for studying the structural connectivity and the quantitative assessment of diffusion restriction as a marker of the neuronal microstructural integrity.[10,11] Single-tensor diffusion tractography assumes that in each voxel, there is 1 representative fiber orientation. This assumption does not hold in areas of high crossing fibers where the orientation of the fibers is not unique and multiple orientations may exist simultaneously. Multitensor tractography (MTT) imaging can estimate ≥2 tensors within a voxel and improves fiber orientation in these regions. Behan et al[9] reported that the eXtended Streamline Tractography reconstruction method is effective in producing more detailed projections of CN VII/VIII than other reconstruction methods. Other publications have proved the efficacy of DTI on CN VII/VIII tractography[8,12-16] and preoperative facial nerve identification.[17-28]

The role of DTI in the study of CNs is increasing,[29-31] and WM microstructural alternations of the auditory neural pathway in patients with VS have been evaluated,[32,33] though not the cisternal segment of the vestibulocochlear nerve, chiefly due to the technical challenges this represents.

This study aimed to identify the microstructural alternations of the cisternal vestibulocochlear nerve and its association with tumor volume in patients with VS. Because a large VS can cause mechanical compression to surrounding structures, we hypothesized that the tumor can cause microstructural alternations along the distal cisternal/intracanalicular (IC) vestibulocochlear nerve and that tumor volume will correlate with microstructural alternations in the affected CN VII/VIII.

MATERIALS AND METHODS
Participants
A retrospective chart review was performed for all patients with VS treated with gamma knife radiosurgery between January 1, 2014, and December 31, 2018, after approval by the institutional review board at Toronto Western Hospital. Overall, 258 patients with VS were enrolled. After we excluded patients with bilateral VSs (9 patients), prior surgical treatment (partial surgical resection or gamma knife radiosurgery, 23 patients), nonavailability of DTI studies (52 patients), and additional technical difficulties related to eddy current and motion correction (12 patients), a total of 162 patients were enrolled. Two other patients with outliers were identified (one had a huge tumor volume compared with other subjects, and the other had marked motion on the clinical DWI sequence), both resulting in improper tract reconstruction, and were also excluded. Thus, a total of 160 patients

were included in our DTI analysis. The analysis focused solely on images obtained before gamma knife radiosurgery treatment.

Tumor Volume
We evaluated tumor size using volumetric measurements. The semiautomatic segmentation function was used with 1-mm-section FIESTA MR images for volumetric measurements. Manual contouring was performed by a trained assistant to delineate the tumor and measure the tumor volume. The tumor volume was measured in milliliters (mL). In all subjects, treatment with gamma knife radiosurgery commenced within 21 days after MR images were acquired.

Image Acquisition
Images were acquired with 3T MR imaging (Signa; GE Healthcare) with an 8-channel phased-array head coil following our routine clinical imaging protocol. DWI scans (60 directions, spin-echo EPI sequence, 1 B_0, $b = 1000$ s/mm^2, one excitation, array spatial sensitivity encoding technique, voxel size $= 0.94 \times 0.94 \times 3$ mm^3, matrix $= 256 \times 256$, TR $= 12,000$ ms, TE $= 86.4$ ms, flip angle $= 90°$, FOV $= 24 \times 24$ cm) were acquired. High-resolution 3D T1-weighted anatomic fast-spoiled gradient recalled and T2-weighted FIESTA scans were obtained for coregistration with MTT images using an inversion recovery fast-spoiled gradient recalled sequence.

Image Processing and Registration
Following eddy current and motion correction with FSL (http://www.fmrib.ox.ac.uk/fsl), the diffusion tensor estimation and derivation of scalar maps (ie, fractional anisotropy [FA], mean diffusivity [MD], radial diffusivity [RD], and axial diffusivity [AD]) were extracted manually using 3D Slicer, Version 4.3 (National Alliance for Medical Image Computing; http://www.slicer.org).[34] The T1-weighted anatomic image was linearly coregistered on the DTI examination with respect to the following landmarks: basilar artery, vertebral arteries, ventral brain stem, tumor configuration, internal auditory canals, fourth ventricle, and cerebellar surface.

Microstructural Analysis
ROIs were placed on the proximal IC segments on both sides where the CNs can be most easily identified anatomically. All seeds were mapped to incorporate a cross-section of the CN of interest and did not exceed 25 mm^2. The ROIs were reconstructed by eXtended Streamline Tractography methods, traced CN VII/VIII tracts were further analyzed on the basis of fiber orientation, and then diffusivity measurements were evaluated. ROIs were used for both tractography construction and diffusivity measurements. The reconstructed tracts were examined in a blinded fashion by a subcommittee of the authors (S.T., P.S.-P.H., and A.M.H.). Two patients with improper tract reconstruction were excluded by consensus among this group.

We defined the CN "vulnerable zone" as the transition between the cisternal and IC segments, which is an anatomic area with increased mechanical compression by the tumor. For the unaffected side, we defined the same vulnerable zone where the nerves were seen to travel in close proximity and were difficult to differentiate anatomically (Fig 1).
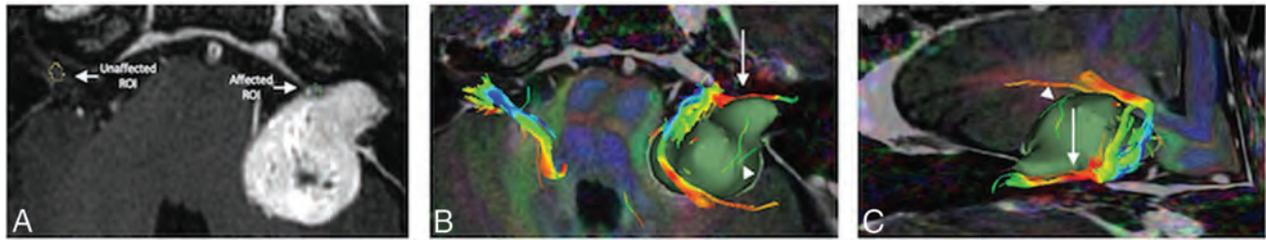
FIG 1. TI-weighted anatomic image (*A*) and 2 fused DTI/TI-weighted images (*B* and *C*) with superimposed representation of tumor volume (green) used to locate the ROIs within the distal cisternal/IC segments (*A*) and to demonstrate the extracted CN VII/VIII tracts (*B* and *C*). Note the detailed correspondence of the tumor and its effect on CN VII/VIII (*white arrows*). The CN VII/VIII can be seen sweeping anterior to the tumor. Also, note scattered, disconnected fibers that are not corresponding to CN configuration in space. These fibers were disregarded anatomically (*arrowheads*).

Criteria of CN Tractography Assessment

The criteria of CN tractography assessment were the following:

1) Anatomic accuracy: we ensured that there was an appropriate representation of the generated tracts in correlation with the CNs on the anatomic T1 image (fast-spoiled gradient recalled) and T2 image (FIESTA) sequences.

2) Delineation of the tumor mass effect on the CNs: special consideration was given to tumor fiber contamination, especially in large tumors where it was difficult to identify the exact anatomy of the affected CNs.

It may be difficult to differentiate the CN VII/VIII on the unaffected side, given the small size of the nerves and their close proximity to each other. Conventional T1-weighted images were used as a guide to identify the orientation and course of both nerves, and tractography reconstruction was easier to obtain compared with the affected side. On the affected side, the ROI was placed at our defined vulnerable zone, where the tumor caused the maximum compression on the nerve against adjacent structures. This placement provides the advantage of a consistent ROI definition for every patient. We avoided an intracanalicular ROI definition because the canal can be narrow with surrounding bone.

Statistical Analysis

All analyses were conducted in SPSS, Version 22.0 (IBM). Multiple comparisons were corrected for all 4 diffusivity metrics for each separate analysis using false discovery rate correction. Statistical significance was determined at $P < .01$ or false discovery rate–corrected P value (q) $< .01$ when adjustments were necessary for multiple comparisons. The average microstructural diffusivity differences between the affected and unaffected CN VII/VIII for each patient were evaluated using paired-sample *t* tests. Continuous variables were summarized as mean (SD). We used the Kolmogorov-Smirnov test to assess normality and the Spearman rank correlation to assess the relationship between DTI metrics and tumor volume.

RESULTS

Clinical Characteristics

We identified 160 surgically naïve patients with unilateral VSs (68 men and 92 women; mean age, 64.2 [SD, 13] years; range, 28–94 years; median, 64.1 years). On clinical assessment, 79 patients (32 men and 47 women) had both tinnitus and sensorineural

hearing loss, 33 patients (15 men and 18 women) had isolated sensorineural hearing loss, 7 patients (6 men and 1 woman) had isolated tinnitus, 3 patients (1 man and 2 women) had sensorineural hearing loss and trigeminal neuralgia pain, and 1 patient (1 woman) had isolated trigeminal neuralgia pain. Unfortunately, we could not access the electronic health records of 37 patients.

Tumor Morphology

Tumor volume ranged from 0.2 to 9.5 mL (mean, 2.3 [SD, 0.6] mL; median, 1.3 mL). Sixty-nine patients had a right-sided tumor, and 91 patients had a left-sided tumor. In all 160 patients, 17 patients had a VS primarily in the cerebellopontine angle, 10 patients had an isolated intracanalicular vestibular schwannoma, and 133 patients had both cerebellopontine angle and IC tumors.

On conventional imaging, the tumor resulted in a significant disruption of the proximal cisternal CN VIII nerve, and it was difficult to trace along the bulk of tumor. The CN VIII fibers distal to the tumor were visualized occasionally. The morphologic assessment is challenging in large tumors with brain stem compression, due to significant anatomic distortion. On MTT, most of the extracted CN VII/VIII tracts had an anterior or anteroinferior relationship to the tumor; the fibers were seen sweeping anterior to the tumor, resulting in color changes on the FA map (Fig 1).

Diffusion Metrics Changes

We found a statistically significant increase in the average FA ($P < .001$, q $< .001$) and significant reduction in the average specific diffusivities (AD: $P < .001$, q $< .001$; RD: $P < .001$, q $< .001$; and MD: $P < .001$, q $< .001$) in all patients with VS at the affected CN VII/VIII compared with the unaffected side (Fig 2). The quantitative summary data of FA and diffusivities are summarized in the Table.

Correlation Analysis

The results of the Pearson correlation coefficient test and the Spearman ρ showed a small-but-significant negative linear relationship between the tumor volume and the specific diffusivities: AD ($r_s = 158 = -0.3$, $P = .001$, q $= 0.004$), RD ($r_s = 158 = -0.24$, $P = .004$, q $= 0.008$), and MD ($r_s = 158 = -0.23$, $P = .007$, q $= 0.009$). FA values did not show a statistically significant correlation
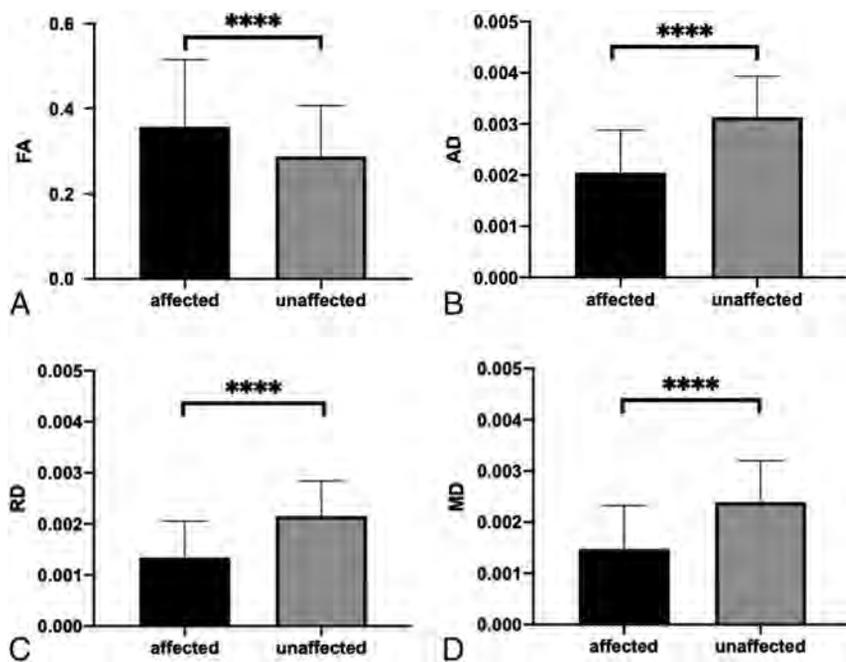
FIG 2. Microstructural properties of the CN VII/VIII in patients with VS. The averaged mean value of the diffusivity measurements in all patients with VS that demonstrate abnormal WM microstructural changes in their VII/VIII nerves (*four asterisks* indicate $P < .0001$ false discovery rate–corrected). DTI-derived values $\pm$ standard error of the mean of affected (*black bars*) and unaffected (*gray bars*) sides. Compared with the unaffected side, patients had significantly higher FA (*A*) and lower AD (*B*), RD (*C*), and MD (*D*) on their affected side. Diffusion is in square millimeters/second.

The mean, median, and range values for averaged FA and diffusivity

	Mean	Median	Range
FA (mm²/s)			
Affected	0.357	0.332	0.108–0.722
Nonaffected	0.287	0.256	0.114–0.746
AD (mm²/s)			
Affected	0.002	0.002	0.001–0.006
Nonaffected	0.003	0.003	0.001–0.005
RD (mm²/s)			
Affected	0.001	0.001	0–0.004
Nonaffected	0.002	0.002	0–0.003
MD (mm²/s)			
Affected	0.001	0.001	0–0.005
Nonaffected	0.002	0.002	0–0.004

with the tumor volume ($r_s = 158 = -0.02$, $P = .870$, q = 0.87) (Fig 3).

DISCUSSION

In the present study, we defined the WM microstructural alteration on the cisternal segment of CN VII/VIII in surgically naïve patients with VS. This study is the first to describe the WM microstructural abnormality characterized by higher FA and lower AD, RD, and MD in the vulnerable zone of the affected side compared with the unaffected side. Furthermore, all specific diffusivities (AD, RD, and MD) demonstrated an inverse correlation with tumor volume but not FA. These results are in line with

our hypothesis that tumors cause microstructural changes which are, in turn, related to tumor volume. While these metrics are not obtained from conventional MR images, MTT is a sensitive tool that provides quantitative diffusion measures and tissue fiber orientation to help develop a better understanding of hearing impairment pathophysiology in patients with VS. It may provide a useful surrogate to measure the alterations in the CN VII/VIII as a result of tumor compression.

The quantitative diffusivity metrics are calculated from the tensors. These maps are generated by mapping the directional components of the major eigenvectors in the $\lambda 1$, $\lambda 2$, and $\lambda 3$ planes. AD indicates diffusion along with the main axis of the ellipsoid, RD is a measure of diffusion along the other 2 orthogonal directions, and MD is equivalent to the average of the eigenvalues.[35,36] These metrics (AD, RD, and MD) have been linked to pathophysiologic mechanisms such as axonal degeneration,[37] demyelination,[38,39] and neuroinflammation.[35]

Because AD is sensitive to axonal integrity, the present study showed a lower AD value in the cisternal segment of the affected vestibulocochlear nerve. This may relate to axonal damage secondary to Wallerian degeneration,[40-42] focal misalignment of neurofilaments, and/or beading in injured axons, which may also result in the reduction of MD.[43-46] Alternatively, the lower AD value can be seen in the assumed increased tumor cellular density in the extracellular matrix.[47,48]

The increase in FA and reduction of MD may be other reliable indicators of damaged fiber integrity, but these are dependent on the brain region and cellular basis.[49,50] Several factors, such as an increase in myelination, microscopic deficits of axonal structures, decreases in axonal diameter, and packing density may contribute to the higher FA.

Although in previous studies, the RD increased in cases of dysmyelination or demyelination,[38,51] we report a reduction on RD value in the affected CN VII/VIII. As the RD indicates diffusion along the two minor orthogonal directions of the traced nerve bundles, we assume that the lower value found in our study may represent a reduction in this cross-sectional area secondary to compression and/or nerve bundle atrophy.

We found that a small inverse relationship exists between the tumor volume and WM microstructural alteration, with a significant degree of agreement. This relationship supports our hypothesis that the tumor volume may cause microstructural alterations in the vulnerable zone of the ipsilateral CN VII/VIII, and it seems to be in line with past studies that confirmed that nerve compression is not the only cause of hearing loss in patients with VS.[52]
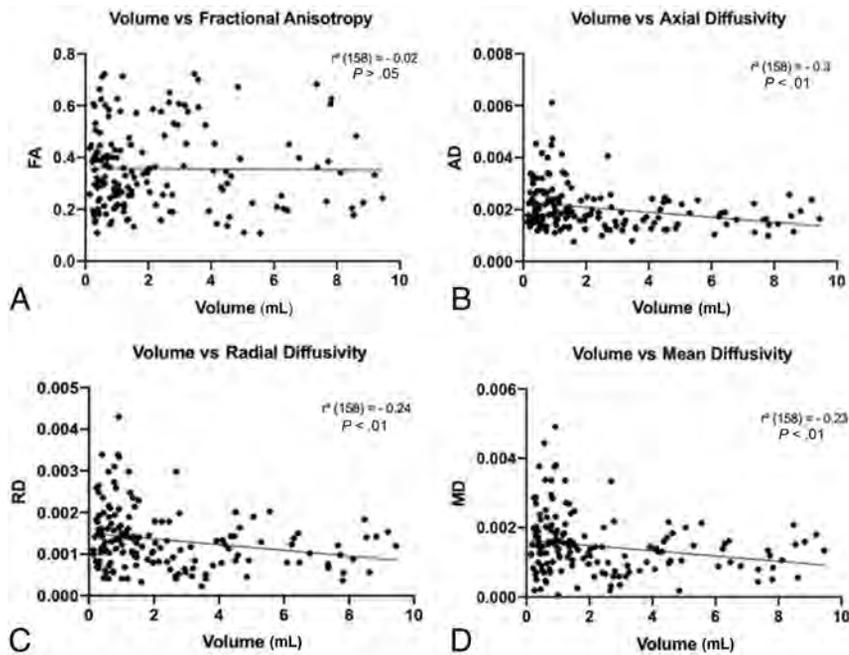
FIG 3. Scatterplots of the relationship between tumor volume and microstructural properties in the affected CNs VII/VIII. A weak-but-significant inverse relationship exists between the tumor volume and the specific diffusivities: axial diffusivity (*B*), radial diffusivity (*C*), and mean diffusivity (*D*), but not with FA (*A*) (significant at $P < .01$; false discovery rate–corrected). Diffusion is in square millimeters/second.

There are some limitations to the current study. The retrospective design made it difficult to acquire patient charts and record all clinical variables. The voxel size for the DWI scan is not equal across all 3 dimensions and is still considered larger than the dimensions of the axons, which may also be inhomogeneous due to tumor distortion. We assume this might be the cause of these variable metric abnormalities. In the future, further studies may focus on WM microstructural alterations along the auditory neural pathway aiming for better understanding and explanation of hearing impairment in this patient group using this noninvasive neuroimaging tool. Moreover, this study paves the way for more studies on posttreatment response prediction to aid in high-risk group identification and clinical prognostication.

CONCLUSIONS

This was a large cohort study that defined the WM microstructural signature of the cisternal CN VII/VIII in surgically naïve patients with VS. We evaluated microstructural alternations in correlation to tumor size volume in the affected vestibulocochlear nerve. Our observations indicate that MTT is a powerful, noninvasive tool that may play a role in the future in treatment planning and clinical decision-making.

ACKNOWLEDGMENTS

REFERENCES

1. Karlan MS, Basek M, Potter GB. **Intracochlear neurilemoma.** *Arch Otolaryngol* 1972;96:573–75 CrossRef Medline
2. Hoshino T, Ishii D. **Intralabyrinthine neurilemmoma.** *ORL J Otorhinolaryngol Relat Spec* 1972;34:117–23 CrossRef Medline
3. Kennedy RJ, Shelton C, Salzman KL, et al. **Intralabyrinthine schwannomas: diagnosis, management, and a new classification system.** *Otol Neurotol* 2004;25:160–67 CrossRef Medline
4. Salzman KL, Childs AM, Davidson HC, et al. **Intralabyrinthine schwannomas: imaging diagnosis and classification.** *AJNR Am J Neuroradiol* 2012;33:104–09 CrossRef Medline
5. Wippold FJ, Lubner M, Perrin RJ, et al. **Neuropathology for the neuroradiologist: Antoni A and Antoni B tissue patterns.** *AJNR Am J Neuroradiol* 2007;28:1633–38 CrossRef Medline
6. Coons SW. **Pathology of schwannomas of the eighth cranial nerve.** *Operative Techniques in Neurosurgery* 2001;4:53–57 CrossRef
7. Johansen-Berg H, Rushworth MF. **Using diffusion imaging to study human connectional anatomy.** *Annu Rev Neurosci* 2009;32:75–94 CrossRef Medline
8. Chen DQ, Quan J, Guha A, et al. **Three-dimensional in vivo modeling of vestibular schwannomas and surrounding cranial nerves with diffusion imaging tractography.** *Neurosurgery* 2011;68:1077–83 CrossRef Medline
9. Behan B, Chen DQ, Sammartino F, et al. **Comparison of diffusion-weighted MRI reconstruction methods for visualization of cranial nerves in posterior fossa surgery.** *Front Neurosci* 2017;11:554 CrossRef Medline
10. Mori S, Zhang J. **Principles of diffusion tensor imaging and its applications to basic neuroscience research.** *Neuron* 2006;51:527–39 CrossRef Medline
11. Kumar M, Rathore RK, Srivastava A, et al. **Correlation of diffusion tensor imaging metrics with neurocognitive function in Chiari 1 malformation.** *World Neurosurg* 2011;76:189–94 CrossRef Medline
12. Hodaie M, Quan J, Chen DQ. **In vivo visualization of cranial nerve pathways in humans using diffusion-based tractography.** *Neurosurgery* 2010;66:788–96 CrossRef Medline
13. Zolal A, Juratli TA, Podlesek D, et al. **Probabilistic tractography of the cranial nerves in vestibular schwannoma.** *World Neurosurg* 2017;107:47–53 CrossRef Medline
14. Yoshino M, Kin T, Ito A, et al. **Diffusion tensor tractography of normal facial and vestibulocochlear nerves.** *Int J Comput Assist Radiol Surg* 2015;10:383–92 CrossRef Medline
15. Yoshino M, Abhinav K, Yeh FC, et al. **Visualization of cranial nerves using high-definition fiber tractography.** *Neurosurgery* 2016;79:146–65 CrossRef Medline
16. Cauley KA, Filippi CG. **Diffusion-tensor imaging of small nerve bundles: cranial nerves, peripheral nerves, distal spinal cord, and lumbar nerve roots—clinical applications.** *AJR Am J Roentgenol* 2013;201:326–35 CrossRef Medline

17. Yoshino M, Kin T, Ito A, et al. **Combined use of diffusion tensor tractography and multifused contrast-enhanced FIESTA for predicting facial and cochlear nerve positions in relation to vestibular schwannoma.** *J Neurosurg* 2015;123:1480–88 CrossRef Medline

18. Mastronardi L, Cacciotti G, Roperto R, et al. **Position and course of facial nerve and postoperative facial nerve results in vestibular schwannoma microsurgery.** *World Neurosurg* 2016;94:174–80 CrossRef Medline

19. Zhang Y, Mao Z, Wei P, et al. **Preoperative prediction of location and shape of facial nerve in patients with large vestibular schwannomas using diffusion tensor imaging–based fiber tracking.** *World Neurosurg* 2017;99:70–78 CrossRef Medline

20. Choi KS, Kim MS, Kwon HG, et al. **Preoperative identification of facial nerve in vestibular schwannomas surgery using diffusion tensor tractography.** *J Korean Neurosurg Soc* 2014;56:11–15 CrossRef Medline

21. Gerganov VM, Giordano M, Samii M, et al. **Diffusion tensor imaging-based fiber tracking for prediction of the position of the facial nerve in relation to large vestibular schwannomas: clinical article.** *J Neurosurg* 2011;115:1087–93 CrossRef Medline

22. Song F, Hou Y, Sun G, et al. **In vivo visualization of the facial nerve in patients with acoustic neuroma using diffusion tensor imaging-based fiber tracking.** *J Neurosurg* 2016;125:787–94 CrossRef Medline

23. Li H, Wang L, Hao S, et al. **Identification of the facial nerve in relation to vestibular schwannoma using preoperative diffusion tensor tractography and intraoperative tractography-integrated neuronavigation system.** *World Neurosurg* 2017;107:669–77 CrossRef Medline

24. Churi ON, Gupta S, Misra BK. **Correlation of preoperative cranial nerve diffusion tensor tractography with intraoperative findings in surgery of cerebellopontine angle tumors.** *World Neurosurg* 2019;127:e509–16 CrossRef Medline

25. d'Almeida GN, Marques LS, Escada P, et al. **Diffusion tensor tractography in the preoperative precise identification of the course of facial nerve in a meningioma of the cerebellopontine angle: technical implications.** *Interdiscip Neurosurg* 2017;9:58–60 CrossRef

26. Taoka T, Hirabayashi H, Nakagawa H, et al. **Displacement of the facial nerve course by vestibular schwannoma: preoperative visualization using diffusion tensor tractography.** *J Magn Reson Imaging* 2006;24:1005–10 CrossRef Medline

27. Yoshino M, Kin T, Ito A, et al. **Feasibility of diffusion tensor tractography for preoperative prediction of the location of the facial and vestibulocochlear nerves in relation to vestibular schwannoma.** *Acta Neurochir (Wien)* 2015;157:939–46 CrossRef Medline

28. Wei PH, Qi ZG, Chen G, et al. **Identification of cranial nerves near large vestibular schwannomas using superselective diffusion tensor tractography: experience with 23 cases.** *Acta Neurochir (Wien)* 2015;157:1239–49 CrossRef Medline

29. Shapey J, Vos SB, Vercauteren T, et al. **Clinical applications for diffusion MRI and tractography of cranial nerves within the posterior fossa: a systematic review.** *Front Neurosci* 2019;13:23 CrossRef Medline

30. Jacquesson T, Frindel C, Kocevar G, et al. **Overcoming challenges of cranial nerve tractography: a targeted review.** *Clin Neurosurg* 2019;84:313–25 CrossRef Medline

31. Ma J, Su S, Yue S, et al. **Preoperative visualization of cranial nerves in skull base tumor surgery using diffusion tensor imaging technology.** *Turk Neurosurg* 2016;26:805–12 CrossRef Medline

32. Rueckriegel SM, Homola GA, Hummel M, et al. **Probabilistic fiber-tracking reveals degeneration of the contralateral auditory pathway in patients with vestibular schwannoma.** *AJNR Am J Neuroradiol* 2016;37:1610–16 CrossRef Medline

33. Kurtcan S, Hatiboglu MA, Alkan A, et al. **Evaluation of auditory pathways using DTI in patients treated with gamma knife radiosurgery**

for acoustic neuroma: a preliminary report. *Clin Neuroradiol* 2018;28:377–83 CrossRef Medline

34. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. **3D slicer as an image computing platform for the quantitative imaging network.** *Magn Reson Imaging* 2012;30:1323–41 CrossRef Medline

35. Alexander AL, Lee JE, Lazar M, et al. **Diffusion tensor imaging of the brain.** *Neurotherapeutics* 2007;4:316–29 CrossRef Medline

36. Beppu T, Inoue T, Shibata Y, et al. **Measurement of fractional anisotropy using diffusion tensor MRI in supratentorial astrocytic tumors.** *J Neurooncol* 2003;63:109–16 CrossRef Medline

37. Song SK, Sun SW, Ju WK, et al. **Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia.** *Neuroimage* 2003;20:1714–22 CrossRef Medline

38. Song SK, Yoshino J, Le TQ, et al. **Demyelination increases radial diffusivity in corpus callosum of mouse brain.** *Neuroimage* 2005;26:132–40 CrossRef Medline

39. Song SK, Sun SW, Ramsbottom MJ, et al. **Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water.** *Neuroimage* 2002;17:1429–36 CrossRef Medline

40. Wakana S, Jiang H, Nagae-Poetscher LM, et al. **Fiber tract-based atlas of human white matter anatomy.** *Radiology* 2004;230:77–87 CrossRef Medline

41. Gao W, Lin W, Chen Y, et al. **Temporal and spatial development of axonal maturation and myelination of white matter in the developing brain.** *AJNR Am J Neuroradiol* 2009;30:290–96 CrossRef Medline

42. Pierpaoli C, Barnett A, Pajevic S, et al. **Water diffusion changes in Wallerian degeneration and their dependence on white matter architecture.** *Neuroimage* 2001;13(6 Pt 5):1174–85 CrossRef Medline

43. Arfanakis K, Haughton VM, Carew JD, et al. **Diffusion tensor MR imaging in diffuse axonal injury.** *AJNR Am J Neuroradiol* 2002;23:794–802 Medline

44. Smith PM, Jeffery ND. **Histological and ultrastructural analysis of white matter damage after naturally-occurring spinal cord injury.** *Brain Pathol* 2006;16:99–109 CrossRef Medline

45. Budde MD, Frank JA. **Neurite beading is sufficient to decrease the apparent diffusion coefficient after ischemic stroke.** *Proc Natl Acad Sci U S A* 2010;107:14472–77 CrossRef Medline

46. Xie M, Wang Q, Wu TH, et al. **Delayed axonal degeneration in slow Wallerian degeneration mutant mice detected using diffusion tensor imaging.** *Neuroscience* 2011;197:339–47 CrossRef Medline

47. Lober RM, Cho YJ, Tang Y, et al. **Diffusion-weighted MRI derived apparent diffusion coefficient identifies prognostically distinct subgroups of pediatric diffuse intrinsic pontine glioma.** *J Neurooncol* 2014;117:175–82 CrossRef Medline

48. Wagner MW, Bell WR, Kern J, et al. **Diffusion tensor imaging suggests extrapontine extension of pediatric diffuse intrinsic pontine gliomas.** *Eur J Radiol* 2016;85:700–06 CrossRef Medline

49. Hoeft F, Barnea-Goraly N, Haas BW, et al. **More is not always better: Increased fractional anisotropy of superior longitudinal fasciculus associated with poor visuospatial abilities in Williams syndrome.** *J Neurosci* 2007;27:11960–65 CrossRef Medline

50. Thomason ME, Thompson PM. **Diffusion imaging, white matter, and psychopathology.** *Annu Rev Clin Psychol* 2011;7:63–85 CrossRef Medline

51. Budde MD, Kim JH, Liang H-F, et al. **Axonal injury detected by in vivo diffusion tensor imaging correlates with neurological disability in a mouse model of multiple sclerosis.** *NMR Biomed* 2008;21:589–97 CrossRef Medline

52. Nadol JB, Diamond PF, Thornton AR. **Correlation of hearing loss and radiologic dimensions of vestibular schwannomas (acoustic neuromas).** *Am J Otol* 1996;17:312–16 Medline

A Radiologic Grading System for Assessing the Radiographic Outcome of Treatment in Lymphatic and Lymphatic-Venous Malformations of the Head and Neck

R. De Leacy, D.V. Bageac, S. Manna, B.S. Gershon, D. Kirke, T. Shigematsu, C. Sinclair, D. Chada, P. Som, A. Doshi, K. Nael, and A. Berenstein

ABSTRACT

BACKGROUND AND PURPOSE: Two-thirds of lymphatic malformations in children are found in the head and neck. Although conventionally managed through surgical resection, percutaneous sclerotherapy has gained popularity. No reproducible grading system has been designed to compare sclerotherapy outcomes on the basis of radiologic findings. We propose an MR imaging–based grading scale to assess the response to sclerotherapy and present an evaluation of its interrater reliability.

MATERIALS AND METHODS: A grading system was developed to stratify treatment outcomes on the basis of interval changes observed on MR imaging. By means of this system, 56 consecutive cases from our institution with formally diagnosed head and neck lymphatic malformations treated by sclerotherapy were retrospectively graded. Each patient underwent pre- and posttreatment MR imaging. Each study was evaluated by 3 experienced neuroradiologists. Interrater reliability was assessed using the Krippendorff α statistic, intraclass coefficient, and 2-way Spearman ρ correlation.

RESULTS: The overall Krippendorff α statistic was 0.93 (95% CI, 0.89–0.95), denoting excellent agreement among raters. Intraclass coefficients with respect to consistency and absolute agreements were both 0.97 (95% CI, 0.96–0.98), illustrating low variability. Every combination of individual rater pairs demonstrated statistically significant ($P < .01$) linear Spearman ρ correlations, with values ranging from 0.90 to 0.95.

CONCLUSIONS: The proposed radiographic grading scale demonstrates excellent interrater reliability. Adoption of this new scale can standardize reported outcomes following sclerotherapy for head and neck lymphatic malformation and may aid in the investigation of future questions regarding optimal management of these lesions.

ABBREVIATIONS: BDL = Berenstein-De Leacy; LM = lymphatic malformation; LVM = lymphatic-venous malformation

Lymphatic malformations (LMs) and lymphatic-venous malformations (LVMs) are low-flow vascular malformations that arise as a result of erroneous vascular development during embryogenesis.[1,2] Both malformations are characterized by distended lymphatic channels, with additional anomalous venous channels present in LVMs. LMs and LVMs reflect 2 distinct classifications; however, clinical differentiation of these subtypes is lacking in much of the literature.

LMs most commonly present in childhood, with an incidence of 1/20,000 children admitted to the hospital compared with 1/100,000 adults admitted.[3] They are unlikely to regress spontaneously and demonstrate growth proportionate to body size.[4] Most of these lesions are diagnosed before 2 years of age, at which point 90% come to attention due symptoms including cosmetic disfigurement, recurrent infection, bleeding, or compression of adjacent structures.[3,5] They represent approximately 5% of benign tumors in infants and children and are located in the head and neck in 66% of cases.[3,6]

Despite nuanced etiologic and pathologic distinctions, treatment approaches to LMs and LVMs are currently similar; thus, the 2 will hereafter be considered as a grouped category (LM-LVMs). Surgical excision of LM-LVMs in the head and neck has proved challenging due to the close proximity of the lesions to vital structures, which often leads to subtotal resection and future recurrence.[7] Particularly in the case of lesions located around the face and upper airway, total resection has the potential to cause

Table 1: Grading system for assessing the radiographic outcome of treatment in LM-LVM malformations of the head and neck

Grade	Description
1	Complete regression of the lesion on cross-sectional imaging
2	Near-complete regression with trace residual of the lesion on cross-sectional imaging
3	Partial regression with <50% estimated volume of residual malformation
4	Partial regression with >50% estimated volume of residual malformation
5	Minimal or no gross interval change
6	Regression of malformation in 1 region with progression into a previously uninvolved/untreated area
7	Gross interval progression
Modifier B	Granulation tissue formation in treatment bed

pronounced deformity and/or functional impairment (respiratory, digestive, and neurologic).[8] As a result, percutaneous sclerotherapy has emerged as a common treatment technique for both macro- and microcystic LM-LVMs in these regions.

Despite the widespread treatment of head and neck LM-LVMs, a standardized grading scale for the assessment of the results is lacking. The radiographic criteria for the evaluation of outcomes have not yet been developed. In 1995, a preoperative staging scale for LMs was proposed by de Serres et al[9] to predict the prognosis and outcome of surgical intervention on the basis of lesion location. More recently, Balakrishnan et al[10] developed a consensus statement recommending standardized clinical outcome measures for studies evaluating the treatment of head and neck LMs. Only one of these measures, lesion volume, is based on radiographic evaluation.

A reliable and reproducible grading system for radiographic treatment outcomes of head and neck LM-LVMs offers the ability to both refine reporting standards and clarify communication between treating physicians. We propose 1 such system based on contrast-enhanced MR imaging and evaluate its interrater reliability in a cohort of patients treated for head and neck LM-LVM with percutaneous sclerotherapy. The soft-tissue detail, absence of ionizing radiation, safety profile, and ubiquity of MR imaging make it an ideal technique on which to develop imaging-based criteria.

MATERIALS AND METHODS
Patient Selection and Treatment
This study was approved by an institutional review board (Mount Sinai Hospital, New York). We retrospectively reviewed 56 cases of fluoroscopically guided percutaneous sclerotherapy for head and neck LM-LVMs from 2005 to 2019. Lesions were initially diagnosed using MR imaging and/or sonography in conjunction with the clinical examination. LMs and LVMs were differentiated by the presence of enhancement on postcontrast T1 sequences. The de Serres stage and lesion architecture (macrocystic, microcystic, or mixed) were determined by preprocedural MR imaging. All patients underwent pre- and posttreatment MR imaging, and patients of all ages were included. Sclerosants used during therapy included bleomycin and doxycycline. To increase the sample of patients with findings representative of lesion progression, we included 5 patients who did not undergo sclerotherapy between MR imaging time points. Raters were blinded to the treatment status of all patients.

The LM-LVM response to therapy was evaluated via the comparison of pre- and postprocedural MR imaging. Change in lesion size was evaluated on axial T2-weighted fat-saturated sequences and in orthogonal planes when available. In the case of LVMs and for the evaluation of granulation tissue response, contrast-enhanced fat-saturated T1 sequences were used.

Imaging Protocol
All patients underwent MR imaging on a 3T Magnetom Skyra MR imaging system (Siemens). A combination of a 12-element head and neck coil was used for radiofrequency signal reception. MR imaging protocol included T1 (TR/TE = 530/17 ms, flip angle = 150°, voxel = 0.6 × 0.6 × 5 [section] mm, axial plane), T2-fat saturation (TR/TE = 3600/90 ms, flip angle = 154°, voxel = 0.6 × 0.6 × 5 [section] mm, axial and coronal planes), and T1 postcontrast with fat saturation (TR/TE = 560/11 ms, flip angle = 180°, voxel = 0.6 × 0.6 × 5 [section] mm, axial and coronal planes). To improve the homogeneity of fat-suppression in the neck, we used the Dixon fat-suppression technique as reported in a prior study.[11] A total of 0.1 mmol/kg of gadolinium was administered for postcontrast imaging.

Grading Scale and Statistical Analysis
The proposed grading system—the Berenstein-De Leacy (BDL) system—is summarized in Table 1. It categorizes treatment responses into 7 distinct grades and includes the descriptive modifier "B," which can be added to any grade to signify granulation tissue formation in the treatment bed. Radiographic improvement is stratified across 4 grades (1–4). The remaining grades (5–7), respectively, signify no change, regression with extension into an untreated area, and gross progression. The grading system is discussed in detail below.

The BDL system was used by 3 neuroradiologists (R.D.L., A.D., K.N.) with expertise in head and neck imaging to grade interval changes in lesion volume in 56 patients with head and neck LM-LVMs. Fifty-one patients underwent fluoroscopically guided percutaneous sclerotherapy within the imaging interval, and 5 patients had no treatment. Grading was conducted in an independent and blinded fashion. Estimation of relative residual lesion volume was based on visual assessment without the aid of automated tools for volumetric analysis. All differences in scoring were resolved by discussion and consensus.

To assess the interrater reliability of the proposed scale, we conducted 3 statistical analyses. First, we calculated the Krippendorff alpha α statistic, a commonly used metric of interrater reliability, treating the grading criteria as an ordinal scale. Next, intraclass coefficients for consistency and absolute agreement were calculated to evaluate the variability of a single outcome grade with respect to the variation across all cases. Finally, the 2-way Spearman ρ correlation was calculated to assess linear correlations between all pairs of raters in a nonparametric fashion. A P value < .05 was set to demarcate statistical significance.

Table 2: Patient age and average treatments

	Median	Range
Age (yr)	6.38	(0.33–73.92)
Imaging interval (mo)	32.50	(1–131)
Treatments within imaging interval (mo)	1.50	(0–13)

Table 3: Patient demographics and LV-LVM characteristics

	No.	(%)
Total	56	
Sex		
Male	28	50.0%
Female	28	50.0%
Classification		
LM	37	66.1%
LVM	19	33.9%
Architecture		
Macrocystic	9	16.1%
Microcystic	19	33.9%
Mixed	28	50.0%
Lesion laterality		
Right	18	32.1%
Left	16	28.6%
Bilateral	22	39.3%
de Serres stage		
I (unilateral infrahyoid)	1	1.8%
II (unilateral suprahyoid)	27	48.2%
III (unilateral suprahyoid and infrahyoid)	8	14.3%
IV (bilateral suprahyoid)	9	16.1%
V (bilateral suprahyoid and infrahyoid)	11	19.6%

Table 4: Frequency of grading outcomes

Final Grade	Count	Frequency (%)
1	3	5.4%
2	8	14.3%
3	17	30.4%
4	8	14.3%
5	7	12.5%
6	6	10.7%
7	5	8.9%
B Modifier	2	3.6%

Table 5: BDL interrater reliability

Interrater Reliability Measure	Statistical Value	CI or P Value
Krippendorff α statistic	0.93	(95% CI, 0.89–0.95)
Intraclass coefficient[a]	0.97	(95% CI, 0.96–0.98)
Spearman ρ (K.N. vs A.D.)	0.90	$P < .001$
Spearman ρ (K.N. vs R.D.L.)	0.93	$P < .001$
Spearman ρ (A.D. vs R.D.L.)	0.95	$P < .001$

[a] Intraclass coefficients calculated for both consistency and absolute agreement yielded identical values.

For all of the above tests, we determined the strength of interrater reliability according to the criteria proposed by Cicchetti and Sparrow[12]: <0.40, poor; 0.40–0.59, fair; 0.60–0.74, good; ≥0.75, excellent.[13] A value of 1.00 indicates perfect agreement, 0 indicates no better than chance, and negative values indicate worse than chance. Statistical analysis was conducted using SPSS, Version 22.0 (IBM).
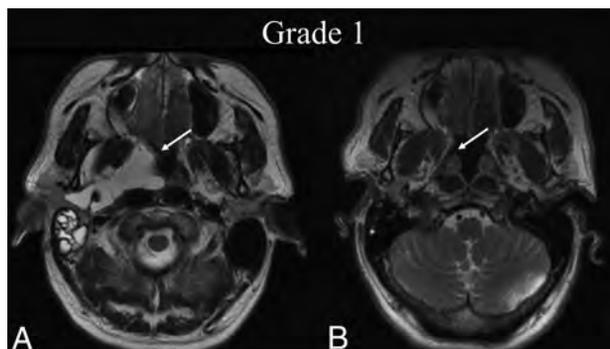
FIG 1. BDL grade 1 response. A 50-year-old man with dysphagia was found to have a de Serres stage II macrocystic LM of the head. *A*, Preprocedural T2-weighted axial MR imaging at the level of the hard palate shows a lesion primarily centered within the right parapharyngeal space (*white arrow*). *B*, Four months later, following 1 sclerotherapy treatment, T2-weighted axial MR imaging at the same level shows BDL grade 1 complete regression of the lesion on cross-sectional imaging (*white arrow*).
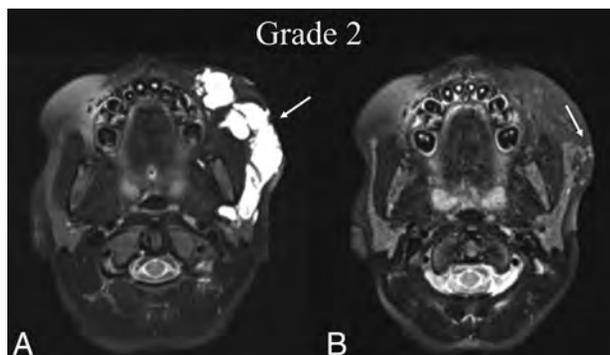
FIG 2. BDL grade 2 response. A 7-year-old girl with a left cheek mass was found to have a de Serres stage II mixed macro-/microcystic LM of the head. *A*, Preprocedural T2 fat-suppressed axial MR imaging at the level of the maxillary alveolar process shows involvement of the left buccal and parotid spaces (*white arrow*). *B*, Seven months later, following 2 sclerotherapy treatments, T2 fat-suppressed axial MR imaging at the same level shows BDL grade 2 near-complete resolution of the LM with trace residuals (*white arrow*).

RESULTS

Pre- and posttreatment MRIs of 56 patients were presented to 3 raters in a random order. The median age of patients on final imaging was 6.38 years (range, 0.3–73.9 years), and 28 patients (50.0%) were female. The median imaging interval was 32.5 months (range, 1–131), and the median number of sclerotherapy treatments received was 1.5 (range, 0–13); 66.1% of lesions were classified as pure LMs. LM-LVMs were localized to the right side of the head and neck in 18 cases (32.1%) and the left in 16 cases (28.6%) and were bilateral in 22 cases (39.3%). Nine LM-LVMs were classified as macrocystic (16.1%), 19 were microcystic (33.9%), and 28 were mixed (50.0%). Anatomic locations representing all preoperative stages of the de Serres criteria were present in the validation cohort. Patient demographics, imaging interval, number of treatments, and LVM characteristics are summarized in Tables 2 and 3.
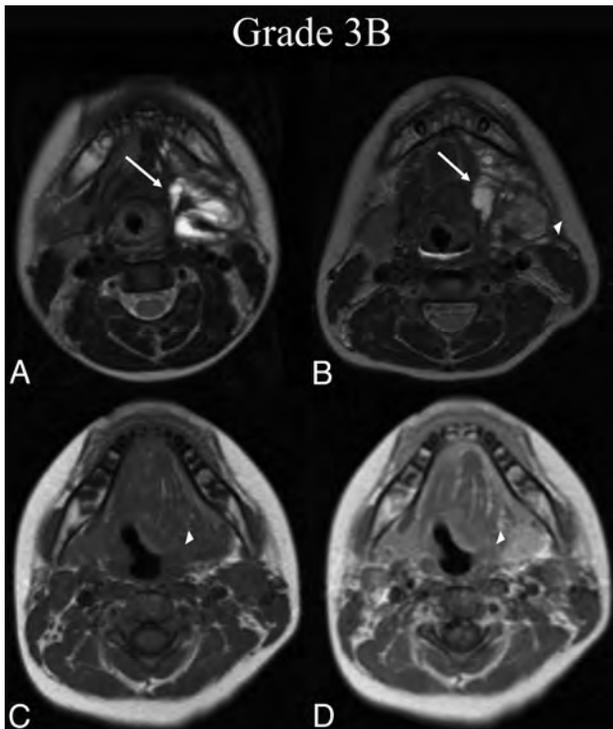
FIG 3. BDL grade 3B response. A 6-year-old boy with dysphagia was found to have a de Serres stage II mixed macro-/microsystic LM of the head. *A*, Preprocedural T2-weighted axial MR imaging at the level of the submandibular glands shows involvement of the left submandibular, sublingual, and parapharyngeal mucosal (not shown) spaces (*white arrow*). *B*, Thirty-seven months later, following 1 sclerotherapy treatment, T2-weighted axial MR imaging at the same level shows BDL grade 3B partial regression, with <50% estimated volume of residual malformation (*white arrow*) and intermediate signal granulation tissue formation within the treatment bed (*white arrowhead*). *C*, At the follow-up time point, note precontrast axial T1-weighted MR imaging isointensity of the suspected granuloma with muscle (*white arrowhead*). *D*, Postcontrast axial T1 at the same level demonstrates enhancement of the suspected granuloma (*white arrowhead*), confirming 3B grading.
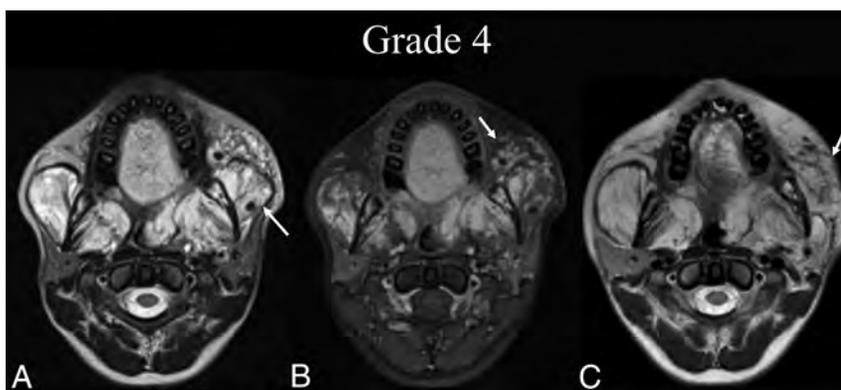

FIG 4. BDL grade 4 response. An 18-year-old woman with an extensive bilateral de Serres stage IV mixed macro-/microsystic LVM of the supra- and infrahyoid neck presents for sclerotherapy after receiving multiple treatments at an outside institution. *A*, Preprocedural T2-weighted axial MR imaging at the level of the alveolar process of the maxilla shows widespread LVM bilaterally (*white arrow*). *B*, Preprocedural T1 fat-saturated postcontrast axial MR imaging at the same level demonstrates enhancement of venous structures within the LVM (*white arrow*). *C*, Thirty-five months later, following 1 sclerotherapy treatment targeting the left buccal region, repeat T2-weighted axial MR imaging at the same level shows BDL grade 4 partial regression (*white arrow*) with >50% estimated volume of residual malformation.

There was unanimous agreement among raters in 39 cases (69.6%), agreement among two-thirds of raters in 16 cases (28.6%), and no agreement on initial grading in 1 case (1.7%). All discrepancies were resolved via discussion. All grades of the BDL scale, including the modifier for granulation tissue, were represented in the validation cohort. The prevalence of each grade is presented in Table 4. Stratification of grade prevalence by lesion architecture is available in the Online Supplemental Data.

The overall Krippendorff α statistic was 0.93 (95% CI, 0.89–0.95), denoting excellent agreement. Both intraclass coefficients with respect to consistency and absolute agreements were 0.97 (95% CI, 0.96–0.98), denoting excellent consistency. A 2-way Spearman ρ correlation was calculated between every permutation of raters: K.N. and A.D. demonstrated a ρ of 0.90 ($P < .001$), K.N. and R.D.L. demonstrated a ρ of 0.93 ($P < .001$), and A.D. and R.D.L demonstrated a ρ of 0.95 ($P < .001$). These values denote a strong and significant linear correlation between individual rater pairs. Measurements of interrater reliability are summarized in Table 5.

DISCUSSION

To the best of our knowledge, the BDL system described and evaluated above is the only grading system to be put forward for assessing the radiologic response to therapy in patients with head and neck LM-LVMs treated by either surgical excision or percutaneous sclerotherapy. Among 3 separate raters evaluating 56 different cases of LM-LVM treatment, it has demonstrated excellent consistency and rates of agreement, suggesting that it may be useful as a tool for both clinical communication and radiographic outcomes reporting.

The recent proliferation of studies evaluating techniques and agents for the treatment of LM-LVMs[14-19] necessitates a standardized method for radiographic outcome, reporting that can provide a common language for accurate comparison among independent trials. The utility of such a system is well-exemplified by the radiographic Response Evaluation Criteria in Solid Tumors for response to treatment in solid tumors, which, since its introduction in 2000, has allowed reliable comparison of different treatment trials during the past 2 decades.[20] Similar benefits extend across a broad range of disease interventions, including intracranial aneurysm embolization (Modified Raymond-Roy Classification), stroke intervention (Modified Treatment in Cerebral Ischemia score), and others.[21,22]

Important steps toward improved standardization have already been taken with regard to preoperative staging and clinical outcome reporting, and this scale is intended to supplement this effort by providing a standard, simple, and reproducible measure of radiographic treatment results.
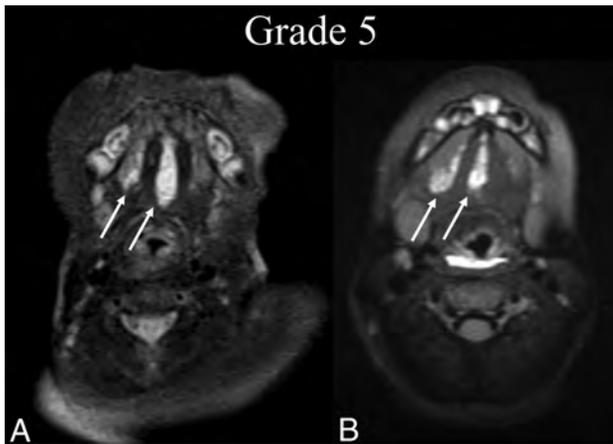
FIG 5. BDL grade 5 response. A 2-month-old boy with tongue swelling was found to have a de Serres stage II mixed macro-/microcystic sublingual LM. *A*, Preprocedural T2 fat-suppressed axial MR imaging at the level of the mandibular symphysis shows involvement of the right and middle aspects of the sublingual space (*white arrows*). *B*, Five years later, following 1 sclerotherapy treatment, T2 fat-suppressed axial MR imaging at the same level shows BDL grade 5 recurrence of the LM without a gross interval change from baseline (*white arrows*).


FIG 6. BDL grade 6 response. An 8-year-old boy with right facial prominence was found to have a de Serres stage II mixed macro-/microcystic LM. *A*, Preprocedural T2 fat-suppressed axial MR imaging at the level of the oropharynx shows trans-spatial involvement of the right neck extending from the right parotid and masticator spaces to the pharyngeal mucosal space (*white arrow*), sparing the retropharyngeal space. *B*, One year later, following 2 sclerotherapy treatments, T2 fat-suppressed axial MR imaging at the same level shows BDL grade 6 regression within the treatment bed and extension into the retropharyngeal space, which was previously uninvolved (*white arrow*).

The 7-grade scale proposed and evaluated here stratifies radiographic improvement across 4 levels, illustrated by case examples in Figs 1–4. These grades include the following: 1) complete regression of the lesion, 2) near-complete repression with trace residual lesion, 3) partial regression with <50% residual lesion, and 4) partial regression with >50% residual lesion. Grade 5 indicates minimal or no interval change in the treated LM-LVM (Fig 5). Regression of the LM-LVM in 1 area, with expansion into a previously uninvolved area is denoted by grade 6 (Fig 6). Finally, gross interval progression is denoted by grade 7 (Fig 7). The formation of granulation tissue within the treatment bed, which

FIG 7. BDL grade 7 progression. A neonate boy with right facial prominence was found to have a de Serres stage II macrocystic LM. *A*, Initial T2 fat-suppressed axial MR imaging at the level of the hard palate shows involvement of the right buccal space (*white arrow*). *B*, Three years later, before any intervention, T2 fat-suppressed axial MR imaging at the same level shows BDL grade 7 progression with clear volume increase and extension into the masticator and parapharyngeal (*white arrow*).

may occur in conjunction with any radiographic grade, is indicated by the modifier B as illustrated in Fig 3. This modifier was omitted from analysis to preserve the statistical power of interrater testing; thus, its interrater reliability has not been evaluated. Nonetheless, it is included as a possible addition to each grade due to the clinical importance of granuloma formation, which is pathologically distinct from the LVM itself and may impede the improvement of cosmetic deficits or mass-related symptoms.

The common standard offered by the BDL system is designed to include a number of advantages over current consensus recommendations that include "LM volume" as the only standardized radiographic indicator of treatment outcome.[10] By treating the LM-LVM response to therapy as a percentage change, it avoids direct estimation of LM-LVM volume, which, in the absence of advanced segmentation methods, is prone to significant error due to the irregular geometry commonly observed in these lesions. Furthermore, by considering the change across an imaging interval, the system captures treatment response as opposed to the absolute volume of the residual lesion after treatment. The BDL system was not designed to assess the clinical efficacy of embolization procedures but rather to provide a standardized system by which radiographic changes in head and neck LM-LVMs can be reported for the purposes of research and clinical communication. Nonetheless, it is structured in a way that allows future validation studies correlated to clinical outcomes.

The primary limitation of this study comes as a result of the validation cohort sample size, which is relatively small due to the rarity with which these malformations present for treatment. As a result, consistency and agreement were evaluated only with respect to the scale as a whole, and not within each grade of the scale individually. Similarly, the descriptive B, which in our proposed grading system represents the modifier used for the present of granulation tissue within the treatment bed, was also excluded from analysis. Further studies evaluating this scale in a larger

cohort could be conducted to overcome this limitation. Other promising areas for future study include validation of the scale with regard to open surgical treatment outcomes, correlation to clinical examination findings, and correlation with meaningful clinical end points, most notably recurrence rates.

CONCLUSIONS

The present study demonstrates that the BDL system for assessing the treatment of head and neck LM-LVMs has excellent consistency and rates of agreement. This is the only grading system to be put forward for assessing the response to therapy in patients with head and neck LM-LVMs treated by either surgical excision or percutaneous sclerotherapy. This system offers an effective means by which to streamline clinical communication and standardize radiographic outcome reporting for the treatment of these lesions.

REFERENCES

1. Wassef M, Blei F, Adams D, et al. ISSVA Board and Scientific Committee. **Vascular anomalies classification: recommendations from the International Society for the Study of Vascular Anomalies.** *Pediatrics* 2015;136:e203–14 CrossRef Medline

2. Defnet AM, Bagrodia N, Hernandez SL, et al. **Pediatric lymphatic malformations: evolving understanding and therapeutic options.** *Pediatr Surg Int* 2016;32:425–33 CrossRef Medline

3. Colbert SD, Seager L, Haider F, et al. **Lymphatic malformations of the head and neck-current concepts in management.** *Br J Oral Maxillofac Surg* 2013;51:98–102 CrossRef Medline

4. Horbach SE, Lokhorst MM, Saeed P, et al. **Sclerotherapy for low-flow vascular malformations of the head and neck: a systematic review of sclerosing agents.** *J Plast Reconstr Aesthet Surg* 2016;69:295–304 CrossRef Medline

5. Cheng J. **Doxycycline sclerotherapy in children with head and neck lymphatic malformations.** *J Pediatr Surg* 2015;50:2143–46 CrossRef Medline

6. Curran AJ, Malik N, McShane D, et al. **Surgical management of lymphangiomas in adults.** *J Laryngol Otol* 1996;110:586–89 CrossRef Medline

7. Perkins JA, Manning SC, Tempero RM, et al. **Lymphatic malformations: review of current treatment.** *Otolaryngol Head Neck Surg* 2010;142:795–803 CrossRef Medline

8. Hartl DM, Roger G, Denoyelle F, et al. **Extensive lymphangioma presenting with upper airway obstruction.** *Arch Otolaryngol Head Neck Surg* 2000;126:1378–82 CrossRef Medline

9. de Serres LM, Sie KC, Richardson MA. **Lymphatic malformations of the head and neck: a proposal for staging.** *Arch Otolaryngol Head Neck Surg* 1995;121:577–82 CrossRef Medline

10. Balakrishnan K, Bauman N, Chun RH, et al. **Standardized outcome and reporting measures in pediatric head and neck lymphatic malformations.** *Otolaryngol Head Neck Surg* 2015;152:948–53 CrossRef Medline

11. Barger AV, DeLone DR, Bernstein MA, et al. **Fat signal suppression in head and neck imaging using fast spin-echo-IDEAL technique.** *AJNR Am J Neuroradiol* 2006;27:1292–94 Medline

12. Cicchetti DV, Sparrow SA. **Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior.** *Am J Ment Defic* 1981;86:127–37 Medline

13. Mazur MD, Taussky P, Shah LM, et al. **Inter-rater reliability of published flow diversion occlusion scales.** *J Neurointerv Surg* 2016;8:1294–98 CrossRef Medline

14. Aluffi Valletti P, Brucoli M, Boffano P, et al. **A single-center experience in the management of head and neck lymphangiomas.** *Oral Maxillofac Surg* 2020;24:109–15 CrossRef Medline

15. Curry S, Logeman A, Jones D. **Sirolimus: a successful medical treatment for head and neck lymphatic malformations.** *Case Rep Otolaryngol* 2019;2019:2076798 CrossRef Medline

16. Malic CC, Guilfoyle R, Courtemanche RJ, et al. **Lymphatic malformation architecture: implications for treatment with OK-432.** *J Craniofac Surg* 2017;28:1721–24 CrossRef Medline

17. Wang S, Du J, Liu Y, et al. **Clinical analysis of surgical treatment for head and neck lymphatic malformations in children: a series of 128 cases.** *Acta Otolaryngol* 2019;139:713–19 CrossRef Medline

18. Wang Y, Tang W, Li X. **Safety and efficacy of surgery combined with bleomycin irrigation for complex cervical-facial lymphatic malformations of children.** *Int J Pediatr Otorhinolaryngol* 2020;128:109724 CrossRef Medline

19. Yamaki T, Sasaki Y, Hasegawa Y, et al. **Percutaneous ultrasound-guided sclerotherapy with polidocanol microfoam for lymphatic malformations.** *J Vasc Surg Venous Lymphat Disord* 2017;5:707–14 CrossRef Medline

20. Therasse P, Arbuck SG, Eisenhauer EA, et al. **New guidelines to evaluate the response to treatment in solid tumors.** *J Natl Cancer Inst* 2000;92:205–16 CrossRef Medline

21. Mascitelli JR, Moyle H, Oermann EK, et al. **An update to the Raymond-Roy Occlusion Classification of intracranial aneurysms treated with coil embolization.** *J Neurointerv Surg* 2015;7:496–502 CrossRef Medline

22. Zaidat OO, Yoo AJ, Khatri P, et al. STIR Thrombolysis in Cerebral Infarction (TICI) Task Force. **Recommendations on angiographic revascularization grading standards for acute ischemic stroke: a consensus statement.** *Stroke* 2013;44:2650–63 CrossRef Medline

MRI Evaluation of the Normal and Abnormal Endolymphatic Duct in the Pediatric Population: A Comparison with High-Resolution CT

R.L. Clarke, B. Isaacson, J.W. Kutz, Y. Xi, and T.N. Booth

ABSTRACT

BACKGROUND AND PURPOSE: An enlarged vestibular aqueduct is the most commonly reported imaging abnormality in children with sensorineural hearing loss. MR imaging is often used to evaluate pediatric sensorineural hearing loss; however, there are no well-established size criteria on MR imaging to diagnose an enlarged endolymphatic duct. The first purpose of the study was to determine a range of normal endolymphatic duct sizes on MR imaging and compare it with that in high-resolution CT. The second purpose was to assess the sensitivity and specificity of MR imaging in diagnosing an enlarged endolymphatic duct in patients with an enlarged vestibular aqueduct on CT.

MATERIALS AND METHODS: Endolymphatic duct midaperture measurements were analyzed in 52 patients with no history of sensorineural hearing loss. Comparison of CT and MR imaging was made in a second cohort of 41 patients with a normal midaperture width on CT. The sensitivity and specificity of MR imaging were then evaluated in a third cohort of 24 patients with a documented enlarged vestibular aqueduct on CT.

RESULTS: In 94 ears, normal endolymphatic duct midaperture measurements ranged from 0 to 0.9 mm on MR imaging. A significant correlation ($P < .001$) and moderate agreement were found between CT and MR imaging in 81 ears with a normal vestibular aqueduct on CT. Twenty-four patients had bilateral ($n = 14$) or unilateral ($n = 10$) enlarged vestibular aqueducts on CT, and the sensitivity and specificity of MR imaging were 97% and 100%, respectively, for a diagnosis of an enlarged endolymphatic duct.

CONCLUSIONS: MR imaging measurements of the normal endolymphatic duct are similar to those established for CT. MR imaging is a useful tool for the diagnosis of enlarged vestibular aqueduct.

ABBREVIATIONS: CHARGE = Coloboma of the eye, Heart defects, Atresia of the choanae, Retardation of growth and/or development, Genital and/or urinary abnormalities, and Ear abnormalities and deafness; ELD = endolymphatic duct; EVA = enlarged vestibular aqueduct; HRCT = high-resolution CT; SNHL = sensorineural hearing loss; VA = vestibular aqueduct

The most common abnormality reported in the literature in children with sensorineural hearing loss (SNHL) is an enlarged vestibular aqueduct (EVA) as demonstrated on high-resolution CT (HRCT). The abnormality is commonly bilateral, and patients typically present with progressive and, at times, sudden hearing loss. The sudden hearing loss may be related to head trauma, sometimes minor in severity.[1] The criteria of an EVA have been revised across time with a normal midaperture width of 1.5 mm on HRCT suggested as a cutoff by Valvassori and Clemis.[2] More recently, a cutoff of 1 mm was suggested on the basis of an audiologic and HRCT correlation (Cincinnati criteria), and further

reductions in the normal width in the healthy population have been reported with oblique reconstruction planes (Pöschl plane).[3,4]

There are differing opinions in the literature as to the optimal imaging technique to evaluate children with SNHL. Previous studies have reported decreased sensitivity in diagnosing EVA with MR imaging compared with CT.[5] By means of 3D fluid-sensitive sequences, the endolymphatic duct (ELD), which is housed inside the vestibular aqueduct (VA), can be visualized. The endolymphatic sac and associated anomalies of the cochlea and vestibule are demonstrated with increased frequency on MR imaging compared with CT.[6] Additionally, MR imaging can determine the status of the cochlear nerve, an important finding for treatment options and the potential for cochlear implantation success.[7,8]

The purpose of this study was to determine a range of normal midaperture ELD sizes in a large cohort of patients without SNHL, to compare measurements of the normal VA and ELD obtained on both HRCT and MR imaging, respectively, and to

evaluate the sensitivity and specificity of MR imaging compared with HRCT in a smaller cohort of patients with EVA, using HRCT as the criterion standard.

MATERIALS AND METHODS

This is a retrospective study approved by the institutional review board and Health Insurance Portability and Accountability Act–compliant.

Healthy patients were selected from a group of patients referred for MR imaging evaluation of cholesteatoma. Inclusion criteria were the availability of a diagnostic axial 3D fluid-sensitive sequence (T2 FSE or T2 sampling perfection with application-optimized contrasts by using different flip angle evolution [SPACE sequence; Siemens]) with coverage including the entirety of the membranous labyrinth as well as a temporal bone HRCT. MR imaging was considered diagnostic if the basal turn of the osseous spiral lamina could be discerned. Exclusion criteria were an unavailable audiogram, any degree of SNHL, or a syndrome associated with hearing loss. The presence of conductive hearing loss was not exclusionary and was assumed to be the result of middle ear inflammation and/or cholesteatoma.

A second cohort of patients that had both a diagnostic temporal bone MR imaging and HRCT was then evaluated. The inclusion criterion was the presence of a normal midaperture width of the VA on HRCT. Patients were included regardless of audiogram or SNHL being present, to compare only agreement between the 2 modalities. High-resolution CT and MR imaging measurements were performed in separate sessions approximately 2 weeks apart by a single reader.

In the third cohort, a search of the electronic medical record was performed for patients in whom temporal bone imaging had been performed with both diagnostic MR imaging and HRCT. The presence of any associated relevant syndromes was noted. The temporal bone HRCTs were evaluated for the presence of an EVA, defined as a midaperture width of ≥1 mm. The dataset included patients with bilateral and unilateral enlarged VA. Two subspecialty-certified neuroradiologists with >20 years of experience (T.N.B.) and 6 years of experience (R.L.C.) in interpreting temporal bone imaging studies independently performed midaperture measurements on both HRCT and MR imaging in separate sessions approximately 2 weeks apart. Additionally, the position and subjective size of the tympanic facial nerve were evaluated on HRCT and categorized as normal or abnormal.

All CTs were obtained using a standard temporal bone protocol on a 128-section multidetector CT scanner (Somatom Definition Flash; Siemens). Axial images were obtained parallel to the lateral semicircular canal with 0.6-mm collimation, 0.5- to 1.0-mm thickness, and a 0.3-mm increment, parallel, using a standard bone window algorithm (window/level, 3400–500/500).

All MR imaging examinations were performed on a 1.5T (Achieva, Philips Healthcare; Aera, Siemens) or 3T scanner (Magnetom Skyra; Siemens). VA measurements were made using an axial 3D fluid-sensitive sequence (T2 FSE or T2 SPACE) (1400TR/250TE/2NEX), with coverage including the entirety of the membranous labyrinth. A variable window width and level were used to optimize for evaluation of the VA.
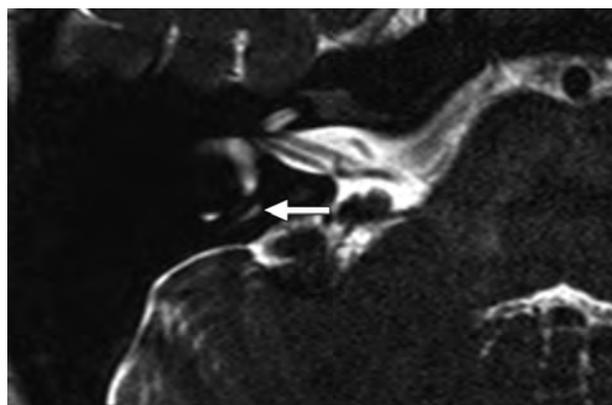
FIG 1. Axial MR imaging shows a normal ELD measuring 0.7 mm at the midaperture (*arrow*).

All studies were reviewed on a PACS workstation with a midaperture diameter of the ELD measured using an electronic caliper. Images were magnified 2–3 times depending on the original FOV. The VA/ELD midpoint was determined in the axial plane on both CT and MR imaging as the point half-way between its origin at the labyrinth near the common crus and its opercular edge. If the operculum could not be well-defined, the posterior wall of the petrous bone at the level of the posterior semicircular canal was used as a surrogate. The width was measured as a line perpendicular to the course of the VA/ELD at the midpoint. If the VA and ELD were not visualized, the width was recorded as 0 mm.

Statistical Analysis

In the healthy cohorts, Spearman and intraclass correlations were performed when comparing HRCT with MR imaging measurements, and mean (SD) was calculated for both MR imaging and HRCT. Sensitivity and specificity with confidence intervals for a diagnosis of an enlarged ELD were calculated using a cutoff of a 1-mm midaperture width. Intraclass correlation coefficients were calculated to assess interrater agreement for both HRCT and MR imaging measurements in the third cohort.

RESULTS

The healthy cohort included 52 patients with an age range of 2–17 years (mean, 10.3 years) and a male/female ratio of 2:1. Ninety-four ears had diagnostic MR imaging studies with an audiogram available for review and no documented SNHL. The ELD midaperture diameter ranged from 0 to 0.9 mm (mean, 0.39 [SD, 0.14] mm). Ninety-eight percent of midaperture widths were ≤0.8 mm (Fig 1).

The second group with both CT and MR imaging available included 41 patients and 81 ears with an age range of 0.3–15 years (mean, 10.8 years; male/female ratio = 1.3:1). The midaperture width of 81 ears with diagnostic imaging of the temporal bone on both HRCT and MR imaging was measured (a single patient had diagnostic imaging of only 1 ear). VA size on CT ranged from 0 to 0.8 mm (mean, 0.51 mm; 95% measuring ≤0.8 mm). In this same group, the ELD midaperture width measured by MR imaging ranged from 0 to 0.8 mm (mean, 0.36; 95% measuring ≤0.8 mm). A significant correlation was present between measurements on
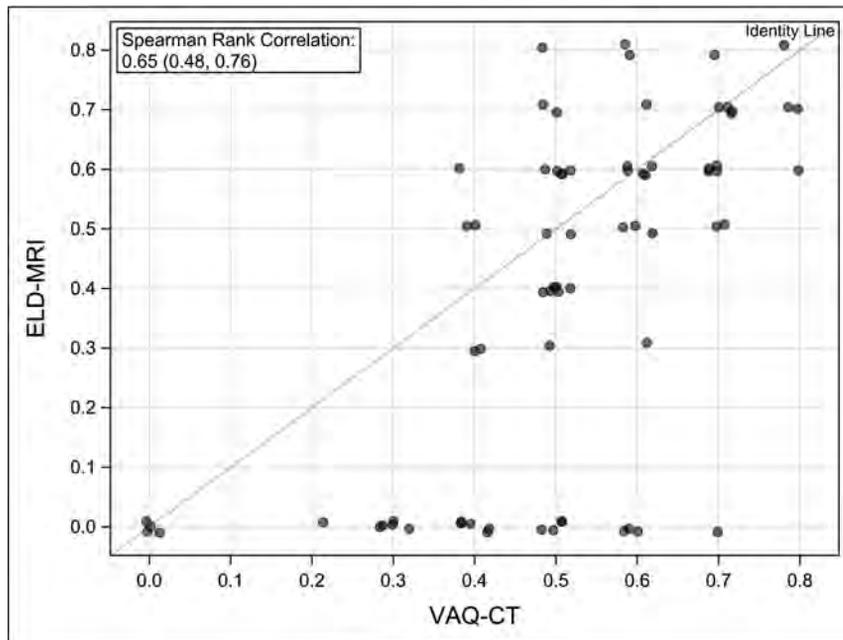
FIG 2. Scatterplot comparing midaperture measurements on CT and MR imaging.
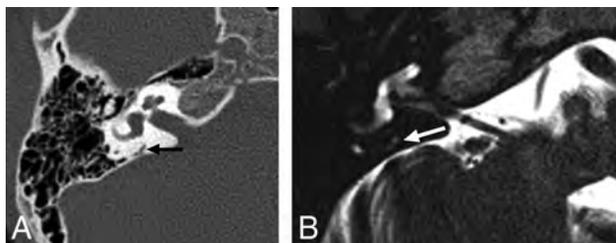
FIG 3. *A*, Axial HRCT shows a normal VA measuring 0.7 mm (*arrow*). *B*, MR imaging in the same patient shows the ELD measuring 0.5 mm (*arrow*).
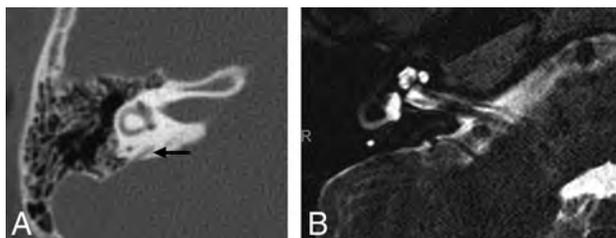
FIG 4. *A*, Axial HRCT shows a normal VA measuring 0.4 mm (*arrow*). *B*, MR imaging in the same patient shows nonvisualization of the ELD.

HRCT and MR imaging, with a Spearman correlation of 0.65 ($P < .001$) (Fig 2). Agreement between the modalities was moderate, with an intraclass coefficient of 0.618 (95% CI, 0.463–0.736). Overall, MR imaging tends to have smaller midaperture ELD widths compared with the VA on HRCT (Fig 3). Many small VAs on HRCT (normal-sized) were not visualized on MR imaging (Fig 4).

The third group had a diagnosis of EVA on HRCT with a diagnostic MR imaging available for review. This group included 24 patients and 48 ears, with an age range from 0.5 to 15 years (mean, 6.6 years; M/F ratio = 1:1.4). Three patients had a diagnosis of Coloboma of the eye, Heart defects, Atresia of the choanae, Retardation of growth and/or development, Genital and/or urinary abnormalities, and Ear abnormalities and deafness (CHARGE) syndrome. Bilateral EVA was present in 14 patients, and unilateral, in 10, with 38 ears having an EVA on CT (Fig 5). The mean VA size was 1.7 mm, and the median VA size was 1.6, with a range of 0–4 mm. By means of a cutoff of 1-mm midaperture width, only 1 ear was misdiagnosed as normal on MR imaging, with the ELD measuring 0.9 mm on MR imaging compared with the VA measuring 1.6 mm on HRCT (Fig 6). All normal VAs on CT were measured as normal ELDs on MR imaging. With HRCT as the standard, sensitivity for demonstrating an enlarged ELD using MR imaging was 97% (95% CI, 86%–100%) with a specificity of 100% (95% CI, 69%–100%) for the first reader. For the second reader, sensitivity was 89% (95% CI, 75%–97%) and specificity was 90% (95% CI, 56%–100%). Intraclass correlation coefficients calculated for interrater agreement were 0.89 (95% CI, 0.81–0.93) on MR imaging and 0.91 (95% CI, 0.84–0.95) on CT.

All patients with nonsyndromic hearing loss had normally positioned tympanic portions of the facial nerve without surgical implications. Three patients with CHARGE syndrome had abnormal facial nerves with bilateral oval window atresia. All 3 patients had hypoplastic facial nerves with 1 displaced into the oval window in both ears. All facial nerves were difficult to visualize with little confidence in the relation of the facial nerve to the round window.

DISCUSSION

The ability to accurately diagnose EVA in children is important, allowing for appropriate counseling of the patient to avoid contact sports, and detection of associated inner ear anomalies, genetic associations, and other imaging findings that may influence surgical planning for cochlear implantation. Advantages of HRCT include its availability and speed as well as lower cost compared with MR imaging. There is excellent osseous detail for visualization of the VA. It also allows assessment of other structures such as the facial nerve canal, which are not well-evaluated on MR imaging and may have implications for surgical planning. However, the risks associated with the use of ionizing radiation in the pediatric population cannot be ignored. Advantages of MR imaging include a lack of radiation exposure and the ability to directly visualize the endolymphatic duct and sac and, most important, the cochlear nerve using high-resolution volumetric T2-weighted sequences. On the other hand, MR imaging is more costly and time-consuming than CT, often requiring the use of
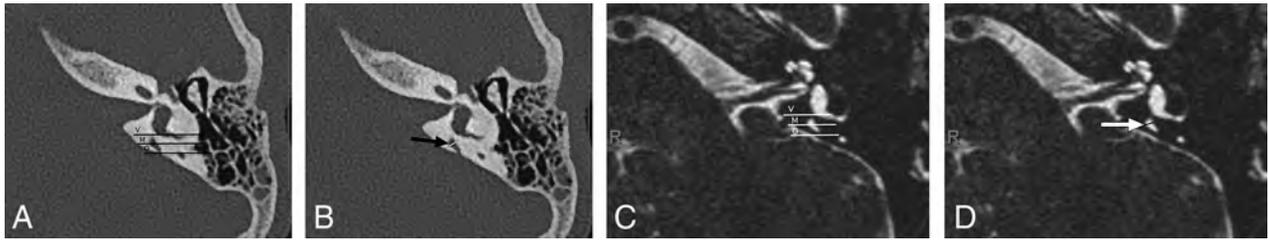
**FIG 5.** Axial HRCT (*A* and *B*) shows the vestibular (V) and opercular (O) planes used to define the VA midpoint (M), and an EVA measuring 1.3 mm (*arrow*). MR imaging (*C* and *D*) in the same patient shows the corresponding vestibular (V), opercular (O), and midpoint (M) planes and an enlarged ELD measuring 1.2 mm (*arrow*).
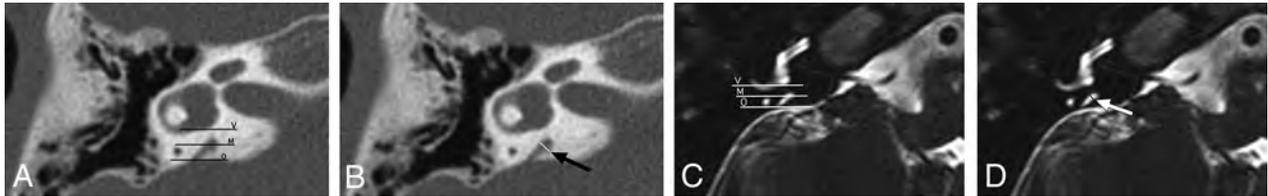

FIG 6. Axial HRCT (*A* and *B*) shows the vestibular (V) and opercular (O) planes used to define the VA midpoint (M), and an EVA measuring 1.6 mm (*arrow*). MR imaging (*C* and *D*) in the same patient shows the corresponding vestibular (V), opercular (O), and midpoint (M) planes and a normal-sized ELD measuring 0.9 mm (*arrow*).

anesthesia or sedation, which comes with additional risks. The decision to use MR imaging versus CT in a patient with suspected EVA must involve all of these factors.

There continues to be debate regarding the diagnostic utility of HRCT versus MR imaging in the evaluation of patients with SNHL and possible EVA. Much of the early research was in support of HRCT as the best imaging technique in this group of patients; however, several more recent studies have suggested an increased role for MR imaging, given the advances in image quality and technology in recent years. Deep et al[9] reported a concordance rate of 88% between HRCT and MR imaging in the diagnosis of EVA in 32 patients when both midpoint and opercular measurements were taken into account. However, more cases were detected by CT alone than by MR imaging alone, and in 12% of cases, the ELD was not visualized on MR imaging when a diagnosis of EVA was made by HRCT measurements. A study by Yang and Liu,[10] which compared midpoint aperture and opercular measurements on HRCT and MR imaging in 25 patients with EVA and an enlarged ELD, showed that measurements at the midpoint were wider on HRCT than on MR imaging, while measurements at the operculum were wider on MR imaging than on HRCT, suggesting a complementary role of the 2 modalities in diagnosing EVA or an enlarged ELD. Another study by Connor et al,[11] which compared CT and MR imaging measurements in 53 subjects, showed 93% agreement between the 2 modalities and no difference in diagnostic ability.

Despite recent studies showing the greater utility of MR imaging, normal ranges for ELD measurements are yet to be established. One study by Dahlen et al,[12] in 1997, evaluating 88 ears using MR imaging showed an average midpoint measurement of 0.8 mm and a range of 0.5–1.4 mm. The authors concluded that a midpoint measurement of ≥1.5 mm on MR imaging is abnormal. This finding corresponds well with the cutoff suggested by

Valvassori and Clemis,[2] but not the more recent cutoff proposed in the Cincinnati criteria. Given the advances in the quality of MR imaging in the decades since that study, further investigation is warranted.

In our cohort of patients with normal HRCT findings and no history of SNHL, the average midaperture ELD measurement on MR imaging was 0.39 mm, with a range of 0–0.9 mm. These values correspond well to the Cincinnati cutoff for HRCT measurements but remain lower than those described by Dahlen et al.[12] This difference may reflect the improved resolution of modern MR imaging resulting in mitigation of the "blooming effect," in which a bright endolymph signal was thought to contribute to an artifactual larger size of the ELD with older MR imaging technology.[12]

Our study also shows a high degree of correlation of midaperture measurements between CT and MR imaging in subjects with normal VAs on CT, with a Spearman correlation of 0.65. Moreover, in patients with known EVA on CT, MR imaging showed a sensitivity of up to 97% and a specificity of up to 100% for an enlarged ELD. Only 1 case was diagnosed by HRCT alone. The findings for the third cohort also demonstrate a high degree of reproducibility of VA/ELD measurements, with good-to-excellent agreement between the 2 readers.

Some have argued that CT and MR imaging are complementary in cases of suspected EVA or an enlarged ELD because of the additional information supplied by HRCT regarding the course of the facial nerve. In our cohort of 24 patients with documented EVA, 3 were found to have an abnormal facial nerve on HRCT. However, all of these had a diagnosis of CHARGE syndrome and would have undergone HRCT regardless. None of the subjects with nonsyndromic hearing loss had an abnormal facial nerve.

Limitations are primarily centered on the limited number of subjects with EVA who had both CT and MR imaging. An increased number of patients are needed to better determine the sensitivity and equivalency because the lower-margin confidence interval for sensitivity lies below 90%.

CONCLUSIONS

In patients without SNHL, ranges of endolymphatic duct measurements by MR imaging are similar to those of CT of the VA with an upper limit of 0.9 mm (98%). MR imaging tends to underestimate the size of a normal VA. MR imaging was able to diagnose an enlarged VA with a sensitivity of 97% and could potentially be used to make this diagnosis.

REFERENCES

1. McClay JE, Booth TN, Parry DA, et al. **Evaluation of pediatric sensorineural hearing loss with magnetic resonance imaging.** *Arch Otolaryngol Head Neck Surg* 2008;134:945–52 CrossRef Medline
2. Valvassori GE, Clemis JD. **The large vestibular aqueduct syndrome.** *Laryngoscope* 1978;88:723–28 CrossRef Medline
3. Boston M, Halsted M, Meinzen-Derr J, et al. **The large vestibular aqueduct: a new definition based on audiologic and computed tomography correlation.** *Otolaryngol Head Neck Surg* 2007;136:972–77 CrossRef Medline
4. Juliano AF, Ting EY, Mingkwansook V, et al. **Vestibular aqueduct measurements in the 45° oblique (Pöschl) plane.** *AJNR Am J Neuroradiol* 2016;37:1331–37 CrossRef Medline
5. Kachniarz B, Chen JX, Gilani S, et al. **Diagnostic yield of MRI for pediatric hearing loss: a systematic review.** *Otolaryngol Head Neck Surg* 2015;152:5–22 CrossRef Medline
6. Davidson HC, Harnsberger HR, Mancuso AA, et al. **MR evaluation of vestibulocochlear anomalies associated with large endolymphatic duct and sac.** *AJNR Am J Neuroradiol* 1999;20:1435–41 Medline
7. Glastonbury CM, Davidson HC, Harnsberger HR, et al. **Imaging findings of cochlear nerve deficiency.** *AJNR Am J Neuroradiol* 2002;23:635–43 Medline
8. Kim BG, Chung HJ, Park JJ, et al. **Correlation of cochlear nerve size and auditory performance after cochlear implantation in postlingually deaf patients.** *JAMA Otolaryngol Head Neck Surg* 2013;139:604–09 CrossRef Medline
9. Deep NL, Carlson ML, Weindling SM, et al. **Diagnosing large vestibular aqueduct: radiological review of high-resolution CT versus high-resolution volumetric MRI.** *Otol Neurotol* 2017;38:948–55 CrossRef Medline
10. Yang L, Liu J. **Comparative analysis of CT and MRI diagnosis of large vestibular aqueduct syndrome (LVAS) in children.** *J Coll Physicians Surg Pak* 2019;29:753–56 CrossRef Medline
11. Connor SEJ, Dudau C, Pai I, et al. **Is CT or MRI the optimal imaging investigation for the diagnosis of large vestibular aqueduct syndrome or large endolymphatic sac anomaly?** *Eur Arch Otorhinolaryngol* 2019;276:693–702 CrossRef Medline
12. Dahlen RT, Harnsberger HR, Gray SD, et al. **Overlapping thin-section fast spin-echo MR of the large vestibular aqueduct syndrome.** *AJNR Am J Neuroradiol* 1997;18:67–75 Medline

Automating Quantitative Measures of an Established Conventional MRI Scoring System for Preterm-Born Infants Scanned between 29 and 47 Weeks' Postmenstrual Age

L. van Eijk, M. Seidel, K. Pannek, J.M. George, S. Fiori, A. Guzzetta, A. Coulthard, J. Bursle, R.S. Ware, D. Bradford, S. Rose, P.B. Colditz, R.N. Boyd, and J. Fripp

ABSTRACT

BACKGROUND AND PURPOSE: Conventional MR imaging scoring is a valuable tool for risk stratification and prognostication of outcomes, but manual scoring is time-consuming, operator-dependent, and requires high-level expertise. This study aimed to automate the regional measurements of an established brain MR imaging scoring system for preterm neonates scanned between 29 and 47 weeks' postmenstrual age.

MATERIALS AND METHODS: This study used T2WI from the longitudinal Prediction of PREterm Motor Outcomes cohort study and the developing Human Connectome Project. Measures of biparietal width, interhemispheric distance, callosal thickness, transcerebellar diameter, lateral ventricular diameter, and deep gray matter area were extracted manually (Prediction of PREterm Motor Outcomes study only) and automatically. Scans with poor quality, failure of automated analysis, or severe pathology were excluded. Agreement, reliability, and associations between manual and automated measures were assessed and compared against statistics for manual measures. Associations between measures with postmenstrual age, gestational age at birth, and birth weight were examined (Pearson correlation) in both cohorts.

RESULTS: A total of 652 MRIs (86%) were suitable for analysis. Automated measures showed good-to-excellent agreement and good reliability with manual measures, except for interhemispheric distance at early MR imaging (scanned between 29 and 35 weeks, postmenstrual age; in line with poor manual reliability) and callosal thickness measures. All measures were positively associated with postmenstrual age ($r = 0.11$–0.94; $R^2 = 0.01$–0.89). Negative and positive associations were found with gestational age at birth ($r = -0.26$–0.71; $R^2 = 0.05$–0.52) and birth weight ($r = -0.25$–0.75; $R^2 = 0.06$–0.56). Automated measures were successfully extracted for 80%–99% of suitable scans.

CONCLUSIONS: Measures of brain injury and impaired brain growth can be automatically extracted from neonatal MR imaging, which could assist with clinical reporting.

ABBREVIATIONS: DGMA = deep gray matter area; dHCP = developing Human Connectome Project; GA = gestational age at birth; ICC = intraclass correlation coefficient; LoA = 95% limits of agreement; LVD = lateral ventricular diameter; PMA = postmenstrual age; PPREMO = Prediction of PREterm Motor Outcomes study; SEM = standard error of measurement; TEA = term-equivalent age

Conventional brain MR imaging scoring of preterm infants classifies injury and impaired growth and contributes to risk stratification and neurodevelopment outcome prognosis.[1] Preterm infants are at risk of adverse motor,[2,3] cognitive,[2,4] and behavioral outcomes.[2,5,6] Although survival rates for premature birth have improved[7,8] and cerebral palsy rates are declining,[9,10] long-term developmental impairments remain concerning.[7,8,11] Neuroimaging is becoming more common in preterm infants before discharge from the neonatal intensive care unit and offers an opportunity for early prognosis. Near term age, the brain undergoes rapid growth.[12,13] Abnormalities identified during this period have potential as predictors of neurodevelopment in children born preterm[14] and can help identify infants at risk of subsequent motor

and cognitive delay. Early detection of these infants allows the efficacy of potential early treatments to be assessed.

A recent standardized conventional scoring system for MR imaging of preterm infants evaluates the severity of cerebral WM, cortical GM, deep GM, and cerebellar and global abnormalities.[15,16] These composite scores of abnormalities and their submeasures, derived from early (29–35 weeks' postmenstrual age [PMA]) or term-equivalent age (TEA) MR imaging, are associated with motor and cognitive outcomes.[1,17-20] George et al[1] ($n = 83$ early; $n = 77$ TEA) found that global brain, WM, and deep GM abnormalities at early and TEA MR imaging and cerebellar abnormalities at TEA MR imaging were associated with neurologic, sensory, motor, and cognitive outcomes at 1-year follow-up. Several studies have also shown associations between abnormalities on TEA MR imaging and cognitive and/or motor outcomes at 2-year follow-up, for WM,[17-19] GM,[17] and global abnormalities[20] and submeasures (ie, ventricular dilation[20] and smaller cerebellar diameter[18,20]). Others found no association with outcomes.[21] Furthermore, abnormalities at TEA MR imaging have been related to longer-term developmental outcomes.[22-25] WM abnormalities were associated with poorer motor outcomes at 5-year follow-up,[22] delayed language development,[23] poorer executive functioning,[23] and lower general intelligence at 4- and 6-year follow-up,[23] as well as poorer cognitive outcomes at 4-,[23] 6-,[23] 7-,[24] and 9-year follow-up. Anderson et al[24] also reported that global brain and deep GM abnormalities at TEA MR imaging were associated with poorer cognitive outcomes at 7-year follow-up,[24] including general intelligence, spelling, math, and motor function. Possible reasons for these discrepancies between studies include differences in the sampled population and MR imaging acquisition, rater variability, sample size, and heterogeneity, including brain injury severity.

Manual scoring is time-consuming and requires expertise in neonatal neuroimaging and is, therefore, impractical for larger data sets and clinical practice. This study aimed to automate the 6 regional measurements of a conventional MR imaging scoring system that are based on distance and area (accounting for approximately half the measures in the scoring system) to improve reproducibility and clinical utility. We automated the extraction of biparietal width, interhemispheric distance, thickness of the corpus callosum (at the genu, midbody, and splenium), transcerebellar diameter, left and right lateral ventricular diameter (LVD), and deep gray matter area (DGMA) for MRIs acquired between 29 and 47 weeks' PMA. Agreement, reliability, and associations between automated and manual measures were examined, as well as for manual test-retest data for comparison. We hypothesized at least a moderate association between automated and manual measures. Finally, we examined the association between each measure with PMA and gestational age at birth (GA), hypothesizing that automated measures would show associations similar to those of the manual measures.

Please address correspondence to K. Pannek, MD, The Australian e-Health Research Centre, CSIRO, Level 5 UQ Health Sciences Building, Royal Brisbane and Women's Hospital, Herston, Queensland 4029 Australia; e-mail: Kerstin.Pannek@csiro.au

MATERIALS AND METHODS

Study Design and Participants
Data were included from the Prediction of PREterm Motor Outcomes (PPREMO) study[26] (longitudinal study of very preterm infants) and the developing Human Connectome Project (dHCP)[27] (cross-sectional study including preterm and term-born infants). PPREMO recruited preterm infants born <32 weeks' GA with no congenital or chromosomal abnormalities and a reference sample of term-born infants (38–42 weeks' GA, birthweight >10th percentile) with an uncomplicated pregnancy, delivery, and postpartum period. Study design and protocol are available.[26] Ethical approval was obtained from the Royal Brisbane and Women's Hospital Human Research Ethics Committee (HREC/12/QRBW/245), The University of Queensland (2012001060), and the trial was registered with the Australian New Zealand Clinical Trials Registry (ACTRN12613000280707). The dHCP study (second release) includes the largest multimodal neonatal open-access data set (http://www.developingconnectome.org/project/). Details of the structural preprocessed data are available.[13] Ethics approval was obtained from the National Research Ethics Committee (REC:14/LO/1169). For both data sets, informed written parental consent was obtained for each infant, and local ethics approval for using the data for the current study was obtained from the CSIRO Health and Medical Human Research Ethics Committee (2020_051_LR).

MR Image Acquisition
PPREMO infants were scanned during natural sleep (no sedation or anesthesia) at the Royal Brisbane and Women's Hospital on a 3T MR imaging scanner (Tim Trio; Siemens) in an MR imaging–compatible incubator with a dedicated neonatal head coil (LMT Lammers Medical Technology). MR imaging noise was attenuated using MiniMuffs (Natus Medical). Infants were monitored with pulse oximetry. T2WI TSE volumes were acquired in the axial plane with the following parameters:[26] TR/TE = 10,580/189 ms, flip angle = 150°, FOV = 144 × 180 mm, matrix = 204 × 256, voxel size = $0.7 \times 0.7 \times 2$ mm³. Infants were scanned at 29–35 weeks' PMA (early MR imaging) and again at TEA.

Images for dHCP were acquired on a 3T MR imaging scanner (Achieva; Philips Healthcare) with a dedicated neonatal head coil.[27] We used the provided motion-corrected and super-resolution reconstructed images,[28] with a resolution of $0.5 \times 0.5 \times 0.5$ mm³.[13]

Conventional MR Imaging Scoring (Manual)
As described in George et al,[16] PPREMO images were manually rated using the standardized MR imaging scoring system.[15] A neurologist with training in radiology and experience with neonatal MR imaging scoring (S.F.) scored images masked to clinical history (except PMA at scanning). Scoring methodology was confirmed by a senior neuroradiologist (A.C.). We manually obtained 6 raw measures of distance and area: biparietal width, interhemispheric distance, thickness of the corpus callosum at 3 locations, transcerebellar diameter, left and right LVD, and DGMA (Table 1).[15] In addition, test-retest ratings were available for 20 infants (both time points): Scans were manually rated again by the same rater (S.F.) 1 month apart and by an independent blinded rater (J.B.), a pediatric radiologist. No manual ratings were available for dHCP.

Table 1: Six manual measures of distance and area

Measure	How to Measure
Biparietal width	Greatest distance between left and right parietal cortices, measured on a single coronal slice identifying bilateral cochlea and basilar truncus
Interhemispheric distance	Distance between crowns of superior frontal gyri, measured on the same coronal slice as biparietal width
Callosal thickness	Thinning at the genu, midbody, splenium
Transcerebellar diameter	Single coronal slice at level of ventricular atrium
Lateral ventricular diameter	Same coronal slice as transcerebellar diameter
Deep grey matter area	Single axial slice showing caudate heads, lentiform nuclei, and thalami

Table 2: Automated measures of distance and area

Measure	How to Measure
Biparietal width	Identify most lateral sagittal slices of parietal GM (*regionprops3*)
	For each GM voxel, calculate distance to voxels in opposite hemisphere
	Calculate maximum distance
Interhemispheric distance	Calculate distance from each voxel of superior frontal gyrus label in left hemisphere to each voxel label in right hemisphere (*bwmorph3, bwdist*)
	Derive minimum distance
Callosal thickness	Combine 2D corpus callosum segmentation of the 11 most medial slices
	Increase resolution (*imresize*) and improve mask (*bwmorph3, bwmorph*)
	Derive skeleton (*bwskel*); for every voxel, calculate normal vector, identify intersection with borders of segmentation (*points2contour, polyfit*)
	Derive distance between upper/lower segmentation borders for every voxel
	Apply smoothing (*nanfastsmooth*), make 3 divisions
	Obtain 97th percentile for each division (*prctile*)
Transcerebellar diameter	Model cerebellum segmentation as 3D ellipsoid (*bwmorph3, regionprops3*)
	Calculate length of principal axis (*regionprops3*)
Lateral ventricular dilation	Identify coronal slice at level of ventricular atrium (*regionprops3*): find maximum surface area using ventricle label (*bwmorph, regionprops*)
	Model 2D ventricle as ellipse, calculate length of minor axes (*regionprops*)
Deep grey matter area	Identify axial slice: centroid of caudate, thalamus, and lentiform nucleus (*bwmorph, regionprops*)
	Combine labels to calculate area (*regionprops*)

Note:—Matlab package used is provided in brackets.

Image Processing

For PPREMO, T2WIs were preprocessed with slice interpolation (up-sampling to a $0.7 \times 0.7 \times 1$ mm^3 resolution) and N4 bias field correction. T2WIs were segmented using the MILXView neuroimaging platform with the UNC neonate atlas[29] and the 20 ALBERT atlas[29,30] used to provide priors and 87 anatomic regions (Milx[31]). PPREMO scans were also segmented with the state-of-the-art dHCP parcellation pipeline. We did not use the dHCP parcellations for extraction of the measures for the PPREMO cohort, however, because several scans with a successful Milx segmentation had no successful dHCP segmentation available due to registration errors in the dHCP pipeline. For dHCP, preprocessed T2WIs and segmentation labels were downloaded.[13] None of the segmentations were manually edited.

T2WIs were rigidly aligned to either an early or TEA PPREMO population template, depending on the age at scanning, using the FMRIB Linear Image Registration Tool (FLIRT; http://www.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT).[32] Segmentations were transformed to this space using the obtained transformation, ensuring that all scans had the same orientation for subsequent analysis steps, while preserving brain size and shape. Because the corpus callosum is a relatively thin region, segmentation errors may influence the results more significantly than other measures using larger segmentations. To evaluate this problem, we extracted callosal thickness measures twice for the PPREMO cohort using segmentations from both pipelines, allowing comparison of quantification between pipelines (Milx versus dHCP).

Extraction of Automated Measures

The automated pipeline (Table 2 and Fig 1) to extract raw measures uses the segmentation labels and Matlab software (MathWorks). A subsample of the PPREMO cohort ($n = 85$) was used to develop the Matlab pipeline. For callosal thickness, we increased the resolution (PPREMO: factor 5; dHCP: factor 4, considering T2WI resolution). The 97th percentile was chosen to identify the distance of the thickest location separately for each division of the corpus callosum, reflecting a location similar to that of the manual measures (in particular the genu and splenium), while ensuring that measurements were not driven by outliers. For each scan, automatically extracted measures were visually inspected and rated as "poor," "usable" (some errors), or "good" (little-to-no error).

Statistical Analysis

Agreement between manual and automated measures is reported using standard error of measurement (SEM), and bias (ie, the mean difference of the measurements), with 95% limits of agreement (LoA). Bland-Altman plots were inspected for bias (separately for early and TEA MR imaging and for the combined sample). Reliability and associations between automated and manual measures were examined by calculating intraclass correlation coefficients (ICCs, type [3, 1]) and Pearson correlation coefficients (r), respectively. For comparison, reproducibility measures for

manual test-retest data were calculated for PPREMO only, because manual measurements were not available for dHCP.

Next, automated measures were also validated by examining their association with PMA, GA, and birth weight (Pearson *r*) in both cohorts using R Statistics (Version 4.0.2; http://www.r-project.org). When one examines these associations, replicating results in an independent cohort is important. The smaller sample size of the PPREMO cohort may result in less power to detect an effect than in the dHCP cohort; conversely, a smaller sample size may also result in overestimating the effect size of associations or false-positives.

For all analyses, only automated measurements with a visual quality rating of "usable" or "good" were included. For reproducibility (ie, agreement and reliability) analyses, preterm-born infants scanned at TEA in the PPREMO cohort were merged with term-born infants. When we examined the association of measures with PMA, GA, or birth weight, outliers in automated and manual measures were removed using the Tukey lowerInner and upperInner fence, which was performed separately for each measure, as well as

FIG 1. Automated MR imaging measures of biparietal width (BPW), interhemispheric distance (IHD), thickness of the corpus callosum (CC), transcerebellar diameter (TCD), left and right LVD, and DGMA. IHD, Blue to orange represents the distance (closest and farthest, respectively) for each voxel to the segmentation in the opposite hemisphere. CC, Upper image, T2WI with the CC segmentation border (red); lower panel, segmentation border (blue), skeleton (red), and distance between the upper and lower borders of the segmentation for each voxel on the skeleton (green). R indicates right; L, left; A, anterior; P, posterior.

separately for scan age <37 weeks and above, and separately within the preterm and term-born group (at TEA MR imaging).

RESULTS

The final sample for PPREMO included 94 infants who underwent MR imaging between 29 and 36 weeks' PMA (78%; Online Supplemental Data), of which 81 infants were scanned again at 38–47 weeks' PMA (77%), and 22 term-born infants (81%). For dHCP, scans of 455 subjects were included (90%; Online Supplemental Data). Demographic details are presented in Table 3.

The 6 automated measures were successfully extracted for 80%–99% of the final scans with a usable segmentation, equivalent to 56%–85% of total scans (Online Supplemental Data) because 14%–30% of the total scans were excluded due to poor quality, realignment errors, poor segmentation, and severe pathology (Online Supplemental Data). Our pipeline was successful for >93% of the suitable PPREMO scans, except for interhemispheric distance (74%) and LVD (left: 78%; right, 61%), and >92% of the suitable dHCP scans, except for callosal thickness measures (75%–88%). Automated biparietal width, transcerebellar diameter, LVD, and DGMA measures showed good agreement, good reliability, and a strong association with manual measures (Online Supplemental Data; SEM = 0.40–2.61, bias = 0.14–5.70, LoA = −4.81–12.92, ICC = 0.77–0.97, *r* = 0.77–0.98, *P* < .001). Findings (ie, agreement, reliability, and associations between automated and manual measures) were similar for both time points (Online Supplemental Data) and in line with the good-to-excellent interrater agreement and reliability for the respective manual measures (except for poorer agreement and reliability for the manual DGMA: SEM = 0.56–0.59, bias = −0.09, LoA = −1.73–1.55, ICC = 0.36–0.67, *r* = 0.36–0.69; Online Supplemental Data). The automated interhemispheric distance showed a moderate agreement, reliability, and association with the manual measure (SEM = 0.73, bias = −0.58, LoA = −2.59–1.44, ICC = 0.59, *r* = 0.59, *P* < .001), but poor reliability and a weak association was found at early MR imaging (SEM = 0.66, bias = −0.44, LoA = −2.27–1.39, ICC = 0.26, *r* = 0.26, *P* < .05), in line with poor interrater reliability and the weak association for manual test-retest data at early MR imaging (SEM = 0.56, bias = −0.04, LoA = −1.57–1.50, ICC = 0.23, *r* = 0.23; Online Supplemental Data).

Automated callosal thickness measures showed poor-to-no agreement, poor reliability, and a weak-to-no association with the manual measures (SEM = 0.77–1.87, bias = 2.76–5.85, LoA = −0.05–10.32, ICC = 0.01–0.24, *r* = 0.03–0.35; Online Supplemental

Table 3: Demographic and clinical characteristics

	PPREMO (*n* = 197)			
	Preterm, Early MR Imaging	Preterm, TEA MR Imaging	Control, TEA MR Imaging	dHCP
No.	94 of 121	81 of 105	22 of 27	455 of 506 (95 preterm)
Sex (female)	61%	63%	50%	45%
PMA at MR imaging (wk)[a]	31.86 ± 1.96 (29–35)	40.71 ± 1.43 (38–47)	41.29 ± 1.25 (39–44)	40.71 ± 2.86 (29–45)
GA (wk)[a]	28.40 ± 2.00 (23–31)	28.50 ± 2.20 (24–31)	40.00 ± 0.98 (38–41)	39.71 ± 2.82 (24–42)
Birth weight (g)[a]	1107 ± 368 (494–1886)	1061 ± 391 (494–1886)	3516 ± 157 (2932–4200)	3250 ± 928 (540–4800)

[a] *P* < .05 for 2-sided *t* test comparing preterm infants at 40 weeks with term-born controls (PPREMO). Continuous variables are given as median ± interquartile range (range). Scans with poor quality (PPREMO, *n* = 39), realignment errors (PPREMO, *n* = 1; dHCP, *n* = 39), and severe pathology (PPREMO, *n* = 16; dHCP, *n* = 12) were excluded.
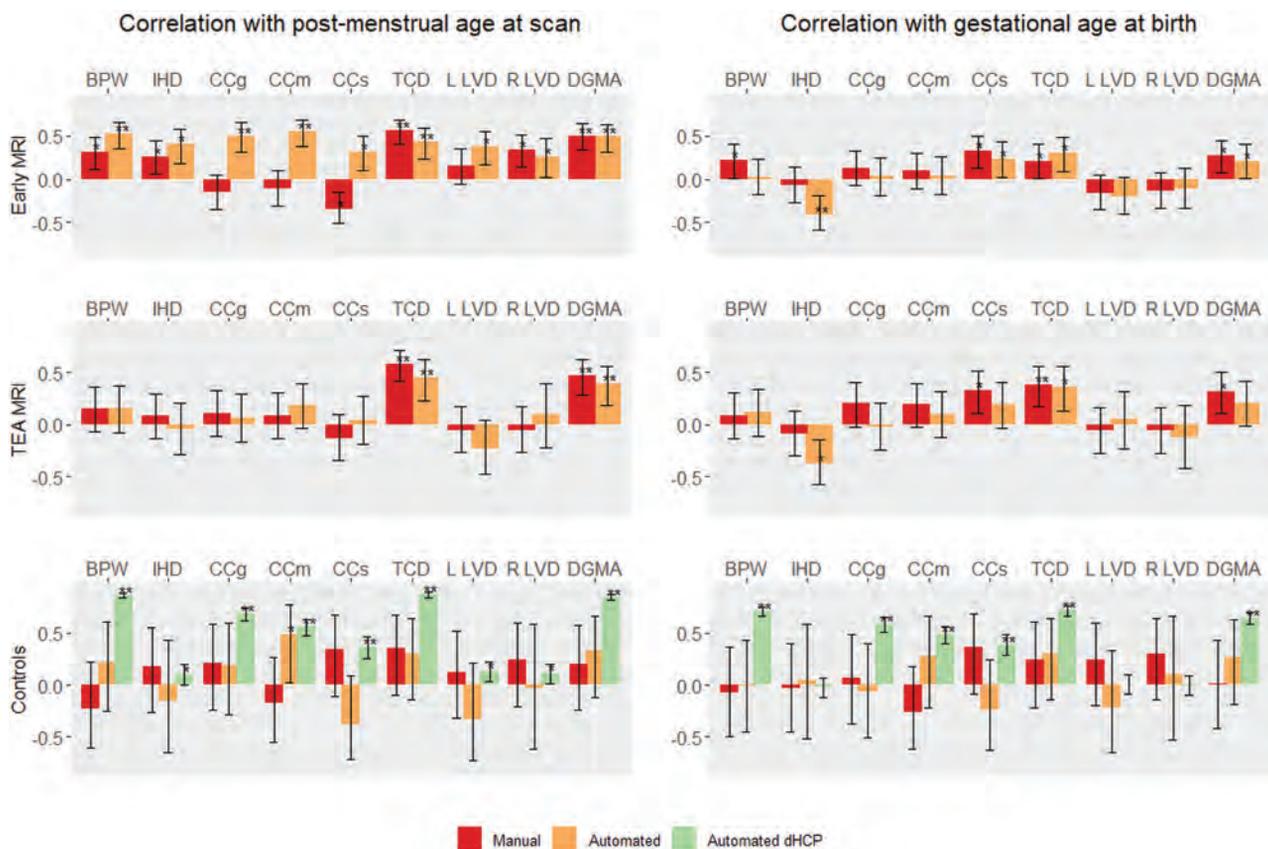
FIG 2. Associations between the 6 raw conventional MR imaging measures with postmenstrual age *(left)* and gestational age at birth *(right)* in the Pearson *r*, displayed separately for preterm-born infants at early MR imaging (*n* = 94) and TEA MR imaging (*n* = 81), as well as term-born controls at TEA MR imaging (PPREMO, *n* = 22; dHCP, *n* = 455). Manual measures were available for PPREMO only. *Double asterisks* indicate *P* ≤ .001; *asterisk*, *P* ≤ .05. BPW indicates biparietal width; CC, callosal thickness (at the genu (CCg), midbody (CCm), and splenium (CCs)); IHD, interhemispheric distance; TCD, transcerebellar diameter; R, right; L, left.

Data), whereas a moderate agreement, reliability, and association were found for manual test-retest data (SEM = 0.23–0.46, bias = −0.34–0.12, LoA = −1.61–1.29, ICC = 0.41–0.70, *r* = 0.41–0.70; Online Supplemental Data). Different measurement techniques likely contributed to the lack of agreement between automated and manual callosal thickness measures. For the automated measures, thickness is derived from multiple measurements and calculated as the distance between the upper and lower segmentation border (measuring perpendicular to the corpus callosum skeleton), whereas the distance for the manual measures is based on a single measurement and measured exactly anterior-posterior (genu and splenium) or inferior-superior (midbody). Care needs to be taken when using manual and automated measures interchangeably because a bias was observed for the callosal thickness measures in the Bland-Altman plots (Online Supplemental Data), with larger automated than manual values for smaller thickness values and smaller automated than manual values for larger thickness values. Examining the distribution of manual and automated measures, however, showed more clustering of time points for automated versus manual measures (Online Supplemental Data). The parcellation software used to extract the corpus callosum segmentation (Milx versus dHCP) influenced the derived measures for the PPREMO cohort slightly, shown by overlapping 95% confidence intervals for reproducibility statistics observed between manual and automated callosal

thickness measures (Online Supplemental Data). The reliability of the callosal thickness measures derived from the 2 segmentations was fair to good (genu ICC = 0.74 [0.66–0.81]; midbody ICC = 0.73 [0.65–0.80]; splenium ICC = 0.54 [0.42–0.65]; *P* < .001).

All automated measures (except right LVD in dHCP) showed a positive association with PMA for both the PPREMO (combining both time points; Online Supplemental Data) and dHCP cohorts (*r* = 0.11–0.94; Online Supplemental Data and Fig 2 left). These associations were similar to those found between the manual measures and PMA for all measures except callosal thickness (Online Supplemental Data). Automated callosal thickness measures showed stronger (and positive) associations with PMA than the manual measures (Online Supplemental Data), whereas some manual callosal thickness measures showed a counterintuitive negative association with PMA (at early MR imaging only) (Fig 2, left).

The associations with PMA were not always consistent across early and TEA MR imaging for the PPREMO cohort (Fig 2, left). All automated measures were positively associated with PMA at early MR imaging, while only transcerebellar diameter and DGMA showed a positive association at TEA MR imaging (a similar trend was observed for the manual measures) (Fig 2, left). In comparison, only callosal thickness at the midbody showed a positive association with PMA in the PPREMO term-born sample, but these results should be seen in light of the

much smaller sample size. In the dHCP, which includes both preterm and term-born infants, all automated measures were positively associated with PMA (Fig 2, left); these associations remained after removing all early MR imaging scans (scans with PMA <37 weeks) and also, when examining the associations between measures with PMA within the term-born group only.

Several automated measures (biparietal width, callosal thickness at the genu and splenium, transcerebellar diameter, and DGMA) also showed a positive association with GA across cohorts ($r = 0.34–0.71$; Online Supplemental Data and Fig 2, right), while the automated interhemispheric distance showed a negative association with GA for the PPREMO cohort only ($r = -0.26$; Online Supplemental Data). In addition, the automated left and right LVD were not associated with GA (Online Supplemental Data and Fig 2, right). Associations were in the same direction as the associations observed for the manual measures (Online Supplemental Data and Fig 2, right). Measures derived from early versus TEA MR imaging showed a similar association with GA (Fig 2, right), except that the associations with GA for the manual biparietal width and automated callosal thickness at the splenium were found at early MR imaging only. Associations between the measures with birth weight were similar to those found for gestational age (Online Supplemental Data).

DISCUSSION

In this study, we developed and validated a method to automatically derive 6 measures of brain growth and development based on raw measurements of distance and area (biparietal width, interhemispheric distance, callosal thickness, transcerebellar diameter, LVD, and DGMA). These measures account for approximately half the measures of the standardized conventional MR imaging scoring system.[15] Conventional MR imaging scores have been linked to outcomes and may assist with early diagnosis and streamlining identified at-risk infants into early intervention. We automated these measures for early and TEA MR imaging to make scoring more time-efficient and objective, enabling its use in clinical and research settings. Automated measures (except for interhemispheric distance and callosal thickness) showed good agreement and reliability as well as a strong association with manual measures in the prospective PPREMO cohort (early MR imaging, $n = 94$; TEA MR imaging, $n = 103$). Furthermore, all automated measures showed similar or stronger associations with PMA compared with the manual measures, which were replicated in an independent cohort including mostly TEA MR imaging (dHCP $n = 455$, except for the right LVD). All measures (except the LVD) showed an association with GA, though not all associations were found across cohorts.

Our findings are in line with those in our previous study of the PPREMO cohort,[16] which reported several positive associations between the manual measures with PMA, as part of validating the manual measures for preterm-born cohorts derived from early and TEA MR imaging. Other studies[15,33] have shown positive associations with PMA for the manual biparietal width, transcerebellar diameter, and DGMA—the measures for which we found the strongest associations. In addition, most associations with PMA were similar for the automated versus manual measures, but the automated callosal thickness measures were

more strongly associated with PMA. The automated callosal thickness measures may be more sensitive than the manual measures because these were derived from multiple measurement points.

Associations with PMA were somewhat different for early versus TEA MR imaging. All measures were associated with PMA at early MR imaging, but only the transcerebellar diameter and DGMA were associated with PMA at TEA MR imaging, in line with others showing no association with PMA for interhemispheric distance, callosal thickness, and LVD when deriving measures at TEA MR imaging[15,33] or in the first year.[34] Regional differences in growth trajectories may explain these inconsistencies between early and TEA MR imaging. Between 27 and 45 weeks' PMA, total brain growth peaks at 35 weeks,[35] but regional differences exist, with WM and subcortical GM peaking earlier (at 33 and 31 weeks, respectively),[35] whereas the cerebellum (at 37 weeks) as well as CSF and cortical GM (both at 38–39 weeks) peak later.[35] Acquiring MR imaging measures at both early and TEA MR imaging may provide additional information on whether brain growth is progressing as expected, which could be valuable for prediction modeling of clinical outcomes.

All automated measures (except LVD) were associated with GA in at least 1 cohort, similar to findings in previous studies using the manual measures,[33,36] except for 1 study[33] finding a negative association between GA and LVD. These findings reflect impaired brain growth in preterm-born infants compared with term-born infants, in line with others[15] who showed smaller biparietal width, callosal thickness, transcerebellar diameter, and DGMA but larger interhemispheric distance and LVD in preterm compared with term-born infants. Furthermore, a recent large volumetric study[35] ($n = 420$) showed larger ventricular volumes and smaller total GM (cortical and subcortical) and WM volumes with lower GA. Another large study[37] ($n = 285$) found smaller volumes for most brain regions with lower GA, but other regions (primary visual, motor, and somatosensory regions) were larger with lower GA, possibly reflecting increased development in response to the ex-utero environment.

There are some limitations to this study. First, the automated method to extract the regional conventional MR imaging measures depends on an accurate segmentation of the brain, resulting in only 70%–86% of the total scans being suitable for our Matlab pipeline (depending on the measure) because we visually inspected segmentations and excluded scans with segmentation errors, poor scan quality (eg, motion), and severe pathology. Different segmentation pipelines may influence the measures to some extent, but differences between measures from different segmentations are likely to be the result of segmentation errors rather than differences between segmentations per se, with the segmentation quality influenced by the age of the templates used and different data used to derive the algorithm. Our aim to automate the Kidokoro measures[15] included the idea that our Matlab pipeline is not specific to 1 particular segmentation software, particularly because different pipelines may work better for a specific patient cohort than others.

Although we aimed to make the automated measures as similar to the manual measures as possible, some differences between measures likely exist due to differences in measurement

techniques. For example, the automated interhemispheric distance is derived from the distance between the voxels of the superior frontal gyri, which may not always measure the distance at the crowns. We focused on 2D measures to enable comparison between manual and automated measures, allowing for manual scoring of scans without a successful brain parcellation, for example in cases with severe pathology. These 2D measures, however, are likely influenced by variation in slice selection (particularly interhemispheric distance and LVD), which could be improved using 3D measures. Furthermore, findings were inconsistent for the manual versus automated callosal thickness measures, with negative associations found for the manual measures with PMA at early MR imaging, but positive associations were found for the automated measures with PMA.

Bland-Altman plots showed a bias. Automated measures may overestimate thickness for smaller values, but manual measures may underestimate thickness. Agreement and reliability for manual callosal thickness measures were poor and poor-to-moderate, respectively (Online Supplemental Data), and other studies[35,38] showed a positive growing trajectory for callosal thickness, suggesting that findings for the manual measures at early MR imaging may be the result of measurement error. Manual measurements of callosal thickness are subject to several challenges because the corpus callosum is very thin, particularly at early MR imaging. Thickness may be difficult to measure with a voxel size of 0.7 mm with the ruler not being sensitive to very small increments, and partial volume, technical artifacts (eg, motion), and the presence of the cavum septi pellucidi may impact the visualization.

Our study provides initial automating of conventional MR imaging scoring for preterm neonates, which enables clinicians to use the scoring system without manually measuring the 6 measures of distance and area, resulting in a more time-efficient and consistent routine neonatal MR imaging clinical scoring system. Our automated method enables extraction for large research data sets, facilitating the development of well-established cutoff scores and more reliable and clinically relevant prediction models to assist clinicians with early diagnosis and intervention. Future work is needed to establish a reference cohort to define the cutoffs for the automated measures and to examine the association between the measures and prospective clinical outcomes.

CONCLUSIONS
We validated an approach to automatically extract approximately half the measures of a conventional brain MR imaging scoring system for preterm-born infants using T2WI acquired between 29 to 47 weeks' PMA, which will enhance the use of this scoring system in both clinical and research settings.

ACKNOWLEDGMENTS

REFERENCES

1. George JM, Fiori S, Fripp J, et al. **Relationship between very early brain structure and neuromotor, neurological and neurobehavioral function in infants born <31 weeks gestational age**. *Early Hum Dev* 2018;117:74–82 CrossRef Medline
2. Allotey J, Zamora J, Cheong-See F, et al. **Cognitive, motor, behavioural and academic performances of children born preterm: a meta-analysis and systematic review involving 64 061 children**. *BJOG* 2018;125:16–25 CrossRef Medline
3. Evensen KA, Ustad T, Tikanmäki M, et al. **Long-term motor outcomes of very preterm and/or very low birth weight individuals without cerebral palsy: a review of the current evidence.** *Semin Fetal Neonatal Med* 2020;25:101116 CrossRef Medline
4. McBryde M, Fitzallen GC, Liley HG, et al. **Academic outcomes of school-aged children born preterm: a systematic review and meta-analysis.** *JAMA Netw Open* 2020;3:e202027 CrossRef Medline
5. Fitzallen GC, Taylor HG, Bora S. **What do we know about the preterm behavioral phenotype? A narrative review.** *Front Psychiatry* 2020;11:154 CrossRef Medline
6. Cheong JL, Doyle LW, Burnett AC, et al. **Association between moderate and late preterm birth and neurodevelopment and social-emotional development at age 2 years.** *JAMA Pediatr* 2017;171:e164805 CrossRef Medline
7. Cheong JL, Spittle AJ, Burnett AC, et al. **Have outcomes following extremely preterm birth improved over time?** *Semin Fetal Neonatal Med* 2020;25:101114 CrossRef Medline
8. Wolke D, Johnson S, Mendonça M. **The life course consequences of very preterm birth.** *Annu Rev Dev Psychol* 2019;1:69–92 CrossRef
9. Jöud A, Sehlstedt A, Källén K, et al. **Associations between antenatal and perinatal risk factors and cerebral palsy: a Swedish cohort study.** *BMJ Open* 2020;10:e038453 CrossRef Medline
10. Galea C, Mcintyre S, Smithers-Sheedy H, et al. Australian Cerebral Palsy Register Group. **Cerebral palsy trends in Australia (1995–2009): a population-based observational study.** *Dev Med Child Neurol* 2019;61:186–93 CrossRef Medline
11. Patel RM. **Short- and long-term outcomes for extremely preterm infants.** *Am J Perinatol* 2016;33:318–28 CrossRef Medline
12. Dubois J, Benders M, Borradori-Tolsa C, et al. **Primary cortical folding in the human newborn: an early marker of later functional development.** *Brain* 2008;131:2028–41 CrossRef Medline

13. Makropoulos A, Robinson EC, Schuh A, et al. **The developing human connectome project: a minimal processing pipeline for neonatal cortical surface reconstruction.** *Neuroimage* 2018;173:88–11 CrossRef Medline

14. Rose J, Vassar R, Cahill-Rowley K, et al. **Neonatal physiological correlates of near-term brain development on MRI and DTI in very-low-birth-weight preterm infants.** *Neuroimage Clin* 2014;5:169–77 CrossRef Medline

15. Kidokoro H, Neil JJ, Inder TE. **New MR imaging assessment tool to define brain abnormalities in very preterm infants at term.** *AJNR Am J Neuroradiol* 2013;34:2208–14 CrossRef Medline

16. George JM, Fiori S, Fripp J, et al. **Validation of an MRI brain injury and growth scoring system in very preterm infants scanned at 29- to 35-week postmenstrual age.** *AJNR Am J Neuroradiol* 2017;38:1435–42 CrossRef Medline

17. Woodward LJ, Anderson PJ, Austin NC, et al. **Neonatal MRI to predict neurodevelopmental outcomes in preterm infants.** *N Engl J Med* 2006;355:685–94 CrossRef Medline

18. Dewan MV, Herrmann R, Schweiger B, et al. **Are simple magnetic resonance imaging biomarkers predictive of neurodevelopmental outcome at two years in very preterm infants?** *Neonatology* 2019;116:331–40 CrossRef Medline

19. Hintz SR, Barnes PD, Bulas D, et al. SUPPORT Study Group of the Eunice Kennedy Shriver National Institute of Child Health and Human Development Neonatal Research Network. **Neuroimaging and neurodevelopmental outcome in extremely preterm infants.** *Pediatrics* 2015;135:e32–42 CrossRef Medline

20. Brouwer MJ, Kersbergen KJ, van Kooij BJ, et al. **Preterm brain injury on term-equivalent age MRI in relation to perinatal factors and neurodevelopmental outcome at two years.** *PLoS One* 2017;12: e0177128 CrossRef Medline

21. Slaughter LA, Bonfante-Mejia E, Hintz SR, et al. **Early conventional MRI for prediction of neurodevelopmental impairment in extremely-low-birth-weight infants.** *Neonatology* 2016;110:47–54 CrossRef Medline

22. Spittle AJ, Cheong J, Doyle LW, et al. **Neonatal white matter abnormality predicts childhood motor impairment in very preterm children.** *Dev Med Child Neurol* 2011;53:1000–06 CrossRef Medline

23. Woodward LJ, Clark CA, Bora S, et al. **Neonatal white matter abnormalities an important predictor of neurocognitive outcome for very preterm children.** *PLoS One* 2012;7:e51879 CrossRef Medline

24. Anderson PJ, Treyvaud K, Neil JJ, et al. **Associations of newborn brain magnetic resonance imaging with long-term neurodevelopmental impairments in very preterm children.** *J Pediatr* 2017;187:58–65 CrossRef Medline

25. Iwata S, Nakamura T, Hizume E, et al. **Qualitative brain MRI at term and cognitive outcomes at 9 years after very preterm birth.** *Pediatrics* 2012;129:e1138–47 CrossRef Medline

26. George JM, Boyd RN, Colditz PB, et al. **PPREMO: a prospective cohort study of preterm infant brain structure and function to predict neurodevelopmental outcome.** *BMC Pediatr* 2015;15:123 CrossRef Medline

27. Hughes EJ, Winchman T, Padormo F, et al. **A dedicated neonatal brain imaging system.** *Magn Reson Med* 2017;78:794–804 CrossRef Medline

28. Cordero-Grande L, Teixeira RP, Hughes EJ, et al. **Sensitivity encoding for aligned multishot magnetic resonance reconstruction.** *IEEE Trans Comput Imaging* 2016;2:266–80 CrossRef

29. Gousias IS, Hammers A, Counsell SJ, et al. **Magnetic resonance imaging of the newborn brain: automatic segmentation of brain images into 50 anatomical regions.** *PLoS One* 2013;8:e59990 CrossRef Medline

30. Gousias IS, Edwards AD, Rutherford MA, et al. **Magnetic resonance imaging of the newborn brain: manual segmentation of labelled atlases in term-born and preterm infants.** *Neuroimage* 2012;62:1499–1509 CrossRef Medline

31. Bourgeat P, Dore V, Villemagne VL, et al. **MilxXplore: a web-based system to explore large imaging datasets.** *J Am Med Inform Assoc* 2013;20:1046–52 CrossRef Medline

32. Jenkinson M, Beckmann CF, Behrens TE, et al. **FSL.** *Neuroimage* 2012;62:782–90 CrossRef Medline

33. Iwata S, Katayama R, Kinoshita M, et al. **Region-specific growth restriction of brain following preterm birth.** *Sci Rep* 2016;6:33995 CrossRef Medline

34. Mongerson CR, Wilcox SL, Goins SM, et al. **Infant brain structural MRI analysis in the context of thoracic non-cardiac surgery and critical care.** *Front Pediatr* 2019;7:315 CrossRef Medline

35. Makropoulos A, Aljabar P, Wright R, et al. **Regional growth and atlasing of the developing human brain.** *Neuroimage* 2016;125:456–78 CrossRef Medline

36. Kidokoro H, Anderson PJ, Doyle LW, et al. **Brain injury and altered brain growth in preterm infants: predictors and prognosis.** *Pediatrics* 2014;134:e444–53 CrossRef Medline

37. Alexander B, Kelly CE, Adamson C, et al. **Changes in neonatal regional brain volume associated with preterm birth and perinatal factors.** *Neuroimage* 2019;185:654–63 CrossRef Medline

38. Teli R, Hay M, Hershey A, et al. **Postnatal microstructural developmental trajectory of corpus callosum subregions and relationship to clinical factors in very preterm infants.** *Sci Rep* 2018;8:7550 CrossRef Medline

Correlation between 2D and 3D Fetal Brain MRI Biometry and Neurodevelopmental Outcomes in Fetuses with Suspected Microcephaly and Macrocephaly

S. Fried, M. Gafner, D. Jeddah, N. Gosher, D. Hoffman, R. Ber, A. Mayer, and E. Katorza

ABSTRACT

BACKGROUND AND PURPOSE: Definitions of fetal microcephaly and macrocephaly are debatable. A better understanding of their long-term prognoses would help guide parental education and counseling. This study aimed to explore the correlation between 2D and 3D fetal brain MR imaging biometry results and the long-term neurodevelopmental outcomes.

MATERIALS AND METHODS: This analysis is a historical cohort study. Fetal brain biometry was measured on 2D and 3D MR imaging using a volumetric MR imaging semiautomated algorithm. We measured and assessed the following brain structures: the supratentorial brain volume and cerebellar volume and cerebellar volume/supratentorial brain volume ratio, in addition to commonly used 2D brain MR imaging biometric variables, including occipitofrontal diameter, biparietal diameter, and transcerebellar diameter. Microcephaly was defined as ≤ 3rd percentile; and macrocephaly, as ≥ 97th percentile, corresponding to −2 SDs and +2 SDs. The neurodevelopmental outcome of this study cohort was evaluated using the Vineland-II Adaptive Behavior Scales, and the measurements were correlated to the Vineland standard scores.

RESULTS: A total of 70 fetuses were included. No significant correlation was observed between the Vineland scores and either the supratentorial brain volume, cerebellar volume, or supratentorial brain volume/cerebellar volume ratio in 3D or 2D MR imaging measurements, after correction for multiple comparisons. No differences were found among fetuses with macrocephaly, normocephaly, or microcephaly regarding the median Vineland standard scores.

CONCLUSIONS: Provided there is normal brain structure on MR imaging, the developmental milestone achievements in early years are unrelated to 2D and 3D fetal brain MR imaging biometry, in the range of measurements depicted in this study.

ABBREVIATIONS: BPD = biparietal diameter; CV = cerebellar volume; HC = head circumference; IQR = interquartile range; OFD = occipitofrontal diameter; STV = supratentorial brain volume; TCD = transcerebellar diameter; US = ultrasound

Measurement of head circumference (HC) is a key element in the routine prenatal ultrasonographic screening of the fetus. Fetal macrocephaly and microcephaly are defined as an ultrasonographic occipitofrontal circumference of >2 SDs above the mean and 3 SDs below the mean, respectively.[1,2] There is a lack of published data regarding the correlation between brain MR imaging biometry and neurodevelopmental outcome.

When macrocephaly appears prenatally, its diagnosis is usually made during the third trimester, especially when no other anomalies are found.[3] There is no consensus with respect to the neurodevelopmental outcomes of macrocephaly. Some studies describe this condition as benign, while others associate it with lower intelligence and motor function.[3-6] Similar results are often observed regarding isolated fetal microcephaly, while definitions of this condition are debatable.[7] Leibovitz et al[8] described the tendency toward overdiagnosis of microcephaly in prenatal ultrasound (US).[8] Stoler-Poria et al[7] showed that prenatal diagnosis of a head circumference between 2 and 3 SDs below the mean does not portend an abnormal neuropsychological development.

When suspicion of an abnormal head circumference arises, in many medical centers, it is a common practice to refer for fetal brain MR imaging to obtain additional information regarding brain development. Detailed anatomic US screening, neurosonography, and genetic counseling are also frequently obtained.[9,10]

A previous study from our group demonstrated significant agreement among US, 2D MR imaging, 3D MR imaging, and the head circumference at birth in the classification of suspected microcephaly and macrocephaly. It raised the question of the additional value of fetal brain MR imaging in the evaluation of isolated fetal microcephaly and macrocephaly.[11] We expand on this previous study and aim to explore this question with the same cohort of fetuses by examining the neurodevelopmental outcomes of fetuses according to their measured brain biometry on 2D and 3D MR imaging.

MATERIALS AND METHODS

This historical cohort study was approved by the Chaim Sheba Medical Center institutional review board (0256–13-SMC). Informed consent was obtained from all women who participated in this study.

Subjects

The subjects of this study were all women with singleton pregnancies who underwent fetal brain MR imaging at our tertiary medical center between 2011 and 2018 during their third trimester. Given the study aims, we recruited the subjects in clusters according to the indication of their MR imaging examinations. These included individuals with either US-suspected microcephaly or macrocephaly and those with non-neurologic reasons such as a sibling with a neurologic abnormality for the normocephaly group.

The inclusion criteria for this study were MR imaging scans that were clear and unambiguous enabling easy assessment of measures, a full neurosonogram examination performed by an obstetrician/gynecologist expert in fetal sonography, and delivery at our medical center. Exclusion criteria were suspicion of intrauterine growth retardation, women with a multiple pregnancy, an associated major brain anomaly, termination of pregnancy, and unwillingness of parents to participate in the interview. We included fetuses with the following findings that are considered a normal or minor variant:[12] a narrow cavum septum pellucidum, dolichocephaly, minor dilation of the subarachnoid space, mega cisterna magna, and short intact corpus callosum, as has been acceptable in similar studies.[13-15]

Information regarding the pregnancy, fetal sex, genetic testing, and US findings was obtained from the medical records.

MR Imaging. This study was based on our routine fetal MR imaging procedure. We used a 1.5T system. Single-shot fast spin-echo T2-weighted sequences in 3 orthogonal planes were performed using a half-Fourier technique (number of excitations = 0.53) with the following parameters: section thickness = 3 or 4 mm, no gap, flexible coil (8-channel cardiac coil). FOV was determined by the size of the fetal head with a range of 24 × 24 cm to 30 × 30 cm, acquisition time = 40–45 seconds, matrix = 320/224, TE = 90 ms, TR = 1298 ms, pixel bandwidth = 122 Hz/pixel. Specific absorption rate values were between 1.1 and 1.7 W/kg.[16]

Semiautomated Algorithm of Volumetric MR Imaging Measurements

We used a semiautomated algorithm developed by our group.[13] The algorithm is based on the initial manually outlined contour of the ROI in the brain parenchyma of the midcoronal section. On the basis of the signal intensity, the contour is then automatically propagated into a more accurate one. The algorithm then proceeds to propagate forward and backward, drawing contours automatically for all the slices. Later, manual adjustments are made to attain maximal precision. The total volume is calculated, summing up the volume of voxels enclosed by each contour.

Measurements

The measurements were performed on the coronal plane. We measured and assessed 2 brain structures, the supratentorial brain volume (STV) and the cerebellar volume (CV), and calculated the ratio between the cerebellar volume and the supratentorial brain volume (CV/STV). These volumes were converted to percentiles according to percentile tables derived from healthy fetuses, obtained from a previous publication from our group.[13] The anatomic boundaries of the brain structures are as follows:[17]

Supratentorial Brain. The supratentorial brain included the parenchyma of the frontal, parietal, occipital, and temporal lobes, including the third ventricle. The lateral ventricles, cerebellum, brain stem, and fourth ventricle were excluded. The boundaries of the parenchyma were defined as the outer edge of the cerebral cortex. The inferior border matched the cortex and the "imagined" line crossing the brain stem between the edges of the tentorium cerebelli.

Cerebellum. The cerebellum included the cerebellar peduncles and vermis and excluded the brain stem and the fourth ventricle.

In addition, 2D MR imaging measurements, performed for clinical purposes by an expert neuroradiologist (E.K.), were obtained from the original MR imaging evaluation, namely, the occipitofrontal diameter (OFD), biparietal diameter (BPD), and transcerebellar diameter (TCD). They were then converted to percentiles according to reference ranges of Tilea et al.[18]

Microcephaly was defined as equal or less than the 3rd percentile, and macrocephaly, as equal or more than the 97th percentile, corresponding the threshold of −2 SDs and +2 SDs.

Neurodevelopmental Outcome

The fetuses were evaluated at least 6 months after birth using the Vineland-II Adaptive Behavior Scales.[19] This is a commonly used interview-based questionnaire that evaluates 4 domains of development of the child from birth to adulthood. These include communication skills, motor skills, socialization skills, and daily living skills.

Because the expression of adaptive behavior changes throughout life, the scores are age-normalized and then added to an adaptive behavior composite,[20] which is characterized by a mean score of 100 and an SD of 15.[21] This study used the published Israeli version of the questionnaire. Because Israeli norms for the full age ranges are not available and because there was no reason to believe that children in the United States and Israel will develop motor and mental skills differently, we used the American

Table 1: Descriptive statistics of the study population

Statistics	
Maternal age at birth (median) (IQR) (yr)	34 (29.8–36.0)
Fetal sex	
Male (No.) (%)	34 (48.6%)
Female (No.) (%)	36 (51.4%)
Abnormal triple test findings (No.) (%)	3 (4.3%)
Abnormal nuchal translucency findings (No.) (%)	3 (4.3%)
Karyotype (No.) (%)	
Not performed	31 (44.3%)
Normal findings	39 (55.7%)
Abnormal findings	0 (0%)
Chromosomal microarray (No.) (%)	
Not performed	48 (68.6%)
Normal findings	21 (30%)
Abnormal findings	1 (1.4%)
Length of pregnancy (mean) (SD) (wk)	38.8 (1.3)

Table 2: 2D and 3D fetal brain MR imaging biometric parameters

Characteristics	Median	IQR
OFD (mm)	98.5	93.0–102.3
OFD (centile)	57.5	28.8–90.0
BPD (mm)	77.0	72.8–80.0
BPD (centile)	40.0	10.0–75.0
TCD (mm)	45.0	41.9–48.0
TCD (centile)	50.0	40.0–75.0
STV (mm^3)	220.294	192,195–248,017
STV (centile)	28.0	4.0–84.3
CV (mm^3)	12.683	10,604–15,114
CV (centile)	48.5	12.8–77.5
CV/STV	0.058	0.052–0.065
CV/STV (centile)	58.5	32.0–88.5

norms. An abnormal Vineland score was defined as a standard score of a subdomain or as an adaptive-level composite standard score of <86 (low or moderately low adaptive level).[22]

Statistical Analysis

Categoric variables were presented as frequency and percentage. Continuous variables were summarized as mean and SD or median and interquartile range (IQR). The Spearman rank correlation coefficient was used to study the association between the Vineland scores and brain parameters. In further analysis, the brain parameters (percentiles) were categorized using the 3rd and 97th percentiles as threshold values for microcephaly and macrocephaly, respectively. The median Vineland scores were compared between healthy fetuses and those with microcephaly and between healthy fetuses and those with macrocephaly using the Mann-Whitney U test. The Benjamini-Hochberg procedure (false discovery rate) was used to adjust the P values to avoid errors due to multiple comparisons. Differences between our cohort and the fetuses that were lost to follow up were studied using the Mann-Whitney U or χ^2 test. All statistical tests were 2-sided, and $P < .05$ was considered statistically significant. SPSS software was used for all statistical analyses (IBM, 2017).

Intraclass Correlation Coefficient

Interobserver and intraobserver variability were evaluated by comparing measurements of 10 fetuses from this cohort made by 2

independent observers with 10 other measurements measured twice by the same observer, respectively. We calculated the intraclass correlation coefficient and its 95% confidence intervals. We defined good agreement as $0.6 \leq$ intraclass correlation coefficient ≤ 0.74 and excellent agreement as intraclass correlation coefficient ≥ 0.75.[23]

RESULTS

A total of 88 fetuses were recruited for the study. Eighteen (20%) fetuses were excluded from the study. Sixteen (18%) fetuses were either lost to follow-up or lost due to parental refusal for an interview. These fetuses were found to be similar to the rest of the cohort (Online Supplemental Data). The other 2 fetuses (2%) were excluded due to termination of pregnancy. The demographic details of the mothers and fetuses are shown in Table 1. Obstetric data of the cohort are presented in the Online Supplemental Data.

MR Imaging Biometrics

The median gestational age at time of MR imaging was 34.3 weeks (IQR, 32.5–36.6 weeks). The 2D and 3D measurements are presented in Table 2. A total of 16 (23%) and 11 (16%) fetuses had STV s of ≤3rd and ≥ 97th percentiles, respectively. This was achieved by recruiting the fetuses in clusters according to the indication for the MR imaging, to have enough data points at the margins of the curve. Of these fetuses, 10 (14%) were ≤1st percentile and 8 (11%) were ≥ 99th percentile.

Neurodevelopmental Outcomes

The Vineland-II Adaptive Behavior scores of our cohort are presented in the Online Supplemental Data. The median age of the children at time of the interview was 55.5 months (IQR, 26.8–75.0 months). The correlations between the biometric parameters and the Vineland composite scores are presented in Figs 1 and 2 for macrocephaly and microcephaly, respectively, as well as in the Online Supplemental Data. We initially found negative correlations between the 2D MR imaging BPD and the daily living skills subdomain score, 3D MR imaging STV and the daily living skills subdomain score, and the 3D MR imaging CV/STV ratio and the motor and daily living skills subdomains scores. However, after we used the Benjamini–Hochberg procedure to adjust the P values for multiple comparisons, all the above correlations were found to be statistically insignificant. Furthermore, no significant differences were found when comparing the median Vineland scores between fetuses with microcephaly and healthy fetuses and fetuses with macrocephaly and healthy fetuses, using the 3rd and 97th percentiles as thresholds, respectively (Tables 3 and 4; Online Supplemental Data).

The only child who had a low adaptive behavior composite score in the cohort had a diagnosis of Soto syndrome, a rare autosomal dominant overgrowth syndrome that is associated with cognitive impairment, which was made at 4 months of age. This fetus did not have extreme biometry, with measurements of the 90th percentile for the OFD and BPD, 10th percentile for the TCD, 98th percentile for STV, 52nd percentile for CV, and 5th percentile for the CV/STV ratio. The 8 children who had moderately low composite scores had quite variable brain volumes: Two had an STV ≤ 3rd percentile, 2 had an STV ≥ 97th percentile, and 4 had a normal STV. Regarding genetics, only 1 fetus had
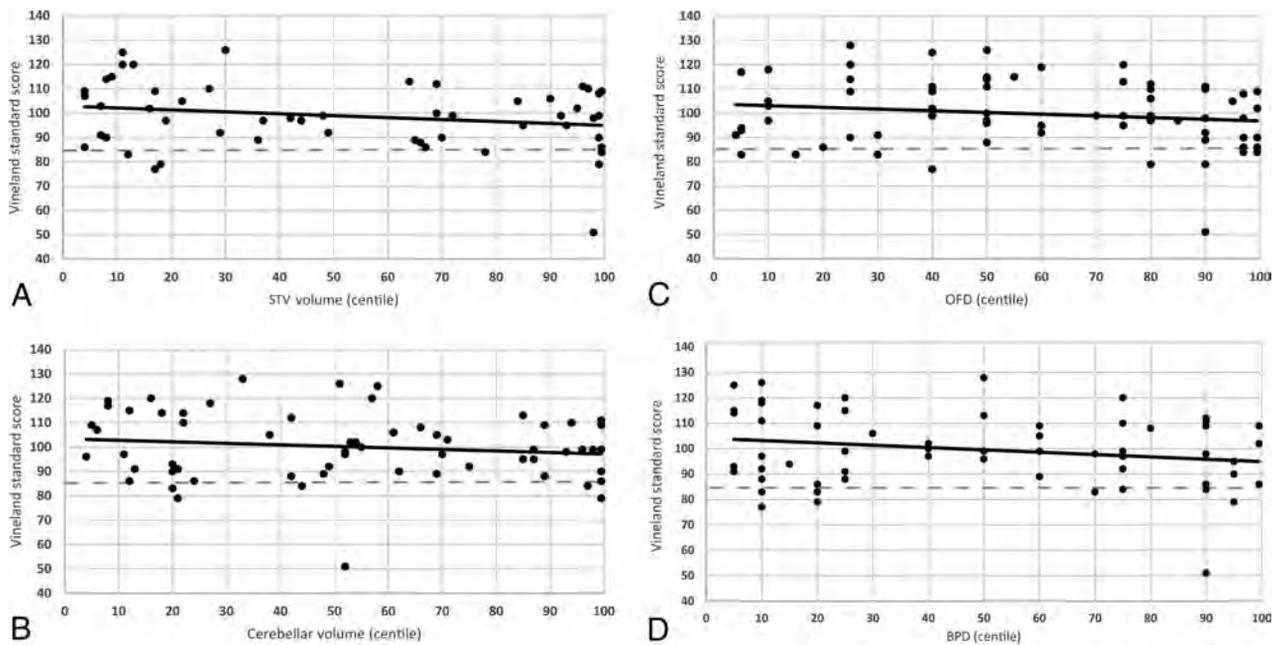
FIG 1. Scatterplots of the Vineland II Adaptive Behavior standard score according to STV (*A*), CV (*B*), OFD (*C*), and BPD (*D*). Macrocephalic and normocephalic biometry.
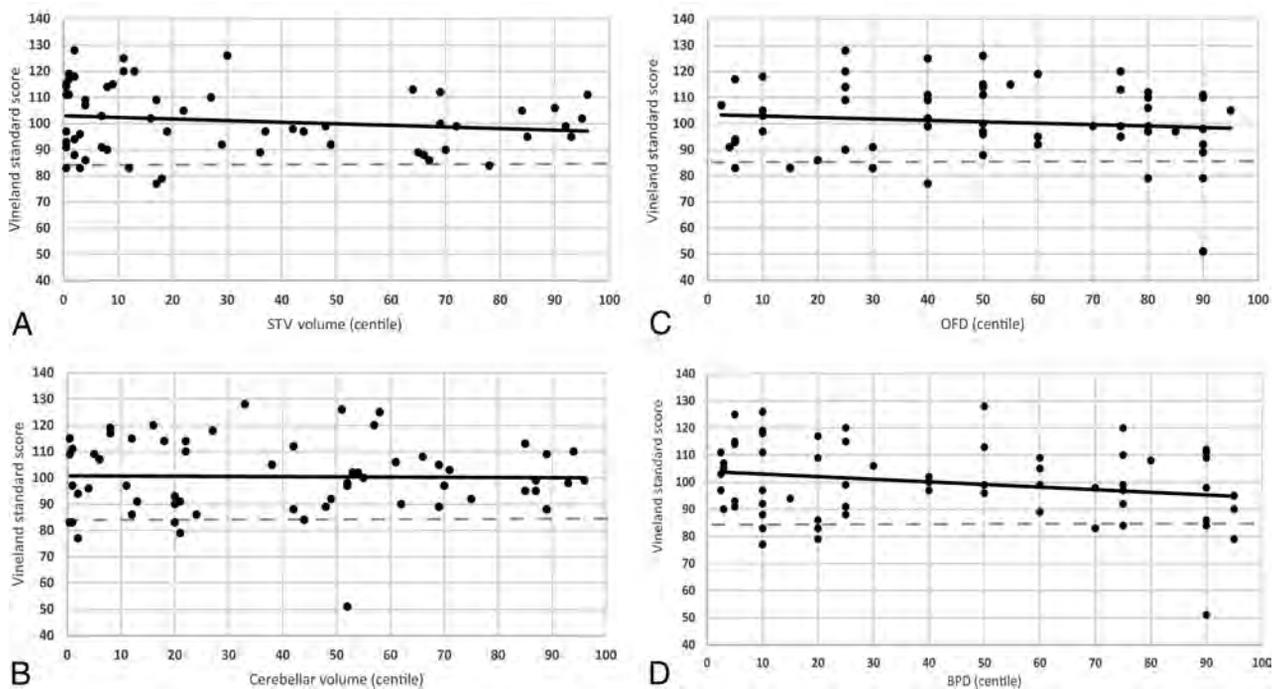
FIG 2. Scatterplots of the Vineland II Adaptive Behavior standard score according to STV (*A*), CV (*B*), OFD (*C*), and BPD (*D*). Microcephalic and normocephalic biometry.

abnormal findings on a chromosomal microarray test. This fetus did not have an abnormal Vineland score.

Interobserver and Intraobserver Variability
All volumetric intraclass correlation coefficients for both inter-observer and intraobserver results were excellent (Online Supplemental Data).

DISCUSSION
In this study, we aimed to find correlations between 2D and 3D fetal brain MR imaging biometry and long-term neurodevelopmental outcomes using the Vineland-II Adaptive Behavior Scales. The results of this study show no correlation among MR imaging STV, CV, the STV/CV ratio, OFD, BPD, or TCD and the neurodevelopmental outcomes when tested as

Table 3: Comparison of Vineland-II Adaptive Behavior standard scores between fetuses with microcephaly and healthy fetuses according to 2D and 3D fetal brain MR imaging biometry[a]

Percentile of biometric value	4–96	≤3	P Value
MR imaging 3D STV (No. of fetuses)	43	16	
Adaptive behavior composite score	99 (109–90)	104 (116–91)	.44
MR imaging 3D CV (No. of fetuses)	53	9	
Adaptive behavior composite score	99 (111–91)	97 (111–83)	.51
MR imaging 2D OFD (No. of fetuses)	58	1	
Adaptive behavior composite score	99 (111–71)	107	NA
MR imaging 2D BPD (No. of fetuses)	61	6	
Adaptive behavior composite score	98 (111–89)	104 (108–95)	.68

Note:—NA indicates not applicable.
[a] Continuous variables are presented as median (IQR). P value refers to the Mann-Whitney U test.

Table 4: Comparison of Vineland-II Adaptive Behavior standard scores between fetuses with macrocephaly and healthy fetuses according to 2D and 3D fetal brain MR imaging biometry[a]

Percentile of biometric value	4–96	≥97	P Value
MR imaging 3D STV (No. of fetuses)	43	11	
Adaptive behavior composite standard score	99 (109–90)	98 (109–84)	.31
MR imaging 3D CV (No. of fetuses)	53	8	
Adaptive behavior composite standard score	99 (111–91)	94 (106–84)	.19
MR imaging 2D OFD (No. of fetuses)	58	11	
Adaptive behavior composite standard score	99 (111–71)	90 (108–86)	.13
MR imaging 2D BPD (No. of fetuses)	61	3	
Adaptive behavior composite standard score	98 (111–89)	102	.93

[a] Continuous variables are presented as median (IQR). P value refers to the Mann-Whitney U test.

continuous variables or when categorized into normal versus abnormal biometry.

When microcephaly or macrocephaly is suspected in a fetus, it is often a challenge to counsel parents, especially when fetuses lack extreme head biometry or additional malformations. Accurate detection of those conditions is important to avoid overdiagnosis that might result in unjustified termination of pregnancy and underdiagnosis that might result in the birth of a child carrying an unfavorable prognosis.

There is no consensus in the literature regarding the neurodevelopmental outcome of infants with macrocephaly. An association between macrocephaly and low intelligence or cognitive or motor function is often reported,[6,24,25] yet other publications describe healthy or even better functioning children.[3,26,27] Information regarding the neurodevelopmental outcome of fetuses diagnosed in utero with macrocephaly is scarce. Biran-Gol et al[5] found that an in utero sonographic finding of a head circumference between 2 and 3 SDs above the mean for gestational age, when no associated brain malformations are present, does not necessarily portend long-term abnormal neuropsychological development. They also acknowledged the low specificity of a fetal US–measured enlarged head circumference, raising the question of whether fetal sonography can accurately predict macrocephaly at birth. Here, we present similar results, because no correlation was found between a large supratentorial brain or cerebellum and neurodevelopmental outcomes of fetuses, meaning that even the finding of a large brain parenchymal volume on MR imaging does not indicate a poor neurologic prognosis.

The definition of fetal microcephaly is also a matter of controversy.[7,28,29] MR imaging evaluation of the fetal brain is usually performed following the sonographic suspicion of microcephaly, allowing the assessment of the brain parenchyma directly and the gathering of more information regarding the development of the fetus's brain.[9]

The neurodevelopmental outcome of fetal microcephaly has been examined in different studies. Kurtz et al[30] argued that the measurement of fetal HC 1–3 SDs below the mean on a US examination carries a reasonable hope that the fetuses will develop normally. Stoler-Poria et al[7] found a negative linear correlation between a smaller HC in utero and later cognitive function but no adverse neuropsychological outcomes in children with fetal HCs 2–3 SDs below the norm. According to the Society of Obstetricians and Gynecologists of Canada, when the HC measurement is lower than 3 SDs below the norm, it should be considered an anomaly.[31] Deloison et al[32] showed, in a relatively large population of fetuses with US-measured HC lower than the 5th percentile, that those with poorer outcomes had other associated findings rather than isolated microcephaly. Our results were similar; we did not find any difference in the neurodevelopmental outcome of fetuses with brain biometry ≤ 3rd percentile, compared with the 4th to 96th percentiles.

A limitation of our study is that MR imaging examinations were performed in a relatively advanced week of gestation compared with those at other centers worldwide. This might be because suspicion of microcephaly and macrocephaly usually does not appear on US before the third trimester.[5,33] This feature might also reflect the local policy that allows termination of pregnancy in the second and third trimesters in cases of suspected severe fetal disabilities including brain pathologies.[34] Furthermore, although the Vineland-II Adaptive Behavior scores are commonly applied in children of birth to toddler age, subtle developmental delays may not be apparent until an older age and might have been missed in this study.[35]

This study is unique in that it focuses on microcephaly and macrocephaly, defined as abnormal head circumference findings, through the prism of measurements obtained on MR imaging. This feature is important because to our knowledge, no definitions of these pathologic conditions, when detected on MR imaging, are available. The strengths of this study are a relatively large cohort of fetuses and a long period of follow-up.

CONCLUSIONS

This study shows that developmental milestone achievements are unrelated to the MR imaging biometry of the fetal brain. This

information can be useful for physicians when counseling parents of fetuses with suspected microcephaly and macrocephaly detected on MR imaging. Further studies with additional data are needed to better reassure parents of fetuses with large and small brain biometry when MR imaging findings are normal.

REFERENCES

1. DeRosa R, Lenke RR, Kurczynski TW, et al. **In utero diagnosis of benign fetal macrocephaly.** *Am J Obstet Gynecol* 1989;161:690–92 CrossRef Medline

2. Chervenak FA, Jeanty P, Cantraine F, et al. **The diagnosis of fetal microcephaly.** *Am J Obstet Gynecol* 1984;149:512–57 CrossRef Medline

3. Day RE, Schutt WH. **Normal children with large heads: benign familial megalencephaly.** *Arch Dis Child* 1979;54:512–17 CrossRef Medline

4. Lewis BA, Aram DM, Horwitz SJ. **Language and motor findings in benign megalencephaly.** *Percept Mot Skills* 1989;68:1051–54 CrossRef Medline

5. Biran-Gol Y, Malinger G, Cohen H, et al. **Developmental outcome of isolated fetal macrocephaly.** *Ultrasound Obstet Gynecol* 2010;36:147–53 CrossRef Medline

6. Petersson S, Pedersen NL, Schalling M, et al. **Primary megalencephaly at birth and low intelligence level.** *Neurology* 1999;53:1254–59 CrossRef Medline

7. Stoler-Poria S, Lev D, Schweiger A, et al. **Developmental outcome of isolated fetal microcephaly.** *Ultrasound Obstet Gynecol* 2010;36:154–58 CrossRef Medline

8. Leibovitz Z, Daniel-Spiegel E, Malinger G, et al. **Prediction of microcephaly at birth using three reference ranges for fetal head circumference: can we improve prenatal diagnosis?** *Ultrasound Obstet Gynecol* 2016;47:586–92 CrossRef Medline

9. Nawathe A, Doherty J, Pandya P. **Fetal microcephaly.** *BMJ* 2018;361:k2232 CrossRef Medline

10. Malinger G, Lev D, Ben-Sira L, et al. **Can syndromic macrocephaly be diagnosed in utero?** *Ultrasound Obstet Gynecol* 2011;37:72–81 CrossRef Medline

11. Gafner M, Fried S, Gosher N, et al. **Fetal brain biometry: is there an agreement among ultrasound, MRI and the measurements at birth?** *Eur J Radiol* 2020;133:109369 CrossRef Medline

12. Barzilay E, Bar-Yosef O, Dorembus S, et al. **Fetal brain anomalies associated with ventriculomegaly or asymmetry: an MRI-based study.** *AJNR Am J Neuroradiol* 2017;38:371–75 CrossRef Medline

13. Ber R, Hoffman D, Hoffman C, et al. **Volume of structures in the fetal brain measured with a new semiautomated method.** *AJNR Am J Neuroradiol* 2017;38:2193–98 CrossRef Medline

14. Ber R, Bar-Yosef O, Hoffmann C, et al. **Normal fetal posterior fossa in MR imaging: new biometric data and possible clinical significance.** *AJNR Am J Neuroradiol* 2015;36:795–802 CrossRef Medline

15. Gafner M, Yagel I, Fried S, et al. **Fetal brain biometry in isolated mega cisterna magna: MRI and US study.** *J Matern Neonatal Med* 2020;1–9 CrossRef Medline

16. Polat A, Barlow S, Ber R, et al. **Volumetric MRI study of the intrauterine growth restriction fetal brain.** *Eur Radiol* 2017;27:2110–18 CrossRef Medline

17. Olshaker H, Ber R, Hoffman D, et al. **Volumetric brain MRI study in fetuses with congenital heart disease.** *AJNR Am J Neuroradiol* 2018;39:1164–69 CrossRef Medline

18. Tilea B, Alberti C, Adamsbaum C, et al. **Cerebral biometry in fetal magnetic resonance imaging: new reference data.** *Ultrasound Obstet Gynecol* 2009;33:173–81 CrossRef Medline

19. Sparrow SS, Cicchetti DV. Balla DA. The Vineland Adaptive Behavior Scales, 2nd ed. In: *Newmark CS. Major Psychological Assessment Instruments.* Allyn & Bacon 2005

20. Chatham CH, Taylor KI, Charman T, et al. **Adaptive behavior in autism: minimal clinically important differences on the Vineland-II.** *Autism Res* 2018;11:270–83 CrossRef Medline

21. Sparrow SS, Cicchetti DV. **Diagnostic uses of the Vineland adaptive behavior scales.** *J Pediatr Psychol* 1985;10:215–25 CrossRef Medline

22. Grinberg A, Katorza E, Hoffman D, et al. **Volumetric MRI study of the brain in fetuses with intrauterine cytomegalovirus infection and its correlation to neurodevelopmental outcome.** *AJNR Am J Neuroradiol* 2019;40:353–58 CrossRef Medline

23. Cicchetti DV. **Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology.** *Psychol Assess* 1994;6:284–90 CrossRef

24. DeMyer W. **Megalencephaly: types, clinical syndromes, and management.** *Pediatr Neurol* 1986;2:321–28 CrossRef Medline

25. Muenchberger H, Assaad N, Joy P, et al. **Idiopathic macrocephaly in the infant: Long-term neurological and neuropsychological outcome.** *Childs Nerv Syst* 2006;22:1242–48 CrossRef Medline

26. Lorber J, Priestley BL. **Children with large heads: a practical approach to diagnosis in 557 children, with special reference to 109 children with megalencephaly.** *Dev Med Child Neurol* 1981;23:494–504 CrossRef Medline

27. Gooskens RH, Willemse J, Faber JA, et al. **Macrocephalies: a differentiated approach.** *Neuropediatrics* 1989;20:164–69 CrossRef Medline

28. Berger I. **Prenatal microcephaly: can we be more accurate?** *J Child Neurol* 2009;24:97–100 CrossRef Medline

29. Gelber SE, Grünebaum A, Chervenak FA. **Prenatal screening for microcephaly: an update after three decades.** *J Perinat Med* 2017;45:167–70 CrossRef Medline

30. Kurtz AB, Wapner RJ, Rubin CS, et al. **Ultrasound criteria for in utero diagnosis of microcephaly.** *J Clin Ultrasound* 1980;8:11–16 CrossRef Medline

31. De Bie I, Boucoiran I. **No. 380: investigation and management of prenatally identified microcephaly.** *J Obstet Gynaecol Canada* 2019;41:855–61 CrossRef Medline

32. Deloison B, Chalouhi GE, Bernard JP, et al. **Outcomes of fetuses with small head circumference on second-trimester ultrasonography.** *Prenat Diagn* 2012;32:869–74 CrossRef Medline

33. Malinger G, Lerman-Sagie T, Watemberg N, et al. **A normal second-trimester ultrasound does not exclude intracranial structural pathology.** *Ultrasound Obstet Gynecol* 2002;20:51–56 CrossRef Medline

34. Gat I, Hoffmann C, Shashar D, et al. **Fetal brain MRI: novel classification and contribution to sonography.** *Ultraschall Med* 2016;37:176–84 CrossRef Medline

35. Salomon-Estebanez M, Mohamed Z, Michaelidou M, et al. **Vineland adaptive behavior scales to identify neurodevelopmental problems in children with congenital hyperinsulinism (CHI).** *Orphanet J Rare Dis* 2017;12:96 CrossRef Medline

Noncontrast MRI Protocol for Selected Pediatric Pituitary Endocrinopathies: A Procedure with High Diagnostic Yield and Potential to Reduce Anesthesia and Gadolinium-Based Contrast Exposure

J. Huang, A. Sarma, N. Gupta, S. Little, and S. Pruthi

ABSTRACT

BACKGROUND AND PURPOSE: Although many pediatric neuroradiology practices empirically use noncontrast brain and pituitary MR imaging for evaluation of growth hormone deficiency, central precocious puberty, and short stature, there are currently insufficient published data to support this practice in an evidence-based fashion. Therefore, the use of contrast-enhanced MR imaging for all pediatric pituitary endocrinopathies remains widespread. We evaluated whether noncontrast MR imaging has adequate diagnostic yield for the evaluation of pediatric growth hormone deficiency, central precocious puberty, and short stature.

MATERIALS AND METHODS: Pituitary MR imaging studies obtained for growth hormone deficiency, central precocious puberty, or short stature in patients 0–18 years of age from 2010 to 2019 were analyzed. Separate blinded review of noncontrast images in cases with abnormalities on the original radiology report was performed by 2 subspecialty-trained pediatric neuroradiologists, with discrepancies resolved by consensus.

RESULTS: Of the 134/442 MR imaging studies obtained for growth hormone deficiency, central precocious puberty, or short stature with hypothalamic-pituitary region abnormalities, there was 70% concordance with the original reports on blinded review of noncontrast images. Twenty-two of 40 discrepancies were deemed unrelated to the indication, and 9 cases originally interpreted as possible microadenoma were read as having normal findings on blinded review. Only 9 of 40 discrepancies required contrast for further characterization.

CONCLUSIONS: In our study, most relevant radiologic findings in patients with growth hormone deficiency, central precocious puberty, and short stature were detectable without contrast, providing evidence that contrast can be avoided in routine MR imaging evaluation of these indications. We propose a "rapid noncontrast pituitary" MR imaging protocol for pediatric patients presenting with growth hormone deficiency, central precocious puberty, or short stature, which may increase efficiency and decrease contrast and anesthesia exposure.

ABBREVIATIONS: CPP = central precocious puberty; GBCA = gadolinium-based contrast agent; GHD = growth hormone deficiency; mDixon = multipoint Dixon; SS = short stature

Currently, there is substantial variability among radiology practices regarding routine intravenous gadolinium-based contrast agent (GBCA) administration for pediatric pituitary MR imaging for common endocrinopathies such as short stature (SS), growth hormone deficiency (GHD), and central precocious puberty (CPP). The clinical significance of retained gadolinium in the brain

is unknown. Therefore, the use of gadolinium-based contrast must be motivated by a compelling contribution to patient care.

Many subspecialty pediatric neuroradiology practices perform noncontrast brain and pituitary MR imaging for GHD, SS, and CPP because many experienced practitioners have intuited that the common causal pathologies are generally evident on noncontrast imaging. Our experience also supports the assertion that the most common anatomic etiologies for GHD, SS, and CPP can be detected without contrast-enhanced sequences in most cases.[1,2] However, this practice has not been uniformly adopted because this experience-based insight has not been tested in an evidence-based fashion or disseminated to the radiology and endocrinology communities at large in a published format.

To our knowledge, only 1 recently published study in the endocrinology literature has directly compared the diagnostic yield

Table 1: Baseline characteristics

Clinical Variable	Total (*n* = 442)	Normal (*n* = 308)	Abnormal (*n* = 134)	*P* Value[a]
Age (yr)	8.2	8.5	7.5	.003
Sex (No.) (%)				.49
Male	233 (53)	159 (52)	74 (55)	
Female	209 (47)	149 (48)	60 (45)	
Clinical indication (No.) (%)				.44
GHD	216 (49)	145 (47)	71 (53)	
CPP	154 (35)	113 (37)	41 (31)	
SS	72 (16)	50 (16)	22 (16)	
Sedation (No.) (%)	225 (51)	147 (48)	78 (58)	.04
Duration of study (hr:min:sec)	0:56:19	0:56:57	0:54:52	.20

[a] The *P* value for quantitative variables was calculated using a *t* test of unequal variance. The *P* value for categoric variables was calculated using a χ^2 test.

of noncontrast-versus-postcontrast MR imaging for the clinical indications of GHD and SS.[3] The only pediatric-specific neuroendocrine imaging scenario described in the most recently published American College of Radiology Appropriateness Criteria is CPP, for which MR imaging of the sella without and with intravenous contrast is listed as the first recommended examination (designated "usually appropriate").[4] Given the relatively limited available published data and guidance on this subject, many radiology practices, therefore, currently follow a "one-size-fits-all" approach to pediatric pituitary imaging, evaluating all neuroendocrine indications with contrast-enhanced brain and pituitary MR imaging, including dynamic contrast-enhanced imaging of the pituitary gland.

The goal of this retrospective study was to evaluate the diagnostic yield of noncontrast-only MR imaging obtained for evaluation of GHD, SS, and CPP in the pediatric population, to determine whether this practice can be adopted as the standard of care. Furthermore, we propose a "rapid noncontrast pituitary" protocol. The ultimate goal was to determine whether this protocol can meet the important objectives of decreasing scan time and potential anesthesia and GBCA exposure while maintaining high test accuracy.

MATERIALS AND METHODS
Data Collection and Analysis
This study was approved by the institutional review board. A search query was performed for all first-time dedicated pituitary MR imaging studies performed in pediatric patients 0–18 years of age from 2010 to 2019 obtained for the clinical indications of GHD, CPP, or SS via retrospective chart review. Exclusion criteria included studies in which contrast was not administered, patients with pre-existing known pituitary abnormalities, patients for whom the MR imaging was ordered for a different clinical indication, and patients who had a pre-existing diagnosis of malignancy.

Separate blinded review of MR examinations with hypothalamic-pituitary region findings on the original radiology report was performed by 2 subspecialty-trained pediatric neuroradiologists. Both pediatric neuroradiologists are subspecialty-trained, with one of them having spent 14 years in clinical practice, and the other, approximately 5 years. Only noncontrast images were reviewed in a blinded fashion, even if contrast-enhanced images were obtained as

part of the original examination. In 37 of 134 cases, there was interobserver variability, which was resolved by consensus.

Statistical Analysis
Statistical analysis was performed using R statistical and computing software 3.4.0 (http://www.r-project.org/).[5] Categorical variables were analyzed using the χ^2 test, and quantitative variables were analyzed using the *t* test of unequal variance.

Protocol
At our institution, the MR imaging brain and pituitary protocol is performed with 1.5 or 3T and includes the following: sagittal T1WI, axial spin-echo T2WI and FLAIR, sagittal spin-echo T2WI, axial DWI, and thin-section sagittal and coronal T1WI focused on the sella (both with a section thickness of 3 mm, no skip). Voxel size is 0.7×0.9 mm, and matrix size is 196×155. In addition, postcontrast sequences are obtained, including axial T1WI of the brain and thin-section sagittal and coronal T1WI focused on the sella (both with a section thickness of 3 mm, no skip). Throughout the 10-year study period, there were minor variations in the protocol depending on the study date, which include sequences such as dynamic contrast-enhanced images through the sella and sagittal T2WI. Our dynamic contrast sequence includes a section thickness of 3 mm with no skip and 10 phases (11.6 seconds per phase). Voxel size is 0.8×1.15 mm, and matrix size is 224×154. Sedation was coordinated with the anesthesiology team as needed.

RESULTS
Demographics
Four hundred forty-two pediatric pituitary MR imaging studies were obtained within the specified time period. For these studies, GHD was the most common clinical indication, accounting for 49%, compared with 35% for CPP and 16% for SS (Table 1). The mean age was 8.2 years (range, 9 months to 17 years). Two hundred twenty-five (51%) studies required sedation. The average scan duration was 56 minutes, 19 seconds (range, 22 minutes, 45 seconds to 1 hour, 59 minutes, 37 seconds).

Radiologic Findings within the Hypothalamic-Pituitary Region
Three hundred eight of 442 studies (69%) demonstrated a normal hypothalamic-pituitary region according to the original MR imaging report. Abnormalities were seen in the hypothalamic-pituitary region in 134 of 442 (30%) studies and outside the hypothalamic-pituitary region in 146/442 (33%) studies. Of the 134 of 442 studies with hypothalamic-pituitary region abnormalities on the original report, 95% (*n* = 127/134) demonstrated pituitary or infundibular pathology and 5% (*n* = 7/134) had hypothalamic region pathology.

The 3 most common radiologic findings on the studies with abnormal findings were pars intermedia or Rathke cleft cyst

| MR Imaging Findings	Studies (No.) (%)
Pars intermedia or Rathke cleft cyst	38 (28)
Subjectively small anterior pituitary	36 (27)
Ectopic or absent posterior pituitary	20 (15)
Empty sella	15 (11)
Postpubertal pituitary	14 (10)
Pituitary stalk interruption	12 (9)
Possible microadenoma	9 (7)
Suprasellar mass or cystic lesion	6 (4)
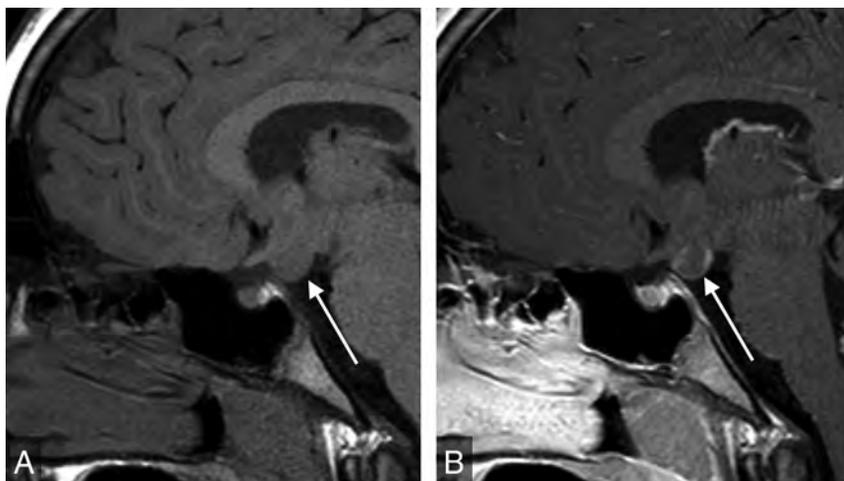
FIG 1. Sagittal T1WI of the sella (A) demonstrates a gray matter–isointense suprasellar mass lesion (arrow). Sagittal postgadolinium T1WI (B) shows enhancement, indicating a diagnosis of neoplasm (hypothalamic glioma) rather than hamartoma.

(28%, $n = 38$), a subjectively small anterior pituitary gland (27%, $n = 36$), and an ectopic or absent posterior pituitary gland (15%, $n = 20$) (Table 2). A small anterior pituitary gland was determined via subjective assessment of the size/height of the anterior pituitary gland, taking known demographic characteristics (patient age and sex) into account. Microadenoma was identified or suspected in 9 cases (7%). Therapeutic decisions were unaffected by the presence of a suspected microadenoma in all cases. The 7 cases of hypothalamic region pathology included 2 cases of hypothalamic hamartoma, 2 cases of suprasellar arachnoid cyst, 1 case of hypothalamic glioma (Fig 1), 1 case of lipoma of the tuber cinereum, and 1 case of mass effect on the hypothalamus due to extra-axial fluid collections.

Concordance

Two subspecialty-trained pediatric neuroradiologists performed separate blinded reviews of only the noncontrast images on the 134 studies with abnormal findings. On blinded review, there was 70% concordance ($n = 94$) with the original MR imaging reports. Of the 40 discrepancies, 58% ($n = 22$) were considered unlikely to produce symptoms of GHD, CPP, or SS or of any clinical significance (eg, pars intermedia cyst). Of the remaining 18 discrepancies, 9 cases were interpreted as having microadenoma or possible microadenoma on the original report and read as having

normal findings on the blinded review of noncontrast images only.

The remaining 9 of 40 discrepancies (23%) required contrast for definitive characterization. Six were extrasellar masses, which included 2 suprasellar arachnoid cysts, 2 hypothalamic hamartomas, 1 hypothalamic glioma, and 1 case of an atypical mass along the pituitary stalk. The 3 cases of stalk abnormalities that required contrast for complete characterization were a case of duplicated stalk and 2 cases of a thickened, enhancing pituitary stalk. Figure 2 summarizes the discrepant findings.

Findings Outside the Hypothalamic-Pituitary Region

Abnormalities outside the hypothalamic-pituitary region were detected in 146 (33%) studies. Most of these were incidental findings, such as nonspecific T2 or FLAIR white matter hyperintensities (10%, $n = 45$) and sinus disease (9%, $n = 39$) (Table 3). Chiari I malformation or cerebellar tonsillar ectopia was detected in 18 cases (4%). Intracranial cysts were found in 4% of studies ($n = 18$); this category includes pineal cysts ($n = 7$), posterior fossa or temporal lobe arachnoid cysts ($n = 6$), and choroid fissure cysts ($n = 2$). Congenital anomalies were noted in 14 cases (3%), including 5 cases with corpus callosum abnormalities and 5 cases with abnormal skull base morphology.

Six studies demonstrated abnormalities for which the differential diagnosis included neoplasm. There were 5 cases of presumed low-grade glioma and 1 of a suspected dysembryoblastic neuroepithelial tumor (Fig 3). Two of these did not have a follow-up study available, while the other 4 were unchanged on follow-up contrast MR imaging. In 2 studies performed for GHD, the suspected mass delayed growth hormone therapy until stability was confirmed on follow-up MR imaging. Two studies were performed for CPP, and the presence of a possible neoplasm did not prevent treatment. The findings in the 2 remaining cases also did not affect management decisions. The patient with suspected dysembryoplastic neuroepithelial tumor did not undergo follow-up imaging or subspecialty referral to neurosurgery due to family preference.

Two patients had cavernous malformations, which were unchanged on follow-up imaging. One patient was incidentally noted to have a cerebral aneurysm, for which no follow-up imaging was available.

DISCUSSION

Neuroendocrine abnormalities include centrally mediated hormonal imbalances and organ-specific pituitary abnormalities. GHD, SS, and CPP are among the most common clinical indications for brain and pituitary MR imaging in our pediatric population.

Concordance using Noncontrast Only Blinded Review of Pituitary MRIs

Legend: ☐ Concordant ☐ Discrepant ■ Contrast Needed

N	Diagnosis	Blind Review without Contrast	Clinical Significance
9	Microadenoma or suspected microadenoma	Normal	None
2	Suprasellar arachnoid cyst	Cystic lesion	Follow up stable
2	Hypothalamic hamartoma	Hypothalamic mass	Follow up stable
2	Borderline thickened and enhancing stalk	Normal	1 case with no follow up; 1 case with normal variant on follow up
1	Hypothalamic glioma	Hypothalamic mass	Follow up stable
1	Atypical mass along stalk	Mass along stalk	Follow up stable
1	Partially duplicated stalk, pituitary aplasia, ectopic PP	Pituitary aplasia, ectopic PP	No follow up

N	Initial Report	Blind Review without Contrast	Clinical Significance
10	Pars intermedia cyst	Normal	None
4	Postpubertal pituitary	Normal	None
3	Small pituitary	Normal	None
3	Heterogeneous pituitary enhancement	Normal	None
2	Partial empty sella	Normal	None

FIG 2. The *black table* (upper right, $n = 18$) indicates discrepant cases between the retrospective noncontrast MR imaging evaluation and the original report that would have required contrast for definitive characterization. Nine of these were diagnosed as microadenoma or suspected microadenoma on original contrast-enhanced MR imaging. The *white table* (lower right, $n = 22$) describes discrepant cases with diagnoses that were deemed unlikely to cause the patient's presenting symptoms (probably clinically insignificant discrepancies). PP indicates posterior pituitary.

Table 3: Additional findings in pituitary MR imaging studies obtained for evaluation of GHD, CPP, or SS

Additional MR Imaging Finding ($n = 146$, 33%)	Studies (No.) (%)
Nonspecific T2 or FLAIR hyperintensities	45 (10)
Sinus disease	39 (9)
Chiari I malformation or cerebellar tonsillar ectopia	18 (4)
Intracranial cysts	18 (4)
Pineal cyst	7
Posterior fossa/temporal lobe arachnoid cyst	6
Choroid fissure cyst	2
Congenital anomaly	14 (3)
Dysmorphic skull base	5
Corpus callosum abnormality	5
Vascular anomaly	7 (2)
Cavernous malformation	2
Cerebral aneurysm	1
Suspected mass	6 (1)
Presumed low-grade glioma or neoplasm	5
Dysembryoblastic neuroepithelial tumor	1
Migrational abnormality	6 (1)

There are many etiologies for SS, which, by definition, occurs in 2.5% of children. GHD accounts for approximately 1%–2% of cases of SS.[6] Most cases of GHD are idiopathic; however, due to the potential for underlying structural abnormalities affecting the hypothalamic-pituitary region, radiologic evaluation is often indicated.[7] Although most patients with GHD have normal MR imaging findings, possible abnormalities include pituitary or in-fundibular hypoplasia or aplasia, ectopic or absent posterior pituitary bright spot, or, rarely, neoplasms such as craniopharyngioma.[8,9] In addition, exclusion of brain tumors that may be exacerbated by growth hormone therapy (eg, glioma) may be an important clinical consideration.[10-12]

The prevalence of CPP is variable geographically, with a female/male ratio of 20:1.[13] A large proportion of cases of CPP are idiopathic, particularly in females. However, in the minority of cases with abnormal brain MR imaging findings, etiologies may include hypothalamic hamartoma or hypothalamic-pituitary neoplasms.[14,15] Thus, radiologic evaluation is often suggested to assess these pathologies.[16]

Our results reveal that most MR imaging findings in GHD, CPP, and SS can be diagnosed on noncontrast sequences. This result confirms the findings of Almaghraby et al,[3] who also concluded that contrast may not be needed in most MR imaging studies performed for GHD and SS. Our study adds CPP to the list of clinical indications for which contrast may not be necessary. Further support for this conclusion includes multiple previous studies verifying that the most common radiologic findings are not likely the etiology of the patient's clinical presentation. For example, pars intermedia or Rathke cleft cysts were seen in 28% of cases in our study and have rarely been implicated in GHD, CPP, or SS.[17] Another relatively common MR imaging finding is a partially empty sella, seen in 11% of cases in our study, which is most frequently of little-to-no clinical significance except in the setting of idiopathic intracranial hypertension.[18] Data such as these may be helpful in the formulation of further widely used evidence-based guidelines such as the American College of Radiology Appropriateness Criteria.
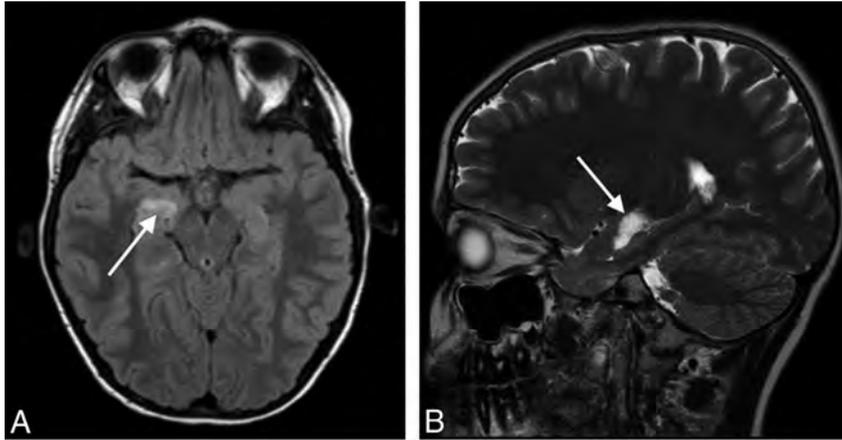
FIG 3. Axial FLAIR *(A)* and sagittal T2WI *(B)* show a hyperintense mass lesion *(arrows)* in the right mesial temporal lobe immediately superior to the hippocampus, most likely representing a low-grade neoplasm. Sagittal T2WI is helpful for localization of this lesion and differentiating it from the choroid plexus in the temporal horn. Although this finding was unrelated to the patient's presentation of GHD, intravenous contrast was needed for complete MR imaging assessment and showed the lesion to be nonenhancing (not shown).
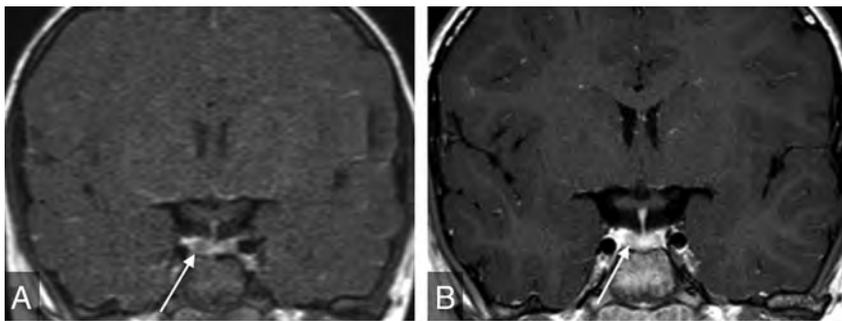

FIG 4. Coronal dynamic postgadolinium *(A)* and delayed *(B)* T1WIs of the sella show heterogeneous enhancement, with hypoenhancement of the right side of the gland *(arrows)*. This finding was considered equivocal for microadenoma, and the patient underwent multiple follow-up examinations despite the finding being considered clinically unrelated to GHD. The finding led to no clinical intervention.

In our study, there were 18 cases in which contrast-enhanced imaging revealed additional findings (eg, microadenoma) or further characterized abnormalities of the hypothalamic-pituitary region that were suspected on noncontrast sequences. In 9 of these cases, suspected microadenomas were identified. In line with previous studies, these were considered noncontributory to the reason for referral and, therefore, clinically insignificant.[19,20] In 7 of these cases, abnormalities were detected on noncontrast-only images, but contrast was required for further characterization.

Microadenomas are a common-but-controversial justification for contrast administration in pediatric pituitary MR imaging and are often cited as a reason to include sequences such as dynamic contrast-enhanced images.[21] They are seldom clinically relevant without signs and symptoms of hormonal hypersecretion (eg, galactorrhea, hyperprolactinemia). In addition, there are multiple pitfalls associated with imaging diagnosis. First, normally occurring hypoenhancing regions within a heterogeneously enhancing normal pituitary gland may simulate microadenoma

(Fig 4). Second, even when a well-delineated lesion is present, functional adenoma cannot be accurately differentiated from nonfunctional adenoma by MR imaging. This finding thus may beget additional lab testing, potentially adding expense, time, and stress to the patient/family experience. Third, if this testing confirms a functional microadenoma, there is still no definitive evidence that the microadenoma is the etiology of the patient's presenting symptoms. For example, the link between microadenoma and CPP remains controversial, and no relationship between follicle-stimulating hormone–secreting adenoma and CPP has been reported.[22] Ultimately, in the setting of GHD, SS, and CPP, the possible microadenoma is usually an incidentaloma.

Thirty-three percent of the MR imaging examinations in our study demonstrated abnormalities outside the hypothalamic-pituitary region, with the large majority not altering patient management (therefore considered incidental). In our study, there were 7 additional findings that were potentially clinically significant, which included 6 suspected low-grade neoplasms and 1 case of cerebral aneurysm. These results are in agreement with those of previous studies confirming that incidental findings on pediatric brain MR imaging studies are common, however rarely clinically significant.[23-25] Many additional findings that may be worthy of follow-up or additional subspecialty referral, such as Chiari I malformation, are detectable by noncontrast examination.

On the basis of the results of our study, we advocate for use of a rapid noncontrast pituitary MR imaging protocol for pediatric patients presenting with GHD, CPP, or SS. This protocol includes coronal and sagittal TSE T1-weighted multipoint Dixon (mDixon) and sagittal T2-weighted thin-section images of the sella, axial T2WI or FLAIR of the brain, and axial DWI (Online Supplemental Data). Depending on institutional preference, thin-section coronal T2WI for evaluation of the optic and olfactory apparatus and hypothalamus may be substituted for or performed in addition to sagittal T2WI, especially in patients with suspected septo-optic dysplasia or Kallman syndrome. The choice of the acquisition technique, such as the section thickness and 2D-versus-3D acquisition, will depend on institutional preference (taking into account the desired length of the rapid examination). When available, scanning should be performed on higher-field-strength magnets (eg, 3T) because increased spatial and image
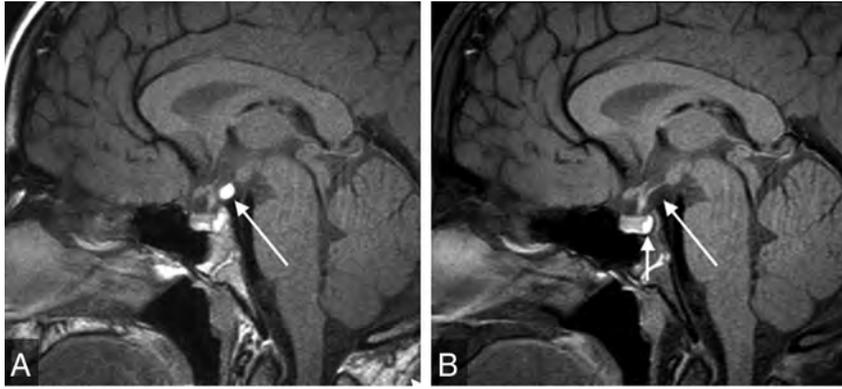
FIG 5. Sagittal mDixon T1WI *(A)* shows a hyperintense suprasellar mass along the posterior aspect of the infundibulum *(long arrow)*. Sagittal mDixon T1-weighted water-only image *(B)* shows complete suppression of this mass, which is diagnostic of lipoma. The neurohypophysis is bright on the water-only images and is highlighted against suppressed fat in the dorsum sellae *(short arrow, B)*. mDixon T1WI is helpful for evaluation of fat-containing sellar region lesions and identification of the normal orthotopic neurohypophysis. Furthermore, it can differentiate a fat-containing lesion from ectopic neurohypophysis.
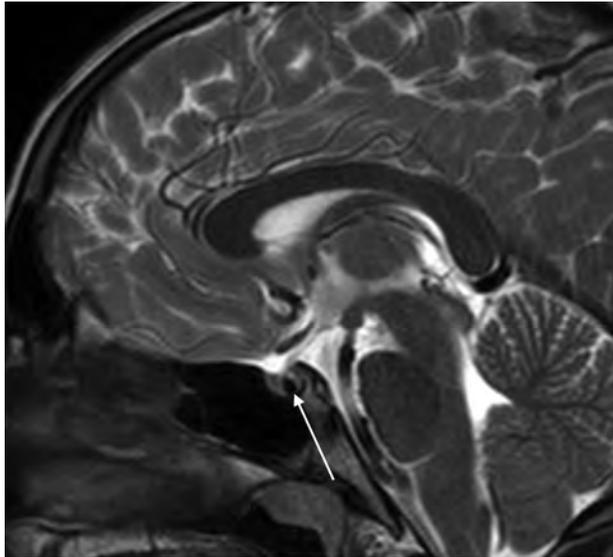

FIG 6. Sagittal T2WI shows a hypointense pars intermedia cyst *(arrow)* that was not definitely visible on sagittal T1WI (not shown). This case highlights this sequence being complementary to T1WI and helpful for identification and characterization of sellar cystic lesions. It is suggested as part of the rapid noncontrast pituitary protocol because it further decreases the needed for intravenous contrast.

Thin-section sagittal T2WI is helpful for characterizing sellar cystic lesions, which often have variable signal on T1- and T2WI (Fig 6). Axial T2WI and/or axial FLAIR (depending on institutional preference) and DWI can be helpful in detecting concurrent brain abnormalities such as signal abnormality concerning for neoplasm, which is a relative contraindication to administration of growth hormone, and cortical malformations that are common in patients with conditions involving the hypothalamic-pituitary axis such as septo-optic dysplasia.[10-12]

Our rapid noncontrast pituitary MR imaging protocol would decrease the total acquisition time, therefore providing a substantial decrease in the duration of anesthesia as well as GBCA administration. In addition, the shorter imaging session length could enable increased patient throughput and entirely eliminate the need for procedural sedation in a subset of patients (acknowledging the contribution of other anesthesia-mitigating factors such as the work of Child Life Specialists). We acknowledge that the discovery of unsuspected lesions that require further evaluation with contrast may inconvenience the small subset of patients who would have to return for an additional contrast-enhanced examination (with or without sedation). However, this would be substantially offset by a much larger pool of patients for whom the noncontrast rapid protocol would provide greater convenience through savings in time, cost, anesthesia, and contrast administration.

The implication that a rapid noncontrast pediatric pituitary MR imaging protocol could decrease scan duration, anesthesia exposure, and exposure to GBCA is particularly important given the current uncertainty about long-term health effects of intravenous GBCA and imaging-related procedural sedation in the pediatric population.[26,27]

There are some limitations to our study. First, these data were collected from a single institution, and our findings may not be representative of those in other regions. Second, there were small variations in MR imaging techniques across the study timeframe. Third, this study evaluates only pituitary MR imaging studies obtained for GHD, CPP, and SS and cannot be generalized to patients presenting with other clinical syndromes or disorders, including diabetes insipidus or excess pituitary hormone production. This study also does not capture results in patients with possible pituitary abnormalities for whom MR imaging studies other than MR imaging of the head and pituitary gland were performed (eg, MR imaging of the brain or MR imaging of the brain and orbits). For example, this issue may explain why our series did not include a patient with craniopharyngioma, which would be an expected pathology in a pediatric cohort including patients with GHD and SS.

contrast resolution may enhance diagnostic performance. However, it may not be feasible for all scans to be performed on 3T systems due to scheduling and other logistic factors. Coronal and sagittal T1WI of the sella is capable of depicting most relevant anatomic abnormalities, including small anterior pituitary and absent or ectopic posterior pituitary glands. In addition, other midline abnormalities such as callosal malformation are readily detectable. At centers capable of performing mDixon sequences, we propose acquiring mDixon pituitary images because the fat-suppressed images are helpful for differentiating the posterior pituitary gland from clival marrow fat, differentiating ectopic posterior pituitary from lipoma, and evaluating extra-cranial pathology (Fig 5).

CONCLUSIONS

Although many pediatric subspecialty practices have already adopted noncontrast brain and pituitary MR imaging for SS, GHD, and CPP, this practice has not yet been uniformly adopted worldwide. An unnecessary contrast-enhanced technique may, hence, lead to excessive GBCA administration, scan length, and procedural sedation. One reason why some may be reluctant to adopt this technique is a paucity of currently available supportive published data. In our study, most relevant radiologic findings could have been detected without GBCA administration. Our study, therefore, provides evidence that GBCAs are not necessary for routine MR imaging evaluation for these indications. Moreover, we propose the use of rapid noncontrast pituitary MR imaging. This practice has the potential to further decrease GBCAs and anesthesia exposure. Future work will focus on prospective evaluation of savings in time, cost, and exposure to GBCAs and anesthesia with this protocol.

REFERENCES

1. Chaudhary V, Bano S. **Imaging of pediatric pituitary endocrinopathies.** *Indian J Endocrinol Metab* 2012;16:682–91 CrossRef Medline

2. Melmed S. **Pituitary-tumor endocrinopathies.** *N Engl J Med* 2020;382:937–50 CrossRef Medline

3. Almaghraby A, Jaju A, Ryan ME, et al. **SAT-LB19 is there a need to use gadolinium contrast for pituitary MRI in the evaluation of pediatric short stature and growth hormone deficiency?** *J Endocr Soc* 2020;4(Suppl 1):SAT-LB19 CrossRef

4. Burns J, Policeni B, Bykowski J, et al. Expert Panel on Neurologic Imaging. **ACR Appropriateness Criteria® Neuroendocrine Imaging.** *J Am Coll Radio l* 2019;16:S161–73 CrossRef Medline

5. R Core Team. *R: A Language Environment for Statistical Computing.* R Foundation for Statistical Computing; 2017

6. Stanley T. **Diagnosis of growth hormone deficiency in childhood.** *Curr Opin Endocrinol Diabetes Obes* 2012;19:47–52 CrossRef Medline

7. Growth Hormone Research Society. **Consensus guidelines for the diagnosis and treatment of growth hormone (GH) deficiency in childhood and adolescence: GH Research Society.** *J Clin Endocrinol Metab* 2000;85:3990–93 CrossRef Medline

8. AlJurayyan RN, AlJurayyan NA, Omer HG, et al. **Pituitary imaging in 129 children with growth hormone deficiency: a spectrum of findings.** *Sudan J Paediatr* 2017;17:30–05 Medline

9. Yesquen P, Clemente M, Campos A, et al. **Brain magnetic resonance imaging in children with isolated growth hormone deficiency and idiopathic short stature diagnoses.** *Horm Res Paediatr* 2019;91 (Suppl 1):437

10. Lea RW, Dawson T, Martinez-Moreno CG, et al. **Growth hormone and cancer: GH production and action in glioma?** *Gen Comp Endocrinol* 2015;220:119–23 CrossRef Medline

11. Clayton PE, Gattamaneni HR, Shalet SM, et al. **Does growth hormone cause relapse of brain tumours?** *Lancet* 1987;329:711–13 CrossRef Medline

12. Raman S, Grimberg A, Waguespack SG, et al. **Risk of neoplasia in pediatric patients receiving growth hormone therapy: a report from the Pediatric Endocrine Society Drug and Therapeutics Committee.** *J Clin Endocrinol Metab* 2015;100:2192–203 CrossRef Medline

13. Partsch CJ, Sippell WG. **Treatment of central precocious puberty.** *Best Pract Res Clin Endocrinol Metab* 2002;16:165–89 CrossRef Medline

14. Cantas-Orsdemir S, Garb JL, Allen HF. **Prevalence of cranial MRI findings in girls with central precocious puberty: a systematic review and meta-analysis.** *J Pediatr Endocrinol Metab* 2018;31:701–10 CrossRef Medline

15. Mogensen SS, Aksglaede L, Mouritsen A, et al. **Pathological and incidental findings on brain MRI in a single-center study of 229 consecutive girls with early or precocious puberty.** *PLoS One* 2012;7: e29829 CrossRef Medline

16. Ng SM, Kumar Y, Cody D, et al. **Cranial MRI scans are indicated in all girls with central precocious puberty.** *Arch Dis Child* 2003;88:414–18 CrossRef Medline

17. Lim HH, Yang SW. **Risk factor for pituitary dysfunction in children and adolescents with Rathke's cleft cysts.** *Korean J Pediatr* 2010;53:759–65 CrossRef Medline

18. Maghnie M, Lindberg A, Koltowska-Häggström M, et al. **Magnetic resonance imaging of CNS in 15,043 children with GH deficiency in KIGS (Pfizer International Growth Database).** *Eur J Endocrinol* 2013;168:211–17 CrossRef Medline

19. Chiu CF, Wang CJ, Chen YP, et al. **Pathological and incidental findings in 403 Taiwanese girls with central precocious puberty at initial diagnosis.** *Front Endocrinol (Lausanne)* 2020;11:256 CrossRef Medline

20. Hirsch W, Zumkeller W, Teichler H, et al. **Microadenomas of the pituitary gland in children with and without hypophyseal dysfunction in magnetic resonance imaging.** *J Pediatr Endocrinol Metab* 2002;15:157–62 CrossRef Medline

21. Derrick KM, Gomes WA, Gensure RC. **Incidence and outcomes of pituitary microadenomas in children with short stature/growth hormone deficiency.** *Horm Res Paediatr* 2018;90:151–60 CrossRef Medline

22. Pedicelli S, Alessio P, Scirè G, et al. **Routine screening by brain magnetic resonance imaging is not indicated in every girl with onset of puberty between the ages of 6 and 8 years.** *J Clin Endocrinol Metab* 2014;99:4455–61 CrossRef Medline

23. Dangouloff-Ros V, Roux CJ, Boulouis G, et al. **Incidental brain MRI findings in children: a systematic review and meta-analysis.** *AJNR Am J Neuroradiol* 2019;40:1818–23 CrossRef Medline

24. Jansen P, Dremmen M, Van Den Berg A, et al. **Incidental findings on brain imaging in the general pediatric population.** *N Engl J Med* 2017;377:1593–95 CrossRef Medline

25. Gupta SN, Gupta VS, White AC. **Spectrum of intracranial incidental findings on pediatric brain magnetic resonance imaging: what clinician should know?** *World J Clin Pediatr* 2016;5:262–72 CrossRef Medline

26. Shah R, D'Arco F, Soares B, et al. **Use of gadolinium contrast agents in paediatric population: Donald Rumsfeld meets Hippocrates!** *Br J Radiol* 2019;92:20180746 CrossRef Medline

27. Davidson AJ, Sun LS. **Clinical evidence for any effect of anesthesia on the developing brain.** *Anesthesiol* 2018;128:840–53 CrossRef Medline

Growth Curves of Subependymal Giant Cell Tumors in Tuberous Sclerosis Complex

J.R. Barnett, J.H. Freedman, H. Zheng, E.A. Thiele, and P. Caruso

ABSTRACT

BACKGROUND AND PURPOSE: Growth of subependymal giant cell tumor and subependymal nodules has not been well-characterized. The purpose of this study was to determine whether growth curves can differentiate subependymal giant cell tumors from subependymal nodules.

MATERIALS AND METHODS: Brain MR imaging of patients with tuberous sclerosis complex were retrospectively reviewed from 2002 to 2018. All lesions in the region of the foramen of Monro were measured. Lesions were categorized on the basis of maximal diameter at the most recent scan: small lesions (<1 cm), indeterminate lesions (>1 cm), and resected lesions (>1 cm and surgically resected). Growth velocity and acceleration on serial imaging were analyzed, and growth rates were calculated between 0 and 20 years of age and compared among the 3 categories.

RESULTS: Forty-one patients were analyzed. The average age at the earliest scan was 5.9 (SD = 5.7) years. One hundred twenty-six small, 27 indeterminate, and 10 resected lesions were measured. Subependymal giant cell tumors grew faster than indeterminate lesions between 6 and 15 years of age. Indeterminate lesions grew faster than small lesions at 0–10 years of age. Resected lesions showed increased velocity and acceleration of growth compared with indeterminate lesions and small lesions on serial imaging.

CONCLUSIONS: Growth differentiates subependymal nodules and subependymal giant cell tumors within the first 20 years of life, and the use of velocity and acceleration of growth may refine the diagnostic criteria of subependymal giant cell tumors. Additionally, 6–15 years of age may be an important period to monitor subependymal giant cell tumors at the foramen of Monro because increased growth may help to identify subependymal giant cell tumors that will continue to grow and result in obstructive hydrocephalus.

ABBREVIATIONS: AP = anterior-posterior; FOM = foramen of Monro; IQR = interquartile range; SEGA = subependymal giant cell astrocytoma; SEN = subependymal nodule; SGCT = subependymal giant cell tumor; SI = superior-inferior; TSC = tuberous sclerosis complex; TV = transverse

Tuberous sclerosis complex (TSC) is an autosomal dominant disorder caused by a mutation in the *TSC1-TSC2* complex.[1] Brain lesions of TSC include subependymal nodules (SENs) and subependymal giant cell tumors (SGCTs), also known as subependymal giant cell astrocytoma (SEGA).[2] Approximately 24% of patients with TSC develop SGCTs, and 88%–95% develop SENs.[2,3]

SGCTs are slow-growing World Health Organization grade I tumors that usually grow near the foramen of Monro (FOM) and can obstruct the flow of CSF leading to increased morbidity and mortality.[2,4,5] SGCTs occur in the same locale as SENs; however, SENs typically remain stable and are not associated with a risk of obstructive hydrocephalus.[3] Currently, consensus guidelines suggest that surveillance with MR imaging be performed every 1–3 years in patients with TSC to monitor for the development of SGCTs because they are a cause for obstructive hydrocephalus.[6]

SGCTs can arise from SENs and are histologically indistinguishable.[2,7] Similarly, MR imaging does not provide a reliable biomarker for the distinction of early SGCTs from SENs. SGCTs and SENs are differentiated by maximal diameter and growth.[6] Consensus guidelines define SGCTs as lesions ≥1 cm in maximum diameter at the caudothalamic groove or a lesion at any location that has shown growth on consecutive imaging.[6] Growth patterns of SGCTs and SENs have not been well-characterized in the medical literature.[3] The purpose of our study was to

characterize the growth of SENs and SGCTs in the region of the FOM and to determine whether growth curves can differentiate SGCTs from SENs.

MATERIALS AND METHODS

Standard Protocol Approvals and Inclusion Criteria

This study was approved by the institutional review board. Patient data were retrospectively reviewed and de-identified, and storage was encrypted and password-protected. The need for written consent was waived.

We retrospectively reviewed the clinical and imaging medical records of patients with TSC seen at our institution between 2002 and 2018.

Inclusion criteria were the following: 1) confirmed diagnosis of TSC by consensus criteria[8] or by genetic testing; 2) at least 3 serial MRIs, including 3D Brain Volume Imaging (BRAVO; GE Healthcare), MPRAGE, and spoiled gradient echo; 3) earliest MR imaging obtained before 20 years of age; and 4) radiologic evidence of SGCTs and/or SENs. For all patients who eventually underwent surgical resection for an SGCT, at least 3 preoperative MRIs had to span at least a 1.5- year surveillance period. For the non-SGCT group, at least 3 MRIs had to span at least a 5-year surveillance period; a longer surveillance period was beneficial to ensure that the nodules did not eventually grow into SGCTs. Exclusion criteria were the following: 1) treatment with mechanistic target of rapamycin inhibitors during the surveillance period, 2) surgical resection of an SGCT before the surveillance period, and 3) nondiagnostic MRIs.

Clinical Data

The clinical records were reviewed for age at earliest MR imaging, sex, genetic testing, and age at surgical resection.

Imaging Data

All nodules at the FOM, small lesions, indeterminate lesions, and resected lesions, were measured from all available brain MRIs. A single researcher (J.R.B.) performed all measurements. The senior neuroradiologist (P.C.) with 17 years' postfellowship experience trained a researcher (J.R.B.) on an initial training data set of 20 MRIs. The research assistant then performed all remaining measurements, and for certain cases when they were in question, the measurements were reviewed and confirmed for accuracy by P.C. The senior neuroradiologist measured 10 nodules in 10 patients independently, and the interreader reliability was calculated using the intraclass correlation coefficient. The intraclass correlation coefficients for anterior-posterior (AP), superior-inferior (SI), and transverse (TV) measurements were 0.94, 0.97, and 0.93, respectively.

All measurements were made using the available T1 volumetric sequence for each MR imaging study date with a BRAVO, MPRAGE, or echo-spoiled gradient echo sequence. The T1 volumetric sequence was brought up in a 3D viewer in the institutional PACS, and the axial, coronal, and sagittal reformats were simultaneously displayed. The AP, TV, and SI measurements were performed.

Imaging Parameters

Parameters for the volumetric T1 sequences were as follows: For the echo-spoiled gradient echo sequence ($n = 110$), the parameters were section thickness = 1 mm, gap = 0 mm, TR = 74 ms, TE = 7 ms, phase = 190–220, frequency = 190–220, number of excitations 1 and FOV = 190–220 × 190–220 mm. For the BRAVO sequence ($n = 1$), the parameters were section thickness = 1 mm, gap = 0 mm, TR = 74 ms, TE = 6.508 ms, phase = 260, frequency = 260, number of excitations 1 and FOV = 260 mm. For the MPRAGE sequence ($n = 160$), the parameters were thickness = 1 mm, gap = 0 mm, TR = 2530 ms, TE = 2530 ms, phase = 192–256, frequency = 192–256, number of excitations 1 and FOV = 192–256 × 192–256 mm.

Lesion Categorization

Lesions were retrospectively categorized using the measurements obtained on the most recent MR imaging during the surveillance period. Small lesions were defined as lesions <1 cm in maximum diameter and may be considered terminologically per the consensus guidelines as synonymous with SENs.[6] Indeterminate lesions were defined as lesions >1 cm in maximum diameter but that never required surgical resection during the surveillance period or at any time point according to the medical records; such lesions are traditionally termed SGCTs.[6] Resected lesions were defined as lesions of >1 cm in maximum diameter, which were surgically resected for an indication of obstructive hydrocephalus, and this term may be considered as synonymous with resected SGCTs.[6]

Volume Calculation and Evaluation

Volume was calculated assuming that the lesion was ellipsoid: [Volume = $4/3\pi$ (A/2)(B/2)(C/2)], where A, B, and C are the maximum dimensions of the lesion measured in 3 perpendicular planes. The volumes of the nodules were calculated on every MR imaging measured during the surveillance period.

Growth Evaluation

The available MRIs for each patient were grouped into 4 age periods: 0–5, 6–10, 11–15, and 16–20 years of age. The first and last volume measurements were taken from each age period, and the following equation was applied to determine growth rate: (Volume at Most Recent Scan – Volume at Earliest Scan) / (Age at Most Recent Scan – Age at Earliest Scan) = Growth Velocity (mm³/year). The median growth velocity and interquartile range (IQR) for each age period were calculated. This approach was then subdivided for each type of lesion (small lesions, indeterminate lesions, and resected lesions). Additionally, the percentage of lesions with growth velocity >4.2 mm³/year in a given age period was calculated; 4.2 mm³ was chosen to reflect an increase in growth of 2 mm in the AP, TV, and SI dimensions using an ellipsoid volume formula.

Finally, growth velocity was found between each MR imaging for small lesions, indeterminate lesions, and resected lesions using the equation (Volume at MR Imaging – Volume of Previous MR Imaging) / (Age at MR Imaging – Age at Previous MR Imaging) = Growth Velocity (mm³/Year). Acceleration of growth was then calculated using the formula (Growth Velocity – Previous Growth Velocity) / (Age at MR Imaging – Age at Previous MR Imaging) =

Acceleration of Growth ($mm^3/year^2$). The median and IQR for growth velocity and acceleration were calculated.

Statistical Analysis

Statistical analysis was conducted using GraphPad Prism software, Version 9 (GraphPad Software). All statistical analyses were conducted with an α of .05. Normality was tested by the D'Agostino-Pearson test. Nonparametric tests were used to compare data that were determined to have a non-normal distribution, and for this data medians with IQRs were reported.

A 2-tailed Mann-Whitney U test was used to evaluate volumes of resected lesions and indeterminate lesions at baseline and the most recent scan. A Kruskal-Wallis test was used to compare growth velocity of all lesions, small lesions, indeterminate lesions, and resected lesions at 0–5, 6–10, 11–15, and 16–20 years of age. Additionally, the following comparisons were made within each age period using a 2-tailed Mann-Whitney U test with a Bonferroni correction: 1) resected lesions versus indeterminate lesions, 2) resected lesions versus small lesions, 3) small lesions versus indeterminate lesions, and 4) small lesions and indeterminate lesions versus resected lesions. For the 0–5, 6–10, and 11–15 years of age periods, 4 comparisons were made. Each P value calculated within these periods was multiplied by 4. A comparison of growth velocity and acceleration on serial imaging was made using a 2-tailed Mann-Whitney U test with a Bonferroni correction to compare all lesion groups. Three comparisons were made for acceleration and velocity. Each P value was multiplied by 3. Finally, a mixed-effects analysis was performed to analyze the interactions between lesion type and age period (0–5, 6–10, 11–15, and 16–20 years of age) on growth velocity. Growth velocity was defined as the dependent variable, and lesion type and age period (0–5, 6–10, 11–15, and 16–20 years) were independent variables with an interaction term between them. Within-subject repeated measurements were accounted for using subject-level random effects.

RESULTS

Clinical Data

Forty-one participants who met the inclusion and exclusion criteria were analyzed. The average age of patients at the earliest MR imaging was 5.9 years of age (range, 0.3–19.2 years of age). Twenty-six patients were female, and 15 were male. Twenty had a *TSC2* mutation, 14 had *TSC1* mutation, 2 had no mutation identified, and 5 were not tested. Seven patients had SGCT resection for an indication of hydrocephalus (Online Supplemental Data). The average age at surgical resection was 13.2 years (range, 2–25.2 years).

Imaging Data

A total of 271 MRIs were analyzed. The average number of MRIs per patient was 6.8 (SD, 2.8). The average time between the earliest and most recent MR imaging was 9 (SD 3.9) years. The average time between MRIs was 1.6 (SD 1.0) years (Online Supplemental Data).

Lesion Categorization

We measured a total of 163 nodules at the FOM: 126 small lesions, 27 indeterminate lesions, and 10 resected lesions. Seven
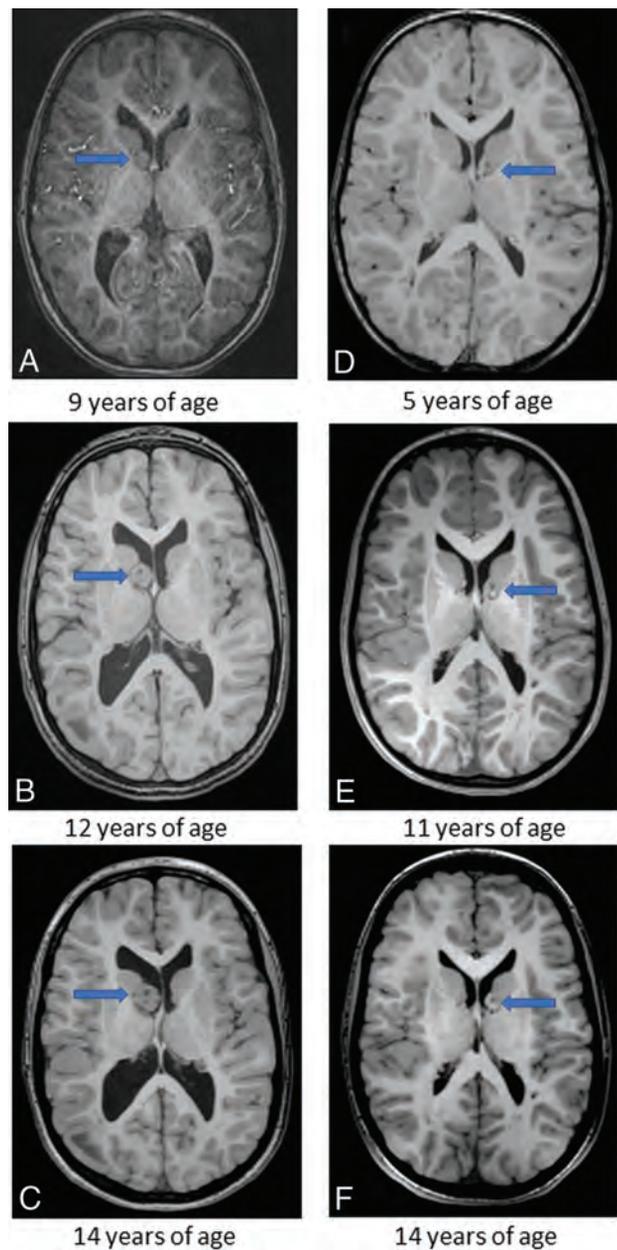

FIG 1. Serial images of patient 11's resected lesion located in the region of the right foramen of Monro (*A–C*) and patient 2's indeterminate lesion located in the region of left foramen of Monro (*D–F*). Patient 11's lesion was resected at 16 years of age. The *arrows* point to the lesion of interest. Cubic measurements are as follows: (AP × TV × SI), *A*, 9.6 mm × 8.8 mm × 10.6 mm; *B*, 16.1 mm × 15.9 mm × 16.3 mm; *C*, 21.5 mm × 16.5 mm × 19.3 mm; *D*, 10.4 mm × 9.2 mm × 11.2 mm; *E*, 10.5 mm × 9.5 mm × 11.7 mm; *F*, 10.4 mm × 9.5 mm × 11.9 mm.

small lesions and 1 resected lesion appeared new—that is, not seen on an earlier MR imaging but appearing on a later MR imaging, in which this interval appearance might, however, have been attributable to technical factors. Of the 7 patients who underwent surgical resection for an indication of obstructive hydrocephalus, 3 had 2 clearly distinct nodules at the FOM and thus were analyzed for growth separately, yielding 10 resected lesions. Examples of serial imaging of a resected lesion and indeterminate lesion can be seen in Fig 1.

Indeterminate Lesions

A, Volume of indeterminate lesions by age at scan.

Resected Lesions

B, Volume of resected lesions by age at scan.
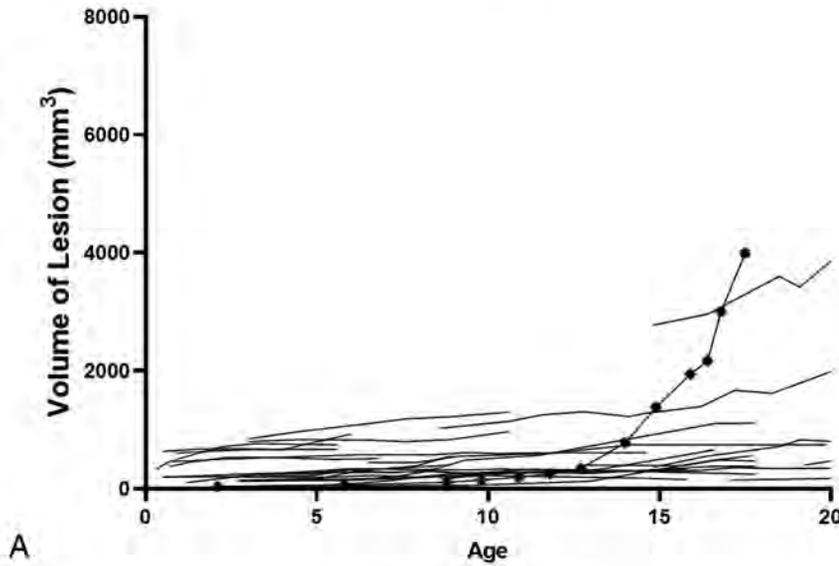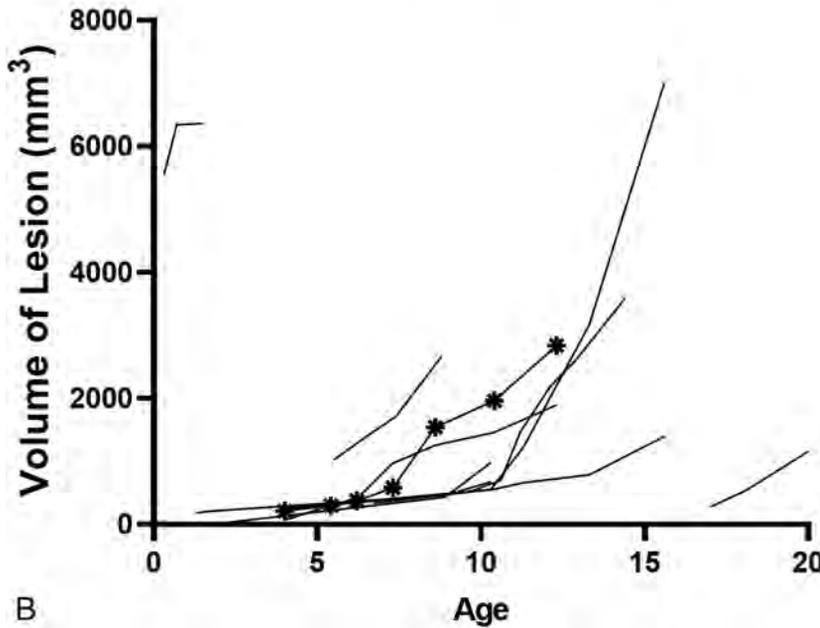
FIG 2. *A,* Volume of indeterminate lesions by age at scan. Each *line* indicates 1 indeterminate nodule. The *marked line* indicates that an indeterminate nodule in patient 23 shows rapid growth. There is currently no evidence of obstructive hydrocephalus. The patient is being closely monitored. *B,* Volume of resected lesions by age at scan. The marked line indicates patient 4's resected lesion. Patient 4's resected lesion had an average growth of 531.3 mm³/year between 6.2 and 7.3 years of age and a period of decelerated growth between 7.3 and 8.7 and 8.7–10.4 years of age, in which the average growth rates were 204.7 mm³/year and 118.9 mm³/year, respectively. At 12.7 years of age, the patient underwent surgical resection for obstructive hydrocephalus.

The average number of lesions measured per patient was 4 (SD, 2.8). The average number of small lesions per patient with at least 1 small lesion measured was 3.5 (SD, 2.5). The average number of indeterminate lesions per patient with at least 1 indeterminate lesion measured was 1.5 (SD, 0.8). The average number of resected lesions per patient with at least 1 resected lesion measured was 1.4 (SD, 3.3).

Volume Evaluation
The volumes of indeterminate nodules and resected lesions at each age recorded can be seen in Fig 2*A,* -*B,* respectively. At the earliest MR imaging, there was no significant difference between indeterminate lesion volume and resected lesion volume ($P = 1.0$). Resected lesions grew substantially during the surveillance period, and at the most recent MR imaging, the resected lesions were significantly larger than the indeterminate lesions ($P < .001$) (Fig 3).

Growth Evaluation
The median growth velocity and IQR for all nodules in the 0–5, 6–10, 11–15. and 16–20 years of age periods were 1.685 mm³/year (IQR = 22.7), 0.67 mm³/year (IQR = 7.1), 1.03 mm³/year (IQR = 7.7), and 1.08 mm³/year (IQR = 10.01), respectively. There was no significant difference in the velocity of growth comparing 0–5, 6–10, 11–15, and 16–20 years of age ($P = .68$). This analysis was repeated for small lesions ($P = .68$), resected lesions ($P = .62$), and indeterminate lesions ($P = .18$). No significant difference in growth velocities was observed.

Lesions types were compared within each age group. The median velocity of growth can be seen in Tables 1 and 2. Small lesions grew significantly slower than resected lesions during 0–5, 6–10, and 11–15 years of age and significantly slower than indeterminate lesions at 0–5, 6–10, and 16–20 years of age. There was a significantly faster velocity of growth exhibited by resected lesions
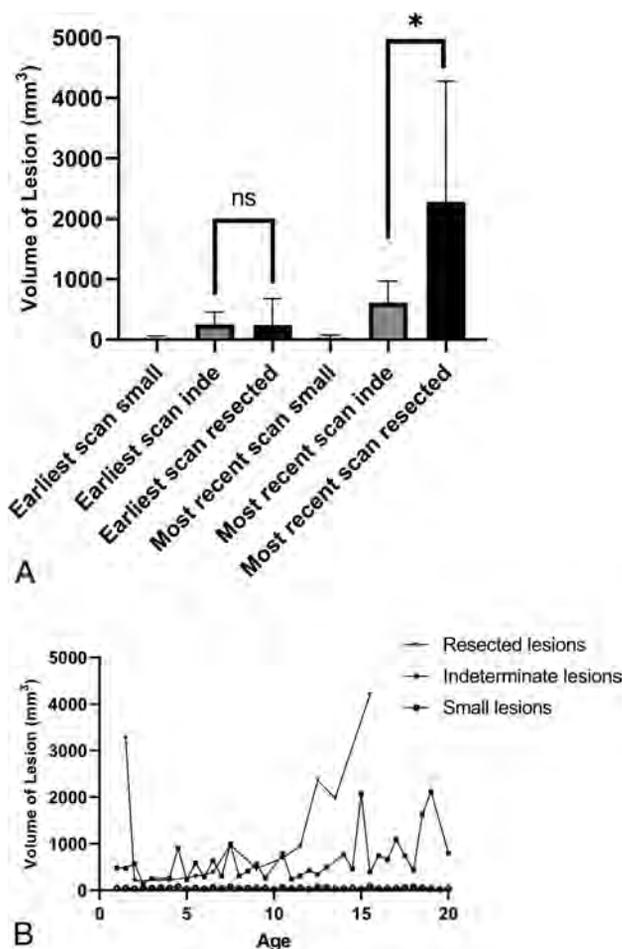
FIG 3. *A*, Median and IQR of lesion volumes of small lesions (small), indeterminate nodules (inde), and resected lesions (resected). The median volume and IQR at the earliest scan: small lesions, 25.22 mm³ (IQR = 50.8 mm³, *n* = 126); indeterminate lesions, 251.9 mm³ (IQR = 319.4 mm³, *n* = 27), and resected lesions, 245.6 mm³ (IQR = 486.5 mm³, *n* = 10). The median volume and IQR at most the recent scan: small lesions, 33.18 mm³ (IQR = 61.8 mm³, *n* = 126); indeterminate lesions, 615.6 mm³ (IQR = 581.1 mm³, *n* = 27); and resected lesions, 2283.0 mm³ (IQR = 3162.0 mm³; *n* = 10). Indeterminate lesion volume and resected lesion volume are compared at earliest scan and most recent scan using the Mann-Whitney *U* test. There is no significant difference in volume at the earliest scan (*P* = 1.0); however, there is a significant difference at the most recent scan. The *asterisk* indicates *P* < .001. *B*, Median lesion volume of each lesion type by age. Each point is the median volume of ≥2 lesions at a given age. NS indicates not significant.

compared with indeterminate lesions between 6–10 and 11–15 years of age. Comparisons of resected lesions with other subgroups could not be made during the 16–20 years of age period because only 1 resected lesion had preoperative MRI during this time period. A mixed-effects analysis found a significant interaction between age periods and lesion type (*P* < .001).

The highest percentage of lesions that grew >4.2 mm³/year occurred between 0 and 5 years of age in all lesion categories. All resected lesions grew in all age periods measured (Table 3).

The median and IQR of velocity of growth and acceleration of growth in each measured time point can be seen in Fig 4*A*, -*B*, respectively. Resected lesions had a significantly faster growth velocity between MR imaging scans compared with indeterminate lesions (*P* < .001) and small lesions (*P* < .001), and indeterminate lesions had significantly faster growth velocities compared with small lesions (*P* < .001) (Fig 4*A*). A comparison of growth acceleration showed significantly higher rates of acceleration in resected lesions compared with indeterminate lesions (*P* = .014) and small lesions (*P* = .004). There was not a significant difference in the acceleration between indeterminate lesions and small lesions (*P* = 1.0).

DISCUSSION

Studies that examine the growth of SGCTs have been limited to case series focused on those that have been surgically resected.[3] Estimates of growth of SGCTs have ranged from 1 mm/year to 65.2 mm/year in maximum diameter.[3] Our study characterizes growth velocities not only for SGCTs but also for SENs and found, on serial imaging, that there were significant differences in growth among resected lesions, indeterminate lesions, and small lesions. Additionally, resected lesions exhibited growth rates distinct from those of all other nodules at the FOM between 6 and 15 years of age. Therefore, measurement of growth rates from 6 to 15 years of age may help to identify those nodules that will continue to grow to a size that would result in obstructive hydrocephalus.

Prior consensus guidelines have proposed static criteria to identify SGCTs. Jóźwiak et al[9] defined SGCTs as tumors of >0.5 cm in maximum diameter, typically localized near the FOM, with documented growth. Roth et al[6] defined SGCTs as lesions of >1 cm in maximum diameter at the caudothalamic groove or a lesion in any location that has exhibited growth on serial imaging regardless of size. In our study, indeterminate

Table 1: Median growth velocities[a]

	0–5 Years of Age	6–10 Years of Age	11–15 Years of Age	16–20 Years of Age
Resected lesions	333.2 mm³/year 58.1–608.23[b] *n* = 2	351. mm³/year IQR = 344.5 *n* = 4	460.3 mm³/year IQR = 647.7 *n* = 5	292.0 mm³/year NA *n* = 1
Indeterminate lesions	32.5 mm³/year IQR = 58.1 mm³/year *n* = 11	18.8 mm³/year IQR = 37.0 mm³/year *n* = 16	15.8 mm³/year IQR = 43.9 mm³/year *n* = 13	15.7 mm³/year IQR = 103.6 mm³/year *n* = 13
Small lesions	0.6 mm³/year IQR = 4.2 mm³/year *n* = 42	0.3 mm³/year IQR = 1.6 mm³/year *n* = 59	0.3 mm³/year IQR = 3.0 mm³/year *n* = 39	0.6 mm³/year IQR = 3.1 mm³/year *n* = 32

Note:—NA indicates not applicable.
[a] Listed are the median growth and IQR for each lesion type for each age period.
[b] Range and not IQR are reported for resected lesions at 0–5 years of age.

Table 2: Growth velocity comparisons[a]

	0-5 Years of Age	6-10 Years of Age	11-15 Years of Age	16-20 Years of Age
Resected vs inde	$P = .62$	$P = .007^b$	$P = .006^b$	NA
Resected vs small	$P = .01^b$	$P < .001^b$	$P < .001^b$	NA
Resected vs inde + small	$P = .03^b$	$P < .001^b$	$P < .001^b$	NA
Inde vs small	$P < .001^b$	$P < .001^b$	$P = .053$	$P = .034^b$

Note:—Inde indicates indeterminate lesions; small, small lesions; resected, resected lesions; NA, not applicable.
[a] P-values calculated using Mann-Whitney U test and Bonferroni correction.
[b] Statistical significance.

Table 3: Percentage of lesions with growth rates > 4.2 mm³/year[a]

	Baseline to Most Recent Scan	0–5 Years of Age	6–10 Years of Age	11–15 Years of Age	16–20 Years of Age
Small lesions	14%	21%	12%	10%	16%
Indeterminate lesions	96%	100%	81%	62%	69%
Resected lesions	100%	100%[b]	100%	100%	100%[c]
All lesions	34%	41%	35%	30%	33%

[a] 4.2 mm³ was chosen to reflect an increase in growth of 2 mm in the AP, TV, and SI dimensions using an ellipsoid volume formula because a 2-mm increase would be an appreciable measurement on MR imaging.
[b] Of the 3 patients with symptomatic SGCTs in this age group, only 2/3 MRIs were considered diagnostic.
[c] Only 1 lesion was available for measurement.

nodules were quite common and not distinguishable from resected lesions by size criteria at the earliest MR imaging but were distinguishable by growth velocities on serial imaging. Moreover, our data clearly show that a considerable percentage of small lesions and indeterminate nodules grew in each age period studied but never resulted in obstructive hydrocephalus.

Our data suggest that comparison of growth velocities can identify SGCTs and that mild growth alone does not reliably identify those nodules that will result in obstructive hydrocephalus. Other studies have recommended that once any growth is proved on serial imaging, resection should be performed.[10,11] However, in our data, enlargement of nonobstructing nodules was frequently observed (Table 3). Hence, a recommendation of resection on the basis of any interval growth may be too nonspecific, but rather, closer monitoring may be indicated.

Prior consensus studies recommend that children with TSC be scanned every 1–3 years.[6] In the 0- to 5-year age group, though a study with larger numbers would be required to address this issue more definitively, our data raise the question of whether a child at 1 year of age who has only a small nodule at the FOM has a low risk of progression to obstructive hydrocephalus in the 0- to 5-year age period.

There are several caveats suggested from our study. First, indeterminate, nonobstructing lesions could exhibit high velocity and acceleration of growth (Fig 2A). Second, a growth deceleration did not preclude eventual progression to SGCT when scanning within the recommended consensus guidelines of every 1–3 years (Fig 2B).

Growth Velocity on Serial Imaging

***
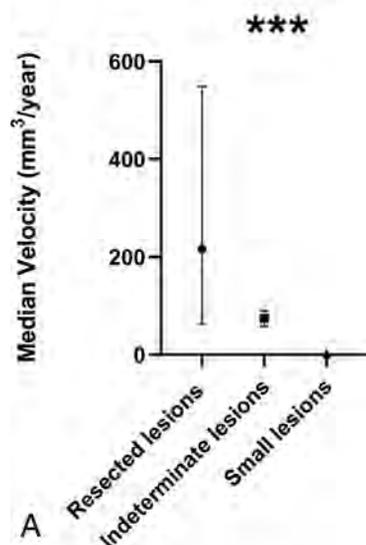
Growth Acceleration on Serial Imaging

**
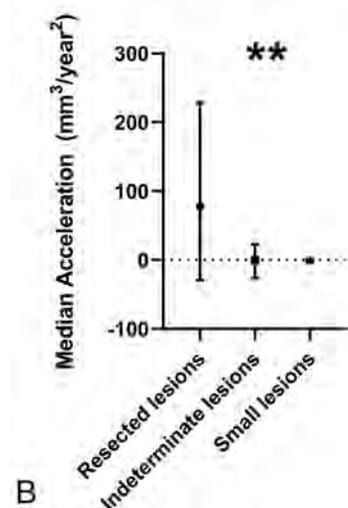
FIG 4. A, Median growth velocity between scans for each lesion type: small lesions, 0.45 mm³/year (IQR = 3.2 mm³/year, n = 789); indeterminate lesions, 21.95 mm³/year (IQR = 71.6 mm³/year; n = 167); resected lesions, 216.1 mm³/year (IQR = 487.7 mm³/year; n = 38). A Mann-Whitney U test with a Bonferroni correction was used to compare the velocity of growth between lesion types. The *triple asterisks* indicate a significant difference in velocity comparing resected lesions and indeterminate lesions (P <.001), resected lesions and small lesions (P <.001), and indeterminate lesions and small lesions (P <.001). B, Median acceleration of growth of all lesion types: small lesions, 0.00 mm³/year² (IQR = 3.4 mm³/year², n = 673); indeterminate lesions, −0.18 mm³/year² (IQR = 70.63 mm³/year², n = 141); resected lesions, 77.01 mm³/year² (IQR = 257.7 mm³/year²; n = 30). A Mann-Whitney U test with a Bonferroni correction was used to compare the acceleration of growth among lesion types. The *double asterisks* indicate significant differences between resected lesions and indeterminate lesions (P = .01) and resected lesions and small lesions (P = .004). There is not a significant difference in the rate of acceleration between indeterminate lesions and small lesions (P = 1.0).

This study has several limitations. First the study is limited by sample size: Although all patients with lesions resected between 2002 and 2018 were included, for the small lesions and the indeterminate nodules, all patients who met the inclusion criteria were randomly selected using a random number generator. Second, this was a retrospective study with a small cohort. Third, we evaluated only lesions that were in the region of the FOM. Fourth, we studied only growth curves and did not look at other MR imaging characteristics. Fifth, most patients in our cohort were studied between 0 and 15 years of age. It is generally believed that SGCTs do not develop after the first 20–25 years of life.[3] It is possible that growth and development of an SGCT could be seen in the future in those patients who have not reached the age of 20–25 years.[3] Finally, our measurement technique of cubic measurements likely overestimates the actual size of lesions.[12] However, the method we implemented provides a practical and reproducible approach to the measurement of nodules at the FOM and allows volume comparisons within our study.

Despite the limitations of this study, our results show clear differences in growth rates across lesion types, which may serve to distinguish obstructing lesions from nonobstructing lesions. A larger prospective study that tracks growth rates may be able to refine current diagnostic criteria of SGCTs. Additionally, our data suggest that obstructing lesions tend to be larger than nonobstructing lesions. Therefore, a future study may be able to define a critical volume range at which a lesion is likely to result in hydrocephalus. The distinction between subependymal nodules and SGCTs has important clinical implications because the management of these lesions is influenced by the ability of a lesion to grow and become symptomatic. Appropriate and consistent categorization will lead to better clinical outcomes.

CONCLUSIONS

SGCTs show growth velocities distinct from those of SENs at the FOM. Measurement of growth velocity and acceleration on serial imaging can help differentiate SGCTs from SENs in the first 20 years of life. Additionally, 6–15 years of age may be an important period to monitor SGCTs at the FOM because increased growth in this period may indicate that the lesions will continue to grow and result in obstructive hydrocephalus.

REFERENCES

1. Curatolo P, Moavero R, de Vries PJ. **Neurological and neuropsychiatric aspects of tuberous sclerosis complex.** *Lancet Neurol* 2015;14:733–45 CrossRef

2. Goh S, Butler W, Thiele EA. **Subependymal giant cell tumors in tuberous sclerosis complex.** *Neurology* 2004;63:1457–61 CrossRef Medline

3. Chan DL, Calder T, Lawson JA, et al. **The natural history of subependymal giant cell astrocytomas in tuberous sclerosis complex: a review.** *Rev Neurosci* 2018;29:295–301 CrossRef Medline

4. Adriaensen M, Schaefer-Prokop CM, Stijnen T, et al. **Prevalence of subependymal giant cell tumors in patients with tuberous sclerosis and a review of the literature.** *Eur J Neurol* 2009;16:691–96 CrossRef Medline

5. Kumari K, Sharma MC, Kakkar A, et al. **Role of mTOR signaling pathway in the pathogenesis of subependymal giant cell astrocytoma: a study of 28 cases.** *Neurol India* 2016;64:988–94 CrossRef Medline

6. Roth J, Roach ES, Bartels U, et al. **Subependymal giant cell astrocytoma: diagnosis, screening, and treatment: recommendations from the International Tuberous Sclerosis Complex Consensus Conference 2012.** *Pediatr Neurol* 2013;49:439–44 CrossRef Medline

7. Nabbout R, Santos M, Rolland Y, et al. **Early diagnosis of subependymal giant cell astrocytoma in children with tuberous sclerosis.** *J Neurol Neurosurg Psychiatry* 1999;66:370–75 CrossRef Medline

8. Roach ES, Sparagana SP. **Diagnosis of tuberous sclerosis complex.** *J Child Neurol* 2004;19:643–49 CrossRef Medline

9. Jóźwiak S, Nabbout R, Curatolo P, participants of the TSC Consensus Meeting for SEGA and Epilepsy Management. **Management of subependymal giant cell astrocytoma (SEGA) associated with tuberous sclerosis complex (TSC): clinical recommendations.** *Eur J Paediatr Neurol* 2013;17:348–52 CrossRef Medline

10. de Ribaupierre S, Dorfmüller G, Bulteau C, et al. **Subependymal giant-cell astrocytomas in pediatric tuberous sclerosis disease: when should we operate?** *Neurosurgery* 2007;60:83–90 CrossRef Medline

11. Fohlen M, Ferrand-Sorbets S, Delalande O, et al. **Surgery for subependymal giant cell astrocytomas in children with tuberous sclerosis complex.** *Childs Nerv Syst* 2018;34:1511–19 CrossRef Medline

12. Stawiski K, Trelińska J, Baranska D, et al. **What are the true volumes of SEGA tumors? Reliability of planimetric and popular semi-automated image segmentation methods.** *MAGMA* 2017;30:397–405 CrossRef Medline

Persistent Trigeminal Artery: A Novel Imaging Finding in CHARGE Syndrome

A. Siddiqui, P. Touska, D. Josifova, and S.E.J. Connor

ABSTRACT

BACKGROUND AND PURPOSE: Coloboma of the eye, Heart defects, Atresia of the choanae, Retardation of growth and/or development, Genital and/or urinary abnormalities, Ear abnormalities and deafness (CHARGE) syndrome is an autosomal dominant genetic disorder with evolving clinical diagnostic criteria. Recently, a number of additional anomalies have been described in this syndrome, which may aid in early diagnosis, particularly in incomplete phenotypes or atypical cases. The persistent trigeminal artery is an embryonic carotid-vertebral anastomosis, rarely seen in the healthy population, with a reported prevalence of 0.4%. Because we had observed the persistent trigeminal artery in patients with CHARGE syndrome, this study aimed to explore the prevalence of the persistent trigeminal artery in this syndrome.

MATERIALS AND METHODS: A retrospective study was performed at our tertiary center. MR imaging studies, clinical records, and genetic results were reviewed for patients diagnosed with CHARGE syndrome between 2006 and 2019. The prevalence of the persistent trigeminal artery in patients with CHARGE syndrome was recorded and compared with other established diagnostic criteria.

RESULTS: Twenty-five patients with CHARGE syndrome were included. The persistent trigeminal artery was demonstrated on MR imaging in 14/25 (56%) patients and was seen more frequently than 4 of 9 other established diagnostic criteria in our cohort. When individual major or minor diagnostic criteria were absent, the persistent trigeminal artery was still demonstrated on MR imaging in 52%–67% of these patients with CHARGE syndrome.

CONCLUSIONS: The prevalence of the persistent trigeminal artery in CHARGE syndrome of 56% is higher than that of some other established diagnostic criteria and much higher than that in the general population. The persistent trigeminal artery may be a useful addition to the expanding phenotype of CHARGE syndrome, supplementing other diagnostic criteria. Radiologists should be aware of this novel finding demonstrable on MR imaging.

ABBREVIATIONS: CHARGE = Coloboma of the eye, Heart defects, Atresia of the choanae, Retardation of growth and/or development, Genital and/or urinary abnormalities, Ear abnormalities and deafness; *CHD7* = Chromodomain Helicase DNA-binding protein; PcomA = posterior communicating artery; PHACE = Posterior fossa/brain malformations, Hemangiomas, Arterial anomalies, Coarctation of the aorta, and Cardiac anomalies and Eye abnormalities; PTA = persistent trigeminal artery

Coloboma of the eye, Heart defects, Atresia of the choanae, Retardation of growth and/or development, Genital and/or urinary abnormalities, and Ear abnormalities and deafness (CHARGE) syndrome is a rare, usually sporadic, autosomal dominant disorder caused by heterozygous mutations in the chromodomian helicase DNA binding protein 7 (*CHD7*) gene in most

cases. While the acronym CHARGE was based on the cardinal features identified when the syndrome was initially described, the clinical diagnostic criteria have since evolved. They now include the 3C triad (Coloboma-Choanal atresia-abnormal semicircular Canals), olfactory hypoplasia, and/or arhinencephaly and rhombencephalic dysfunctions, which are now considered the most important and constant clues to the diagnosis.[1-5]

The persistent trigeminal artery (PTA) is an embryonic carotid-vertebral anastomosis that is normally transiently seen in early fetal life. Because it is rarely reported in the healthy population, with a prevalence of 0.4%,[6] and is easily recognized on MR imaging, it would constitute a potentially useful syndromic association. Because we had observed the PTA on MR imaging studies in patients with CHARGE syndrome, our hypothesis was that the PTA would have a high prevalence in CHARGE syndrome

and would represent a useful diagnostic criterion. Hence, our primary objective was to determine the prevalence of the PTA on MR imaging in patients with CHARGE syndrome, and our secondary objective was to compare the prevalence of the PTA with that of other diagnostic criteria.

MATERIALS AND METHODS
Patients
A retrospective study on CHARGE syndrome was performed at Guy's and St. Thomas' NHS Foundation Trust with local institutional approval and the need for informed consent was waived. The radiology information data base (CRIS, Healthcare Software Solutions, UK) was searched for MR imaging studies performed between 2006 and 2019 containing the terms "CHARGE-syndrome" or "CHARGE syndrome" in the report or clinical details. The clinical records and imaging studies were reviewed, and a diagnosis of CHARGE syndrome was confirmed on the basis of either genetics or a defined clinical diagnostic criterion subcategorizing into typical, partial/incomplete, or atypical CHARGE.[5] The demographic data, clinical criteria, and results of genetic testing were recorded.

Imaging Analysis
Imaging was performed on both 1.5 ($n = 21$) and 3T ($n = 6$) MR imaging systems from the same manufacturer (Philips Healthcare), with 2 patients having undergone scanning on both systems. Most 1.5T MR imaging studies were performed on a single scanner, an Achieva 1.5T (Philips Healthcare), while the other studies were performed on a 3T Achieva scanner (Philips Healthcare). Axial T2WI was available for all patients (4- to 5-mm section thickness on 1.5T and 2-mm thickness on 3T). 20/25 patients also had additional thin-section (1- to 1.4-mm section thickness) heavily T2WI through the posterior fossa. Time-of-flight MR angiography was acquired in 6/25, including in 2 patients who did not have thin-section T2WI.

Imaging was reviewed by 2 experienced neuroradiologists (A.S. with 13 years' consultant/senior experience and S.E.J.C. with 18 years' consultant/senior experience), and agreement was reached by consensus. The presence of a PTA was recorded if a vessel was seen connecting the intracavernous ICA to the basilar artery. When present, the PTA was subclassified as either medial or lateral, depending on its origin from the medial or lateral surface of the ICA and/or a course relative to the abducens nerve or lateral border of the dorsum sellae. Hypoplasia of the basilar artery below the level of the PTA and the presence, absence or hypoplasia of the posterior communicating artery (PcomA) (when an MRA sequence was available) were also recorded.

Statistics
Statistical calculations were performed using SPSS Statistics software (IBM). The prevalence of the PTA, along with the other major and minor diagnostic criteria, in patients with CHARGE syndrome was calculated. The prevalence of each diagnostic criterion, along with the prevalence of the PTA was compared; 95% confidence intervals were derived using a 1-sample binomial (Clopper-Pearson) test. The Fisher exact test was used to compare PTA prevalence between defined groups. A P value $< .05$ was deemed significant.

RESULTS
Patient Characteristics and Diagnostic Criteria for CHARGE
Twenty-five patients (male = 16; female = 9) with a diagnosis of CHARGE syndrome were included in the study. Ages ranged from 2 days to 22 years (median = 1 month), with 12 neonates (0–1 months), 5 infants (1–12 months), and 8 older children/adults (1–22 years). Most neonatal MR imaging studies were performed for dysmorphology or cardiovascular problems, whereas MR imaging studies in infants and older children were performed for various other indications, including cranial nerve dysfunction and hearing loss. On the basis of clinical criteria,[5] 22/25 had a clinical diagnosis of typical CHARGE syndrome, 2/25 had atypical CHARGE syndrome (but confirmed genetically), and 1/25 had limited clinical information (also confirmed genetically). Genetic testing for mutations in the CHD7 gene was performed and confirmed in 17/25 cases; the remaining 8 cases had clinical confirmation of the diagnosis but had either not undergone genetic testing or the results were unavailable. With the Fisher exact test, there was no significant difference between the presence or absence of the PTA between the genetically confirmed and clinically confirmed cases ($P = .5$). Similarly, there was no significant difference in the prevalence of the PTA between patients with typical and atypical CHARGE syndrome ($P = .7$).

Imaging Findings of the PTA
The PTA was seen in 56% ($n = 14$) of patients with CHARGE syndrome (Figs 1 and 2). When seen, the PTA was medial in 79% (11/14) and lateral in 21% (3/14) of patients. The basilar artery below the PTA was hypoplastic in 36% (5/14) of cases, and the PcomA was absent or hypoplastic in all 6 patients who had undergone MRA. The PTA was seen in 8/14 males (57%) and 6/14 females (43%) and was unilateral in all patients. When present, it was also seen on standard 5mm thick-section T2WI of the head in 8/14 (57%). In 3 neonates who had undergone subsequent follow-up imaging in later childhood, the PTA was demonstrable on each occasion. The PTA was also seen in 4/8 patients who did not have genetic testing. Other imaging findings of CHARGE syndrome (Figs 1 and 2) were also recorded.

Comparison between the Presence of PTA and Other Diagnostic Criteria
The Table shows the comparison between the prevalence of the PTA and other diagnostic criteria in patients with CHARGE syndrome. In particular, the prevalence of the PTA (56%, 14/25) was higher than some of the established major and minor clinical criteria such as choanal atresia (40%, 10/25), gonadal anomalies (40%, 10/25), orofacial/laryngeal clefting (32%, 8/25), and tracheoesophageal fistula (8%, 2/25). Regarding an additional value when a major criterion was absent, the PTA was seen in 55% (5/9) of patients with CHARGE syndrome without documented colobomas and 53% (8/15) of patients without documented choanal atresia. Comparison with minor criteria also showed that the PTA was seen in 57% (4/7) of patients without cardiovascular anomalies, in 67% (10/15) of patients without gonadal anomalies, in 65% (11/17) of patients without orofacial/laryngeal clefting, and in 52% (13/23) of patients without a tracheoesophageal
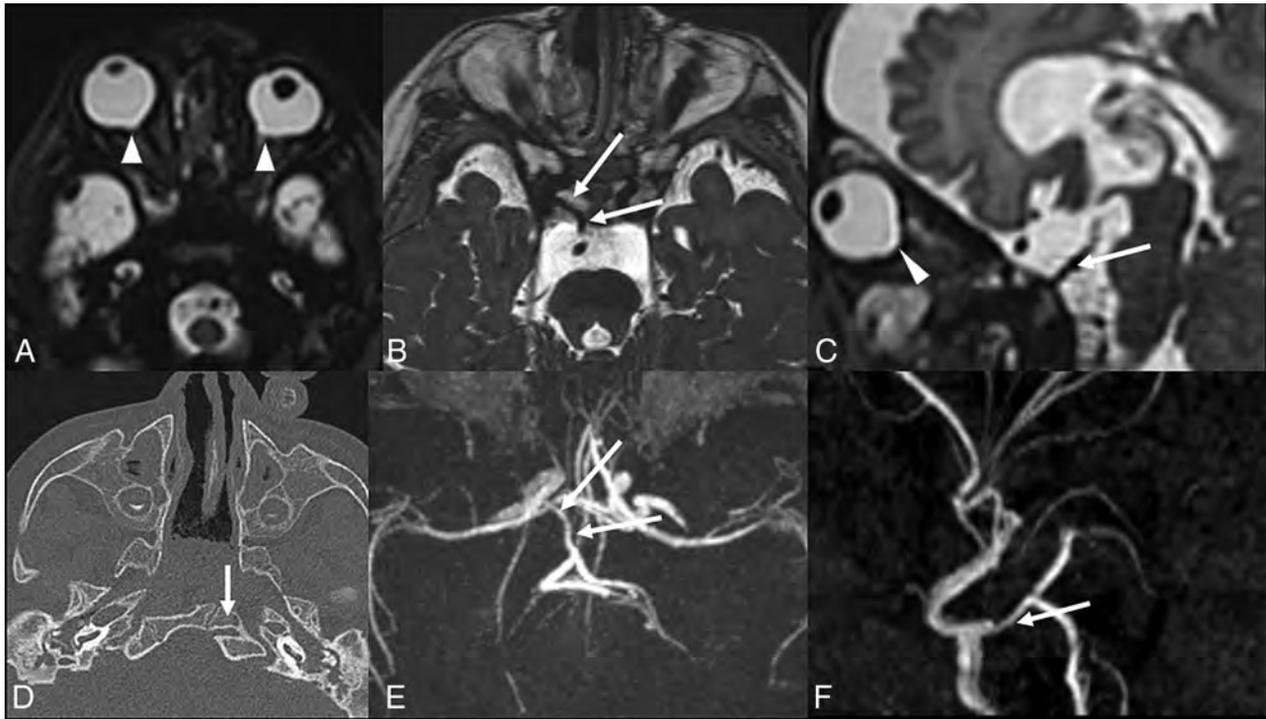
FIG 1. PTA in CHARGE syndrome, medial variant. Axial T2WI (*A* and *B*), sagittal T2WI (*C*), axial CT (*D*), and MRA transverse (*E*) and lateral (*F*) views in a neonate demonstrate a PTA (*long arrows*) connecting the intracavernous ICA to the basilar artery. Note the bilateral ocular colobomas (*arrowheads*), absent semicircular canals, and clival cleft (*arrow* in *D*).
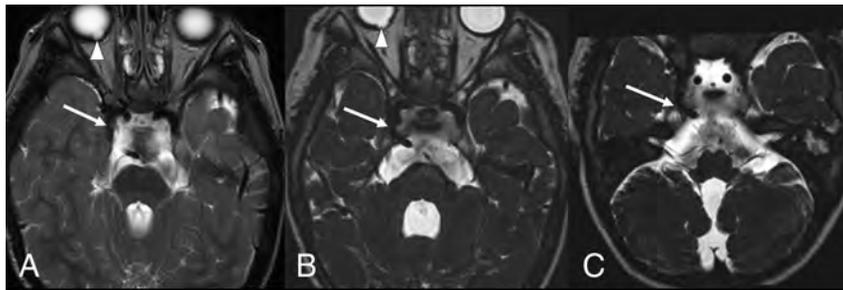

FIG 2. PTA in CHARGE syndrome, lateral variant. Standard 5-mm brain axial T2WI (*A*) and thin-section submillimetric T2WI (*B* and *C*) show a lateral-type PTA (*long arrows*), bilateral coloboma (*arrowheads*), and aplasia of the semicircular canals. The PTA is less clearly evident on the standard brain T2WI (*A*) compared with the thin-section images (*B* and *C*). Note the hypoplastic basilar artery below the PTA.

fistula, further supporting the additive complementary value of this finding.

DISCUSSION

The PTA was demonstrated on MR imaging in 56% (14/25) of cases in our cohort of patients with CHARGE syndrome. The prevalence of the PTA in patients with CHARGE syndrome was higher than 4 of 9 other established diagnostic criteria. The PTA was still present in most patients in whom specific major criteria (colobomas and choanal atresia) or minor criteria (cardiovascular anomalies, gonadal anomalies, orofacial/laryngeal clefting, tracheoesophageal fistula) were absent.

CHARGE Syndrome

CHARGE syndrome (or Hall-Hittner syndrome) was first described as an association of congenital anomalies in 1979.[1,2] The acronym CHARGE is based on the cardinal features identified when the syndrome was delineated: Coloboma, Heart malformation, choanal Atresia, Retardation of growth and/or development, Genital and/or urinary anomalies, and Ear anomalies.[3] Diagnosis is based on clinical criteria, which have been updated across time, with emphasis on the 3C triad (Coloboma, Choanal atresia, and abnormal semicircular Canals).[4,5] It was subsequently labeled as the CHARGE syndrome because the association of multiple anomalies was pathogenetically related following the identification of the *CHD7* gene.[7] Since that time, several new clinical findings have been added to the clinical spectrum of CHARGE syndrome, such as dysmorphic features, rhombencephalic dysfunction, cranial nerve dysfunction, hypoplasia of the semicircular canals, olfactory hypoplasia and/or arhinencephaly, feeding difficulties, esophageal anomalies, facial clefts, and hypothalamic-hypophyseal dysfunction.[8-13] The clinical spectrum of *CHD7* mutations and CHARGE continues to expand, and recent literature has focused on the presence of venous and skull base anomalies.[14,15]

Prevalence of PTA versus other CHARGE clinical criteria and of the PTA in cases in which the clinical criterion was absent

Clinical Criterion	Prevalence in Literature[12,25]	Prevalence in Our Study (95% CIs)	Prevalence of PTA When the Clinical Criterion Was Absent
PTA		56%, 14/25 (95% CI, 0.35–0.76)	
Coloboma	75%–90%	64%, 16/25 (95% CI, 0.42–0.82)	55%, 5/9 (95% CI, 0.21–0.86)
Choanal atresia	35%–65%	40%, 10/25 (95% CI, 0.21–0.61)	53%, 8/15 (95% CI, 0.27–0.79)
CHARGE ear	95%–100%	100%, 25/25 (95% CI, 0.86–1)	
Cranial nerve dysfunction	40–>95%	100%, 25/25 (95% CI, 0.86–1)	
Cardiovascular anomalies	50%–85%	72%, 18/25 (95% CI, 0.51–0.88)	57%, 4/7 (95% CI, 0.18–0.9)
Gonadal/genitourinary anomalies	50%–70%	40%, 10/25 (95% CI, 0.21–0.61)	67%, 10/15 (95% CI, 0.21–0.61)
Clefting, orofacial/larynx	15%–20%	32%, 8/25 (95% CI, 0.15–0.53)	65%, 11/17 (95% CI, 0.38–0.86)
Growth deficiency/developmental delay	70–>90%	100%, 25/25 (95% CI, 0.86–1)	
Tracheoesophageal fistula/anomalies	15%–20%	8%, 2/25 (95% CI, 0.01–0.26)	56%, 13/23 (95% CI, 0.34–0.77)

Genetically, CHARGE syndrome is a monogenic autosomal dominant disorder with an incidence of approximately 1:8500 to 1:12,500 in North America.[5,12,16] It results from a dysblastogenetic and dysneurulative process and could be related by a common pathogenetic mechanism resulting in disturbed neural crest development.[12] Mutations in the gene *CHD7* (in 8q12) were identified as causative for CHARGE syndrome in most (approximately two-thirds) patients with a clinical diagnosis of CHARGE syndrome.[7] *CHD7* belongs to a large family of evolutionarily conserved proteins thought to play a role in chromatin organization and is a regulatory element that potentially affects a large number of developmental pathways, explaining the pleiotropic nature of its phenotypic spectrum.[12]

While congenital cardiovascular anomalies are commonly seen in CHARGE syndrome, anomalies of the head and neck vessels have been rarely described, with occasional case reports of isolation of the carotid artery, isolation of the left subclavian artery from the pulmonary artery, and complex cervical arteriovenous fistulas.[17-19]

PTA

The PTA is one of the primitive carotid-vertebrobasilar anastomoses; it is the largest of these anastomotic vessels seen in fetal life and is seen in approximately 0.1%–1% of angiographic studies,[20,21] with a recent meta-analysis demonstrating a pooled prevalence of 0.4%.[6] In utero, the trigeminal artery supplies the basilar artery before development of the PcomA and vertebral arteries, with subsequent involution. The PTA arises from the junction between the petrous and cavernous ICA and runs posteriorly along the trigeminal nerve or crosses over or through the dorsum sellae. Vertebral, posterior communicating, and caudal basilar arteries are often hypoplastic. In 1959, Saltzman[22] reported 8 cases and proposed an angiographic classification for the PTA into 2 main types. In Saltzman type 1, or the fetal variant of the PTA, the basilar artery proximal to the insertion of the PTA may be hypoplastic and the PcomA may be absent, making the PTA a very important vessel supplying the entire basilar artery system distal to the anastomosis and providing flow to a large part of the posterior circulation, whereas in Saltzman type 2, or the adult variant, the PTA is a relatively less important vessel, with the posterior cerebral arteries receiving their blood supply predominantly through patent PcomAs and the basilar artery filling by the vertebral arteries.[21,23] While we were not able to

classify all PTAs according to the Saltzman system, in our series, the basilar artery was hypoplastic below the PTA in 36% (5/14) of patients and the PcomAs were hypoplastic in 6/6 patients who had undergone an MRA, suggesting that these were probably type 1, thereby making the PTA a potentially significant source of supply to the posterior circulation.

The exact etiology of the persistence of this vessel into postnatal life is unclear. The PTA is the most prominent of the carotid-basilar anastomoses, existing in early fetal life (4- to 5-mm embryonic stage) and providing an important source of flow to the developing rhombencephalon, arising from the developing ICA to connect with the paired longitudinal neural arteries, which would eventually form the basilar artery. Growth of the posterior fossa structures, interposition of the developing basisphenoid cartilages, and evolution of the carotid and vertebrobasilar systems result in obliteration of the PTA at around the 7- to 14-mm embryonic stage, with the PcomA superseding the PTA from above and the vertebral arteries developing from below.[21] The PTA has been reported in association with vascular malformations, aneurysms, trigeminal neuralgia/compression, and PHACE syndrome, with additional implications in surgical and endovascular treatment-planning, making it a relevant finding to note, with significance potentially emerging later in life.[20,21,24] One of our patients had a mild ICA hypoplasia contralateral to the PTA, but we did not find any vascular malformations or aneurysms in our study.

Association of the PTA and CHARGE Syndrome

To the best of our knowledge, this is the first report describing a compelling association of the PTA with CHARGE syndrome. This finding was readily identified on imaging studies and particularly when thin-section T2WI was available. While it is important for the neuroradiologists to recognize this association, it may also have an additional supplementary value when other criteria are lacking, particularly in cases of atypical or partial/incomplete CHARGE syndrome. The presence of the PTA on neuroimaging studies, particularly in the context of a child with complex cardiovascular anomalies or rhombencephalic dysfunction, should prompt a review for other hitherto unsuspected imaging features of CHARGE syndrome.

The prevalence of PTA of 56% (14/25) in cases of CHARGE syndrome is particularly high, given the relatively low prevalence of the PTA (0.4%).[6] Similarly, we have not found any reference to the PTA or intracranial arterial anomalies in the context of

CHARGE syndrome, including large reviews on this subject.[25,26] This could potentially be due to the lack of thin-section imaging of the posterior fossa or lack of MR angiographic imaging. The PTA was clearly identified on the thin-section T2WI performed for assessment of the inner ears or when thin-section T2WI of the head was available, but it could also be demonstrated in 57% (8/14) on the standard T2WI of the head (Fig 2). There was no particular correlation with sex. We also found that a substantial proportion of patients with CHARGE syndrome, who did not have one of the major or minor diagnostic criteria, had the presence of the PTA, as shown in the Table, in more than half of the cases. This is a strong argument for including the PTA as an additional diagnostic criterion of CHARGE syndrome, to further support the other diagnostic criteria, increase diagnostic confidence, and direct appropriate investigations. Indeed, 2 patients who did not have a typical CHARGE diagnosis (but were genetically confirmed) had a PTA.

CHD7 is expressed ubiquitously in fetal and adult tissues, including eye, olfactory epithelium, inner ear, and the vascular system.[7] The in situ hybridization analysis of the *CHD7* gene during early human development has shown a good correlation between the specific *CHD7* expression pattern and the developmental anomalies observed in CHARGE syndrome.[12] Therefore, it is plausible that mutations in this gene in some could result in vascular maldevelopment and persistence of embryonic vessels. Embryologically, the PTA is a significant source of supply to the developing hindbrain in very early fetal life at around 4–6 weeks but has a very limited time span of approximately a week or so, with normal regression of the PTA paralleling the simultaneous development of the PcomA and vertebral arteries, which then take over the function of providing flow to the posterior circulation and the rapidly developing hindbrain.[21,27,28] It is also well-established that hindbrain dysfunction is a characteristic feature of CHARGE syndrome; therefore, the finding of parallel anomalous vascular development is not surprising. Embryologically, CHARGE is thought to be a complex neurocristopathy, related to abnormalities of neural crest cells and abnormal differentiation of the cephalic mesoderm and ectoderm occurring between 3 and 9 weeks of gestation.[5] This timing also corresponds to and overlaps the time of development and regression of the PTA, thereby potentially explaining its persistence in CHARGE.

PTA versus Other Criteria in CHARGE
Compared with the other accepted diagnostic criteria, the PTA fares well, and the prevalence of PTA in our series (56%, 14/25) is higher than some major and minor criteria (Table). It is well-documented that semicircular canal aplasia is seen in almost all cases of CHARGE syndrome, as was seen in 100% of our cases. However, regarding other major criteria, ocular colobomas were seen in only 64% (16/25), and choanal atresia, in 40% (10/25) of our cohort. Previous studies have indicated the prevalence of ocular colobomas to be 75%–90% and choanal atresia to be 35%–65%.[12] The prevalence of minor criteria in our study was variable, and the PTA also compared favorably. This finding is a strong argument in support of including the PTA as an additional diagnostic criterion. It is also topographically distinct from the other established clinical criteria and adds the missing arterial angle to the pre-existing broad clinicoradiologic phenotype of CHARGE syndrome. In addition to our findings of the high prevalence of the PTA, recent studies have also demonstrated a high prevalence of clival pathology in CHARGE syndrome, with characteristic clival clefts found in nearly 90% of cases, and cerebellar malformations/heterotopia, in >70%, further arguing for expansion of the existing clinical criteria.[15,29,30]

Limitations
There are limitations to a retrospective study. In particular, there were variable MR imaging sequences and technical parameters, and MRA was available in only 6/25 patients. However, most of our patients (20/25) had undergone thin-section T2WI, primarily for assessment of the vestibulocochlear nerves and inner ears, and the PTA was well-demonstrated on this sequence. In 3 patients without MRA or thin-section T2WI, the axial head T2WI did not show an obvious PTA, and we acknowledge the possibility that a small PTA may have been present. When one compares the utility of the PTA with other diagnostic criteria, one must consider their prevalence in other syndromic and non-syndromic populations. We did not look at the prevalence of the PTA in other syndromes or similar conditions; therefore, we are unable to assess the specificity of this finding and how frequently it is present in other related syndromes, though we did not find any reference to the PTA in chromosomal syndromes overlapping with CHARGE syndrome such as 22q11.2 deletion or cat eye syndromes. We also acknowledge that the available clinical data may have been limited in some of our patients. Additional prospective multi-institutional studies would be helpful in further validation of this interesting finding and bringing into focus the association of arterial anomalies with CHARGE syndrome, which may have significance in later life.

CONCLUSIONS
This is the first report of an association between the PTA and CHARGE syndrome. With the identification of a high prevalence of the PTA in CHARGE syndrome (56%), which is higher than some other established diagnostic criteria and much higher than in the general population, our study adds another interesting and topographically distinct imaging finding to the existing broad phenotype of CHARGE syndrome.

REFERENCES
1. Hittner HM, Hirsch NJ, Kreh GM, et al. **Colobomatous microphthalmia, heart disease, hearing loss, and mental retardation: a syndrome.** *J Pediatr Ophthalmo l Strabismus* 1979;16:122–28 CrossRef Medline
2. Hall BD. **Choanal atresia and associated multiple anomalies.** *J Pediatr* 1979;95:395–98 CrossRef Medline
3. Pagon RA, Graham JM, Zonana J, et al. **Coloboma, congenital heart disease, and choanal atresia with multiple anomalies: CHARGE association.** *J Pediatr* 1981;99:223–27 CrossRef Medline

4. Blake KD, Davenport SL, Hall BD, et al. **CHARGE association: an update and review for the primary pediatrician.** *Clin Pediatr (Phila)* 1998;37:159–73 CrossRef Medline

5. Verloes A. **Updated diagnostic criteria for CHARGE syndrome: a proposal.** *Am J Med Genet A.* 2005;133:307–08 CrossRef Medline

6. Brzegowy K, Peôkala PA, Zarzecki MP, et al. **Prevalence and clinical implications of the primitive trigeminal artery and its variants: a meta-analysis.** *World Neurosurg* 2020;133:e401–11 CrossRef Medline

7. Vissers LE, van Ravenswaaij CM, Admiraal R, et al. **Mutations in a new member of the chromodomain gene family cause CHARGE syndrome.** *Nat Genet* 2004;36:955–57 CrossRef Medline

8. Lanson BG, Green JE, Roland JT, et al. **Cochlear implantation in children with CHARGE syndrome: therapeutic decisions and outcomes.** *Laryngoscope* 2007;117:1260–66 CrossRef Medline

9. Song MH, Cho HJ, Lee HK, et al. **CHD7 mutational analysis and clinical considerations for auditory rehabilitation in deaf patients with CHARGE syndrome.** *PLoS One* 2011;6:e24511 CrossRef Medline

10. Tellier AL, Cormier-Daire V, Abadie V, et al. **CHARGE syndrome: report of 47 cases and review.** *Am J Med Genet* 1998;76:402–09 CrossRef Medline

11. Blustajn J, Kirsch CF, Panigrahy A, et al. **Olfactory anomalies in CHARGE syndrome: Imaging findings of a potential major diagnostic criterion.** *AJNR Am J Neuroradiol* 2008;29:1266–69 CrossRef Medline

12. Sanlaville D, Verloes A. **CHARGE syndrome: an update.** *Eur J Hum Genet* 2007;15:389–99 CrossRef Medline

13. Morimoto AK, Wiggins RH, Hudgins PA, et al. **Absent semicircular canals in CHARGE syndrome: radiologic spectrum of findings.** *AJNR Am J Neuroradiol* 2006;27:1663–71 Medline

14. Friedmann DR, Amoils M, Germiller JA, et al. **Venous malformations of the temporal bone are a common feature in CHARGE syndrome.** *Laryngoscope* 2012;122:895–900 CrossRef Medline

15. Mahdi E, Whitehead MT. **Clival malformations in CHARGE syndrome.** *AJNR Am J Neuroradiol* 2018;39:1153–56 CrossRef Medline

16. Issekutz KA, Graham JM, Prasad C, et al. **An epidemiological analysis of CHARGE syndrome: preliminary results from a Canadian study.** *Am J Med Genet A* 2005;133:309–17 CrossRef Medline

17. Dashti SR, Spetzler RF, Park MS, et al. **Multimodality treatment of a complex cervicocerebral arteriovenous shunt in a patient with CHARGE syndrome: case report.** *Neurosurgery* 2010;67:E208–09 CrossRef Medline

18. Bech AP, Op Den Akker J, Matthijsse PR. **Isolation of the left subclavian artery from the pulmonary artery in a patient with CHARGE association.** *Congenit Anom (Kyoto).* 2010;50:200–02 CrossRef Medline

19. Ghalili K, Issenberg HJ, Freeman NJ, et al. **Isolated left carotid artery in CHARGE association: diagnosis and repair.** *Ann Thorac Surg* 1990;50:130–32 CrossRef Medline

20. Azab W, Delashaw J, Mohammed M. **Persistent primitive trigeminal artery: a review.** *Turk Neurosurg* 2011;22:399–406 CrossRef Medline

21. Meckel S, Spittau B, McAuliffe W. **The persistent trigeminal artery: development, imaging anatomy, variants, and associated vascular pathologies.** *Neuroradiology* 2013;55:5–16 CrossRef Medline

22. Saltzman GF. **Patent primitive trigeminal artery studied by cerebral angiography.** *Acta Radiol* 1959;51:329–36 CrossRef Medline

23. McKenzie JD, Dean BL, Flom RA. **Trigeminal-cavernous fistula: Saltzman anatomy revisited.** *AJNR Am J Neuroradiol* 1996;17:280–82 Medline

24. Diana F, Mangiafico S, Valente V, et al. **Persistent trigeminal artery aneurysms: case report and systematic review.** *J Neurointerv Surg* 2019;11:1261–65 CrossRef Medline

25. van Ravenswaaij-Arts CM, Hefner M, Blake K, et al. **CHD7 disorder.** *GeneReviews* October 2, 2006. Updated September 17, 2020. https://www.ncbi.nlm.nih.gov/books/NBK1117/. Accessed May 4, 2021

26. Zentner GE, Layman WS, Martin DM, et al. **Molecular and phenotypic aspects of CHD7 mutation in CHARGE syndrome.** *Am J Med Genet A* 2010;152:674–86 CrossRef Medline

27. Fields WS. **The significance of persistent trigeminal artery: carotid-basilar anastomosis.** *Radiology* 1968;91:1096–1101 CrossRef Medline

28. Menshawi K, Mohr JP, Gutierrez J. **A functional perspective on the embryology and anatomy of the cerebral blood supply.** *J Stroke* 2015;17:144–58 CrossRef Medline

29. de Geus CM, Bergman JE, van Ravenswaaij-Arts CM, et al. **Imaging of clival hypoplasia in CHARGE syndrome and hypothesis for development: a case-control study.** *AJNR Am J Neuroradiol* 2018;39:1938–42 CrossRef Medline

30. Wright JN, Rutledge J, Doherty D, et al. **Cerebellar heterotopias: expanding the phenotype of cerebellar dysgenesis in CHARGE syndrome.** *AJNR Am J Neuroradiol* 2019;40:2154–60 CrossRef Medline

A Longitudinal Analysis of Early Lesion Growth in Presymptomatic Patients with Cerebral Adrenoleukodystrophy

E.J. Mallack, G. Askin, S. van de Stadt, P.A. Caruso, P.L. Musolino, M. Engelen, S.N. Niogi, and F.S. Eichler

ABSTRACT

BACKGROUND AND PURPOSE: Cerebral adrenoleukodystrophy is a devastating neurological disorder caused by mutations in the *ABCD1* gene. Our aim was to model and compare the growth of early cerebral lesions from longitudinal MRIs obtained in presymptomatic patients with progressive and arrested cerebral adrenoleukodystrophy using quantitative MR imaging–based lesion volumetry.

MATERIALS AND METHODS: We retrospectively quantified and modeled the longitudinal growth of early cerebral lesions from 174 MRIs obtained from 36 presymptomatic male patients with cerebral adrenoleukodystrophy. Lesions were manually segmented using subject-specific lesion-intensity thresholding. Volumes were calculated and plotted across time. Lesion velocity and acceleration were calculated between sequentially paired and triplet MRIs, respectively. Linear mixed-effects models were used to assess differences in growth parameters between progressive and arrested phenotypes.

RESULTS: The median patient age was 7.4 years (range, 3.9–37.0 years). Early-stage cerebral disease progression was inversely correlated with age ($\rho = -0.6631$, $P < .001$), early lesions can grow while appearing radiographically stable, lesions undergo sustained acceleration in progressive cerebral adrenoleukodystrophy ($\beta = 0.10$ mL/month^2 [95% CI, $0.05-0.14$ mL/month^2], $P < .001$), and growth trajectories diverge between phenotypes in the presymptomatic time period.

CONCLUSIONS: Measuring the volumetric changes in newly developing cerebral lesions across time can distinguish cerebral adrenoleukodystrophy phenotypes before symptom onset. When factored into the overall clinical presentation of a patient with a new brain lesion, quantitative MR imaging–based lesion volumetry may aid in the accurate prediction of patients eligible for therapy.

ABBREVIATIONS: CALD = cerebral adrenoleukodystrophy; HSCT = hematopoietic stem cell transplantation; LS = Loes score; t_0 = time-zero

X-linked adrenoleukodystrophy is a devastating neurologic disorder caused by mutations in the *ABCD1* gene, which lead to an accumulation of very long chain fatty acids in plasma and tissue.[1] Multiple phenotypes emerge with no genotype-phenotype relationship having been established.[2] Most patients will develop cerebral adrenoleukodystrophy (CALD), with the highest incidence occurring in childhood.[3] Lesions most often occur in the splenium (60%–80%) or genu (10%–15%) of the corpus callosum and spread confluently into the surrounding subcortical white matter.[4,5] More than 80% of children with CALD will experience inflammatory demyelination, "progressive CALD," followed by rapid neurodegeneration and death without treatment in 2–3 years.[6,7] Conversely, 15%–20% of children and most adults will undergo spontaneous arrest of disease, "arrested CALD" without evidence of brain inflammation and are ineligible for hematopoietic stem cell transplantation (HSCT).[8]

HSCT is most successful when initiated in the window before the onset of neurologic symptoms.[7,9-11] However, the presymptomatic window is narrow,[12,13] and only a minority of patients with

progressive CALD are identified in time for HSCT.[7] HSCT itself is associated with significant toxicity, morbidity, and mortality, the latter 2 of which increase with disease burden.[7,10] Early identification and treatment of CALD is a challenge.[7,12] Most important, the addition of adrenoleukodystrophy to the Recommended Uniform Screening Program has provided a new opportunity to alter the natural history of the disease by monitoring for CALD from birth.[14]

The MR imaging score, or Loes score (LS), is the standard metric for brain demyelination in CALD.[15] Previous work has quantified the increase in the MR imaging score per year stratified by lesion distribution.[5] However, the LS is not specific nor predictive of progression early in CALD and underestimates the true burden of cerebral disease.[4,16] There is an explicitly stated need for early biomarkers to aid in the accurate prediction of patients with CALD appropriate for therapy.[7,17]

Few prior studies focused on the kinetics of 3D CALD lesion growth in presymptomatic patients. We recently demonstrated that lesion growth is rapid when lesions are small, the growth rate is inversely correlated with age, and older patients tend to stabilize and undergo spontaneous arrest of disease.[4] The primary aim of this study was to model the growth of early cerebral lesions from longitudinal MRIs obtained in presymptomatic patients with CALD using lesion volumetry, a quantitative MR imaging approach. We hypothesized that the trajectories of lesion growth diverge between the progressive and arrested phenotypes before symptom onset. Hence tracking the growth of de novo cerebral lesions may enable the early diagnosis of patients with CALD appropriate for HSCT and thereby facilitate clinical decision-making around this progressive disease.

MATERIALS AND METHODS
Standard Protocol Approvals, Registrations, and Consents
Patient data were retrospectively reviewed and de-identified, and storage was encrypted and password-protected. Due to anonymization, consent was waived. This study was approved by each institution's ethics/institutional review board.

Subjects
We performed a retrospective review of patients with CALD evaluated at 3 institutions (Massachusetts General Hospital, Amsterdam University Medical Center, and Weill Cornell Medicine) to select cases for study inclusion. The inclusion criteria were the following: 1) confirmed diagnosis of adrenoleukodystrophy by genetic testing (*ABCD1* gene mutation) or high levels of plasma very long chain fatty acids;[18] 2) early-stage CALD[7,19] defined as no cerebral symptoms as defined by the neurologic functional score (range, 0–25) and the LS on first MR imaging with abnormal findings (ie, MR imaging at diagnosis) between 0.5 and nine; 3) ≥2 available brain MRIs per subject; 4) at least 1 axial T2-weighted sequence per MR imaging; 5) MRIs obtained in the pretreatment time period (no history of hematopoietic stem cell or gene therapy); and 6) pattern 1 or 2 lesion distributions only (genu or splenium of corpus callosum).

Patients were divided into 2 study cohorts by clinical diagnosis: progressive CALD versus arrested CALD. Diagnosis of arrested CALD was defined by ≥2 consecutive MRIs spanning a minimum of 6 months with no increase in LS and no contrast enhancement.[8] Analyses were limited to callosal lesions to compare growth parameters accurately between subjects over similar neuroanatomic boundaries. Included patients could have a neurologic functional score of >0 if the symptoms were attributed to adrenomyeloneuropathy. Treatment with Lorenzo Oil did not exclude patients from participation in the study.[1] MR imaging surveillance every 6 months between 3 and 10 years of age, published by the New York State Newborn Screening workgroup, was followed.[14] Otherwise, MR imaging and clinical follow-up were completed per clinical determination.

MR Imaging Parameters
MR imaging studies of the brain in the Amsterdam University Medical Center cohort were performed on a 3T MR imaging unit (Ingenia; Phillips Healthcare). MR imaging studies in the Weill Cornell Medicine cohort were performed on a 3T MR imaging unit (MR750; GE Healthcare). One-hundred forty-two of the 150 MR imaging studies in the Massachusetts General Hospital cohort were performed on a 3T MR imaging unit (Tim Trio [Siemens] until 2017, Prisma 2017 to present, [Siemens]). Eight studies obtained before 2015 were performed on 1.5T MR imaging units (Signa HDxt, GE Healthcare; or Ingenia, Philips Healthcare). Each subject was scanned longitudinally according to their site's standard clinical protocol. Axial conventional T2, conventional FLAIR, or sampling perfection with application-optimized contrasts by using different flip angle evolution (SPACE, Siemens) FLAIR volumetric sequences were analyzed. Parameter ranges were the following: TR = 5000–6500 ms, TE = 100–393 ms, frequency = 192–512, phase = 75%–100%, section thickness = 1–5 mm, gap = 0–1 mm, FOV = 180–220 × 180 – 220 mm.

MR Imaging Analysis
Loes Score. Loes scores were assigned by neuroradiologists (P.A.C., A.L., K.B.) or neurologists (F.S.E., P.L.M, M.E.) with a minimum of 10 years of experience in cerebral adrenoleukodystrophy. Reviewers evaluated the images according to the Loes scoring system for abnormal signal hyperintensity or atrophy involving specific brain structures known to be involved in CALD.[15]

Lesion Pattern. The lesions were subdivided into 5 patterns according to their primary anatomic distribution as previously published:[5] 1) parieto-occipital lobe white matter and splenium of the corpus callosum; 2) frontal lobe white matter and genu of the corpus callosum; 3) frontopontine or corticospinal projection fibers; 4) cerebellar white matter; and 5) simultaneous parieto-occipital and frontal white matter involvement. Only patients in whom the MR imaging showed at least either pattern 1 or pattern 2 lesions were included in this study. Pattern 3 lesions were included only if they were continuous with a frontal or parieto-occipital lesion.

Lesion Volume Calculation. 3D Slicer software (https://www.slicer.org/) was used to create 3D label maps of the lesion by thresholding the label intensity to match the lesion intensity on the source axial T2, conventional FLAIR, or SPACE FLAIR sequence.[20,21] Lesions were then manually segmented section-by-section according to the subject-specific threshold value. Lesion

volumes were automatically calculated from the compiled section-by-section lesion segmentations (ie, area of the segmentation × section thickness) and are reported in milliliters (mL).

Interrater Reliability. To assess the interrater reliability of lesion segmentation between 2 raters (E.J.M. and A.L.) on the dataset, we calculated the intraclass correlation coefficient using a 2-way random-effects model with a consistency definition.

Statistical Analyses
Descriptive Statistics. All continuous variables are reported as median and range. Discrete variables are summarized as frequency (%). The Wilcoxon rank sum test was used to calculate differences in median values among skewed samples. Two-tailed P values < .05 were considered statistically significant. SPSS Statistics (Version 25; IBM) was used to perform this section of the analyses.

Population-Level Analysis of Cerebral Lesion Growth. Raw lesion volumes versus age were plotted. Lesion velocity was defined as a change in lesion volume between 2 sequential MRIs divided by the difference in months between scans and is reported in milliliters per month (mL/month). Lesion acceleration was defined as the difference in lesion velocities between the 2 scans divided by the time interval between scans and is reported in milliliters squared per month (mL/month^2). The Spearman rank correlation coefficient (ρ) was calculated to quantify the relationship between patient age and lesion velocity. Similarly, ρ was calculated to quantify the correlation between patient age and lesion acceleration.

Longitudinal Analysis of Individual Cerebral Lesion Growth. Time-zero (t_0) was defined as the date of first abnormal MR imaging. Fold change was calculated as the ratio of lesion volume on the nth MR imaging at t_n over the initial volume at t_0. Fold change in lesion volume for each patient was then plotted from t_0. Similarly, lesion velocity was plotted for each patient from t_0.

Lesion Growth Curve Modeling. The trajectory of each dependent variable was visualized using the nonparametric local-weighted regression curve with 95% confidence intervals and a smoothing parameter (α) of 0.75. Linear mixed-effects models were used to assess the difference in fold change and lesion velocity between the progressive and arrested phenotypes. Individual differences in baseline lesion volume and number/intervals of follow-up scans were addressed by assuming random intercepts for each patient. The model was fit with the main effects for phenotype group and time as well as an interaction term between phenotype and time:

$$\mathcal{Y}_{ij} = \beta_0 + \beta_1 group_i + \beta_2 time_{ij} + \beta_3 time_{ij} group_i + b_i + \varepsilon_{ij},$$

$$Group \begin{cases} 1 = progressive \\ 0 = arrested \end{cases}$$

$$\mathcal{Y}_{ij} progressive = \beta_0 + \beta_1 group_i + (\beta_2 + \beta_3) time_{ij} + b_i + \varepsilon_{ij},$$

$$\mathcal{Y}_{ij} arrested = \beta_0 + \beta_2 time_{ij} + b_i + \varepsilon_{ij},$$

where

$$\varepsilon \sim N(0, \sigma^2),$$

$$b_i \sim N(0, \sigma_b^2).$$

The regression coefficient (β), (95% CI), and P value are reported for each parameter estimate from the linear mixed effect (LME) models.

Restricted Data Set Analysis To understand the growth of small lesions from the earliest time points, we recalculated the nonparametric local-weighted regression curves and linear mixed-effects model analyses from a subset of the data meeting more restricted criteria: LS ≤ 2 on the first MRI with abnormal findings.[13] To illustrate lesion growth across time, we plotted raw volumes in months from t_0. This section of the analysis was performed by a biostatistician using R statistical and computing software, Version 3.5.3 (http://www.r-project.org/).

Data Availability
Following publication, any data not published within this article will be anonymized and shared by request from any qualified investigator.

RESULTS
Clinical Characteristics
Thirty-six patients met the inclusion criteria, resulting in 174 scans available for analysis. The total LS range across all MRIs with abnormal findings was 0.5–7.5. At the time of first MR imaging with abnormal findings, patients had a median age of 89.1 months (range, 47.2–449 months), with a median MR imaging LS of 1 (range, 0.5–7.5, Online Supplemental Data). Twenty-three patients were diagnosed with progressive CALD, accounting for 101 MRIs in the study. The median age at first MR imaging with abnormal findings was 71.8 months (range, 47.2–137.5 months), with an initial LS of 1 (range, 1–7). Thirteen patients underwent spontaneous arrest of cerebral disease, accounting for the remaining 73 MRIs. The median age at first MR imaging with abnormal findings was 185.6 months (range, 96.3–449 months), with an initial LS of 3 (range, 0.5–7.5). Patients with arrested CALD were older at the time of diagnosis (15.5 versus 6.0 years of age, P < .001).

A total of 23 patients met the more restricted criteria (LS ≤ 2 at first MR imaging with abnormal findings), resulting in 125 MRIs eligible for analysis. The LS range was 0.5–6. Nineteen patients with progressive CALD were identified, accounting for 90 MRIs. They were 69.5 months of age (range, 47.2–137.5 months) at diagnosis and had an initial LS of 1 (range, 1–2). Four of the patients with arrested CALD were identified, accounting for the remaining 35 MRIs. They were 136.3 months of age (range, 96.3–185.6 months) and had an initial LS of 1 (range, 0.5–2). Patients with arrested CALD were older at the time of diagnosis (11.4 versus 5.8 years of age, P = .003).

Lesion Segmentation
The intraclass correlation coefficient between lesion segmentations produced by 2 raters (E.J.M. and A.L.) on 39 MRIs from the data set was 0.874 (95% CI, 0.772–0.932; P < .001), indicating a
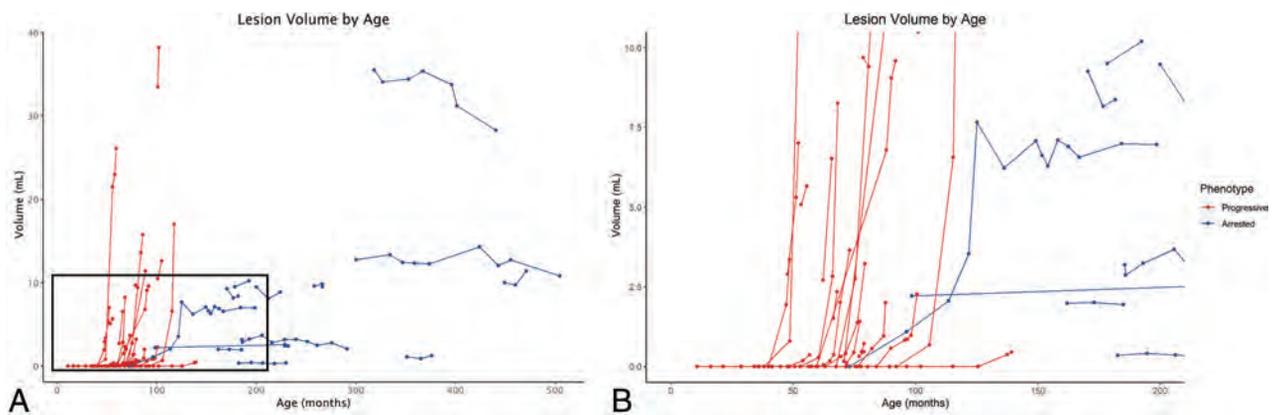
FIG 1. *A*, Patient-specific lesion trajectories of each cohort plotted against patient age. *B*, Zoomed view of outlined area in 1*A*): 0–18 years of age on the x-axis and 0- to 10-mL lesion volume on the y-axis. Note the exponential growth trajectories for patients with progressive CALD.
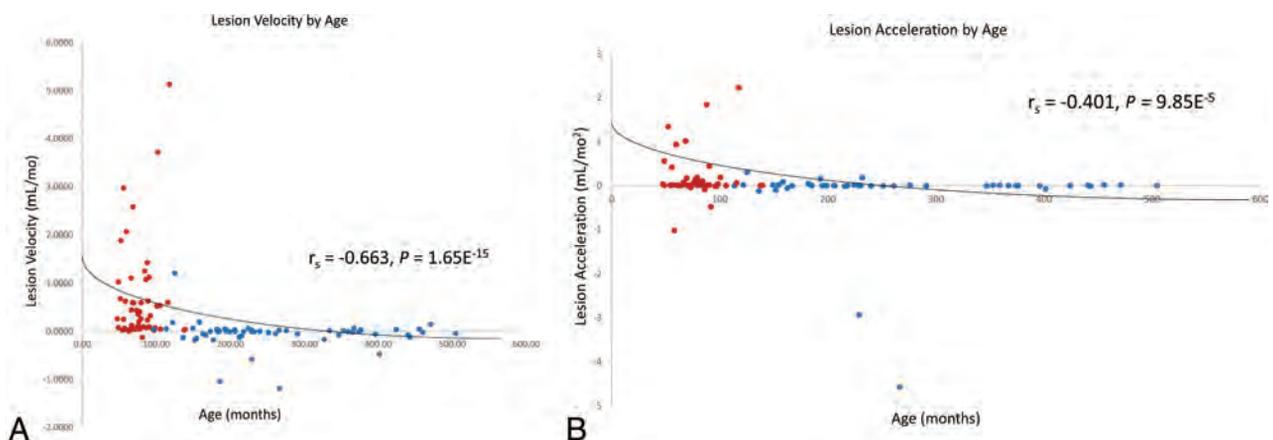

FIG 2. *A*, Lesion growth velocities over sequential MR imaging pairs and *B*, acceleration over sequential MR imaging triplets, stratified by age in patients with progressive CALD (red) and arrested CALD (blue). Presymptomatic lesion progression is inversely correlated with age ($P < .001$).

significant, good-to-excellent degree of reliability between segmentations produced by different raters.

Disease Progression is Inversely Correlated with Age

Patient-specific lesion volumes were plotted across time against patient age in months (Fig 1). There was a moderate-to-strong negative correlation between patient age and lesion velocity ($n = 112$, $\rho = -0.6631$, $P < .001$, Fig 2*A*). Accordingly, lesion acceleration inversely correlated with patient age ($n = 89$, $\rho = -0.4010$, $P < .001$, Fig 2*B*).

The Loes Score is Insensitive to Early Lesion Growth

There was a very strong correlation between lesion volume and LS ($n = 132$, $\rho = 0.8086$, $P < .001$). The median volume of a lesion with an LS of 1 was 0.51 mL (interquartile range, 0.36–1.46 mL; range, 0.05–8.26 mL). The median volume of a lesion with an LS of 2 was 3.18 mL (interquartile range, 2.45–3.67 mL; range, 0.68–11.41 mL). The median fold change on longitudinal MRIs with an LS of 1 (ie, $\Delta LS = 0$) was 1.83 ($n = 16$; interquartile range, 1.47–5.71; range, 1.11–29.14). Similarly, the median rate of lesion growth on longitudinal MRIs with an LS of 1 was 0.08 mL/month ($n = 30$; interquartile range, 0.05–0.25 mL/month; range, 0.01–2.58 mL/month; Fig 3).

Lesions in Progressive CALD Undergo Early, Sustained Acceleration

The 2-way interaction term between phenotype and time was significant for the rate of lesion growth (Table). For every 1 month, lesion velocity increased by 0.10 mL/month in patients with progressive CALD ($\beta = 0.10$ [95% CI 0.05–0.14], $P < .001$). There was a marginally larger effect in the restricted cohort ($\beta = 0.11$ [95% CI, 0.06–0.15], $P < .001$). The rate of lesion growth in patients with progressive CALD increased with time, consistent with exponential growth kinetics.

Lesion Growth Trajectories Diverge between Phenotypes before the Onset of Cerebral Symptoms

The 2-way interaction term between phenotype and time was significant for lesion growth measured by fold change (Table). Lesion volume increases by a factor of 2.5 from baseline every month in patients with progressive CALD (fold change $\beta = 2.49$ [95% CI, 2.10–2.89], $P < .001$). This model is limited to MRIs of asymptomatic patients, with a median initial LS of 1 (range, 1–7), and an LS range of 1–7 on subsequent MRIs (Fig 4).

The results of the phenotype and time interaction from the restricted cohort were similar to those of the full cohort (Table): There
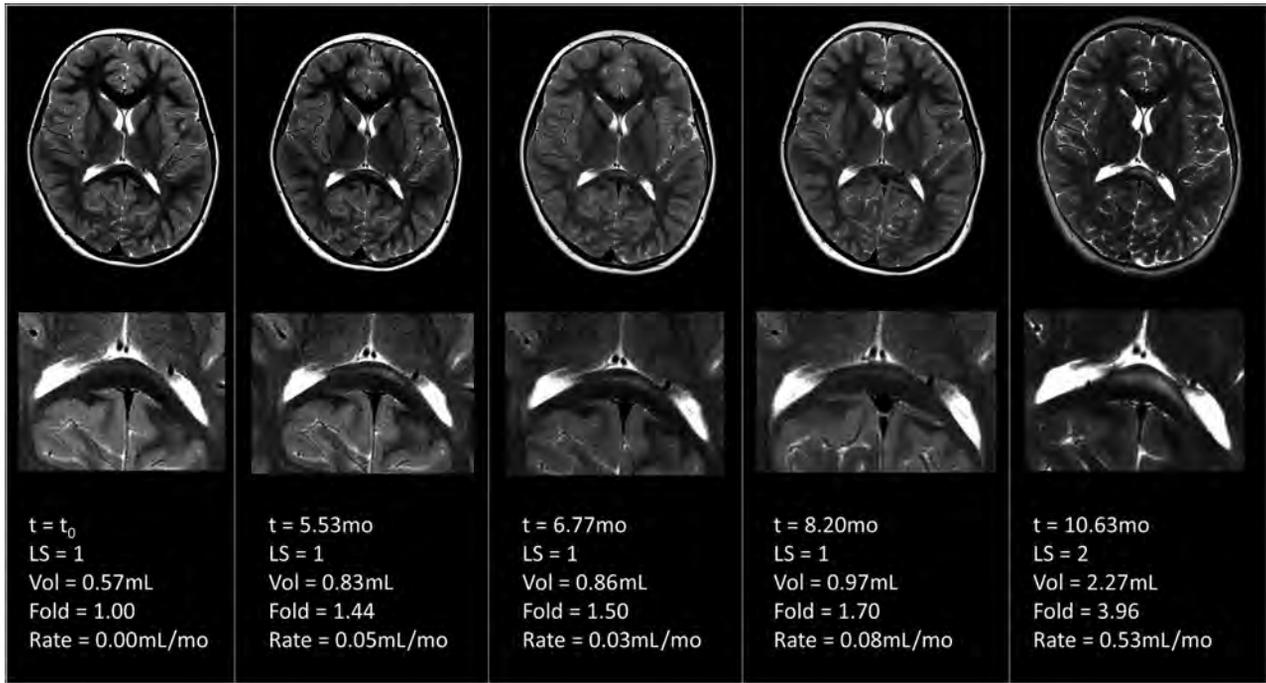
FIG 3. Serial MRIs demonstrating accelerated lesion growth without a change in the LS in patient 11.

Estimates produced by the linear mixed-effects models for fold change and lesion velocity in the full and restricted cohorts

	Fold Change			Lesion Velocity		
	Estimate	95% CI	*P*	Estimate	95% CI	*P*
Full cohort						
Time	0.01	−0.03–0.04	.768	0.00	−0.00–0.00	.989
Phenotype (progressive)	−3.00	−6.94–0.94	.142	0.340	−0.10–0.77	.136
Time*phenotype (progressive)	2.49	2.10–2.89	<.001	0.100	0.05–0.14	<.001
Restricted cohort						
Time	0.01	−0.06–0.08	.752	0.00	−0.01–0.01	.899
Phenotype (progressive)	−3.43	−11.08–4.21	.386	0.15	−0.47–0.78	.632
Time*phenotype (progressive)	2.54	2.04–3.04	<.001	0.11	0.06–0.15	<.001

Note:—Asterisk indicates the interaction between phenotype and time in the linear mixed effects model.

was a marginally larger effect for fold change in lesion volume across time ($\beta = 2.54$ [95% CI, 2.04–3.04], $P < .001$). This model is limited to MRIs of asymptomatic patients with a median initial LS of 1 (range, 1–2) and an LS range of 1–6 on subsequent MRIs (Fig 4).

In both linear mixed-effects models, the main effect term for phenotype was not significant, indicating that the 2 groups did not differ at onset. This is expected because baseline MRIs with abnormal findings have a common corrected value for both phenotypes (ie, fold change = 1, lesion velocity = 0 mL/month). The median rate of lesion growth directly following the first MR imaging with abnormal findings in the progressive group was significantly faster than that in the arrested group (0.448 mL/month versus 0.003 mL/month, $P = .001$).

To approximate real-time Newborn Screening MR imaging surveillance of early CALD, we plotted raw volumes from t_0 in the restricted cohort of patients (Fig 5).

DISCUSSION

Our study aimed to explore cerebral lesion growth as a potential early biomarker for progressive CALD. The implementation of the adrenoleukodystrophy Newborn Screening has provided the opportunity to monitor for new cerebral lesions from birth.[14,22] The early, accurate diagnosis of progressive disease is paramount because HSCT is most successful when initiated in the window before the onset of neurologic symptoms.[7,9,10,23] This may be equally important for other treatment approaches as well; a recent trial of autologous hematopoietic stem cell gene therapy has produced encouraging results in boys with early CALD.[19] Because symptoms accelerate with time in patients with progressive CALD,[3,5,6] we hypothesized that disease acceleration could be radiographically captured before symptom onset. We completed the primary aim by modeling the 3D expansion of early cerebral lesions from longitudinal MRIs obtained in presymptomatic patients with progressive and arrested CALD. Our results demonstrate that disease progression is inversely correlated with age, early lesions can grow while appearing radiographically stable, lesions undergo sustained acceleration in early-stage progressive CALD, and growth trajectories diverge between phenotypes in the presymptomatic time period.
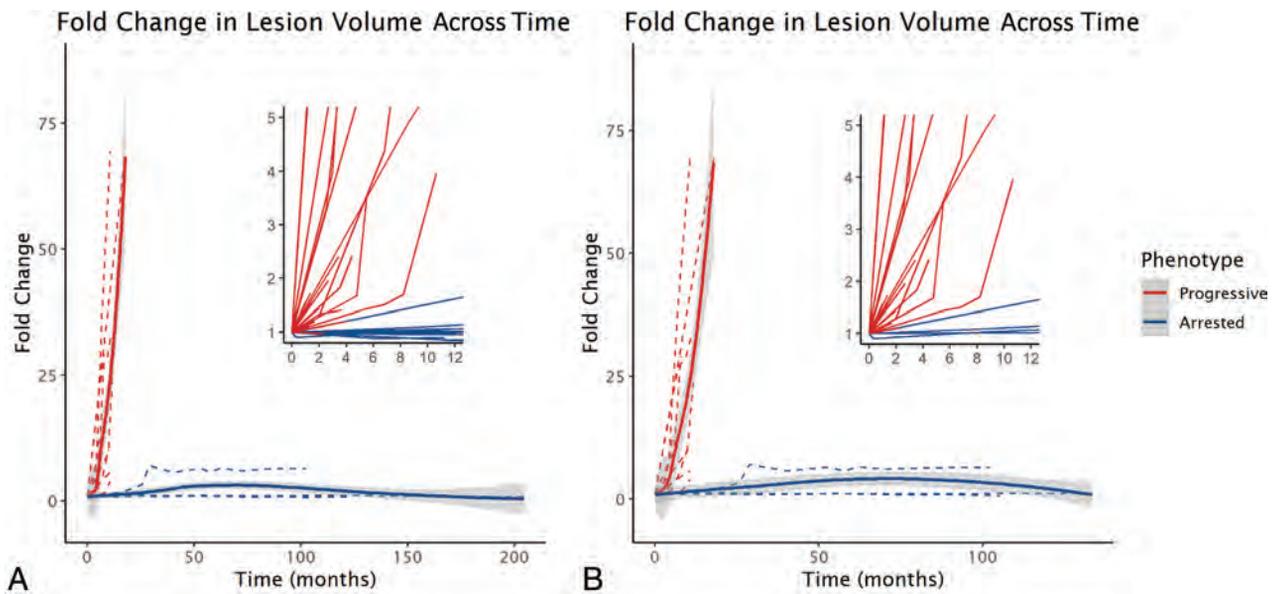
FIG 4. *A*, Lesion growth trajectories from the time of the first abnormal MR imaging finding in presymptomatic patients with progressive (*n* = 23, red) versus arrested (*n* = 13, blue) CALD. Lesion volume increased 2.5× per month (fold change β = 2.49 [95% CI, 2.10–2.89], *P* < .001) in the patients with progressive CALD. *B*, In the cohort of patients with an initial LS ≤ 2 (restricted cohort; *n* = 19 progressive, *n* = 4 arrested), lesion growth was similar (β = 2.54 [95% CI, 2.04–3.04], *P* < .001) for the patients with progressive CALD.
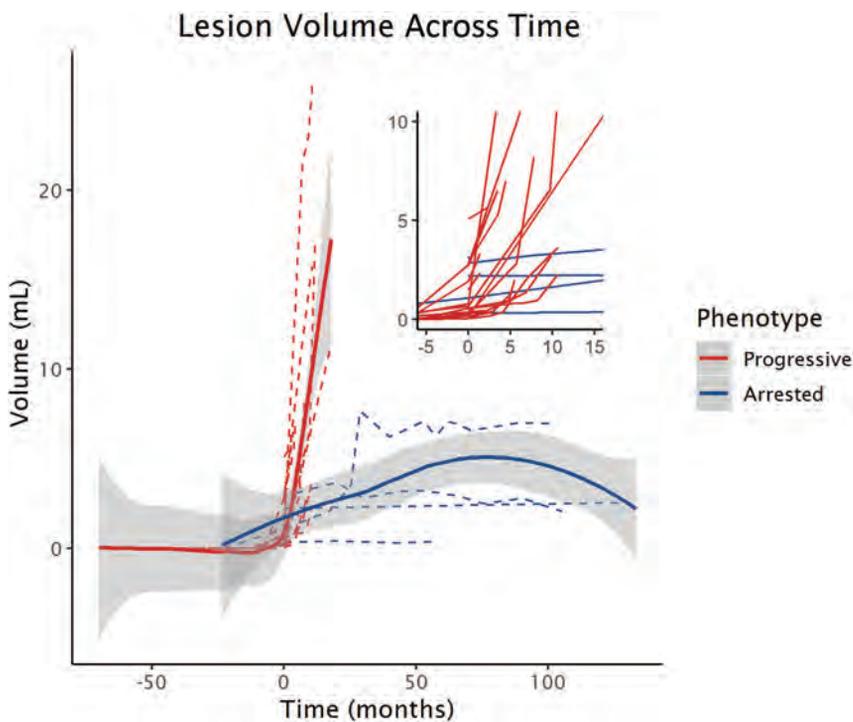

FIG 5. Nonparametric local-weighted regression curves demonstrating the early lesion growth trajectories between phenotypes measured by raw lesion volume across time in the cohort of patients with an initial LS ≤ 2.

To further understand lesion expansion from the earliest time points, thus approximating what may be found by Newborn Screening MR imaging surveillance, we repeated the analysis on the subset of longitudinal MRIs that had an LS ≤ 2 at diagnosis. In this restricted cohort, patients with both progressive and arrested CALD were diagnosed at a younger age with smaller lesions. Again, all measures of growth differed between phenotypes, indicating that trajectories diverge very early in disease and may be detected by calculating lesion velocity and fold change from the MR imaging following diagnosis.

Suspicion for progressive CALD should be highest in younger patients whose lesions follow an exponential growth curve. Our recent work demonstrated the inverse relationship between rates of demyelinating lesion growth and age in a cohort of asymptomatic patients with CALD.[4] The current study validates our previous results in a larger, international cohort. Biologically, our findings are consistent with the timing of dynamic myelination into late adolescence and adulthood[24] and the suggestion that following transplantation, more developed white matter tracts in adolescence protect against further neurocognitive decline compared with children with CALD undergoing the procedure.[17] Stated another way, younger children with younger myelin may be more vulnerable to cerebral disease progression than older patients with more developed myelin.

The proposed model has identified features of early lesion growth that have the potential to inform the timely selection of patients appropriate for rescue therapy. This feature is pertinent

to recent observations that altered cerebral microvascular perfusion in perilesional normal-appearing white matter precedes lesion progression.[16] Early patient identification would help mitigate risk and maximize neurologic outcomes, decrease the number of MRIs required to identify the correct patients for treatment, reduce patient exposures to anesthesia and contrast agents, and widen the presymptomatic treatment window. Patients treated expeditiously with an LS ≤ 2 have demonstrated superior posttransplantation neurocognitive and neuropsychiatric outcomes and less posttransplantation disease progression on MRI.[13] Lesion volumetry may also prove to be a more sensitive method to monitor for disease activity posttreatment. In terms of feasibility, the biomarker itself is based on the acquisition of standard T2-weighted images and, therefore, makes the analysis paradigm possible across all platforms. Conceptually, an analysis of growth kinetics can be applied to other leukodystrophies, as it has been in other white matter diseases.[25,26]

Longitudinal imaging within subjects provides strength to the analysis. The statistical analyses accounted for the correlation between repeat patient measures, thus allowing each patient to act as an internal control for his or her lesion-specific changes across time. Sequential pair and triplet data points are required to calculate lesion growth velocity and acceleration, respectively, which improve the accuracy of the study conclusions. However, important limitations apply. First, we analyzed retrospective data and are limited by when subjects came to medical attention. The result is lesion onset (the transition from a normal MR imaging finding with a LS = 0 to an abnormal MR imaging finding with a LS ≥ 0.5) is not available for most of the arrested CALD cohort. This issue limits the comparison with the patients with progressive CALD for whom lesion onset was factored into the model. Only a standardized, prospective imaging surveillance program will reveal the true growth trajectory and behavior of cerebral adrenoleukodystrophy lesions across time. Second, there were an unequal number of scans per patient and unequal time between repeat measures. This intra- and intersubject heterogeneity precluded the use of more preferable methods of longitudinal mathematic modeling, including fitting more complex, nonlinear models. Third, the number of subjects, specifically in the arrested cohort, is limited. As with many rare diseases, the generalizability of the study conclusions is limited due to the number of available cases. Fourth, while most scans were sequentially obtained on the same MR imaging scanner for each subject, 5 subjects had their 1–2 MRIs obtained on a 1.5T unit. This may have introduced variability in the interpretation of disease onset and the degree of white matter involvement. If present, the variability is small: Sicotte et al[27] demonstrated a 10% average difference in total white matter lesion volume between 1.5T and 3T scanners. A similar effect may have been introduced by performing some measurements on the axial T2-weighted sequence versus 3D T2 FLAIR sequences. Fifth, our analysis did not include isolated pattern 3, 4, or 5 lesions because they are exceedingly rare in this age group.[5] Sixth, manual lesion segmentations are time- and resource-intensive, making integration into clinical workflow difficult. Adaptation of pre-existing volume measurement methodologies or automated deep learning approaches to lesion segmentation may support workflow integration.[28,29] Finally, the proposed models require validation before entertaining their use in clinical practice. While our models provide insight into early lesion behavior, we emphasize caution when interpreting the results in the context of clinical practice. If validated, our results lay the foundation for diagnostic criteria aimed at identifying early-stage progressive CALD in a fashion similar to that of other progressive neurologic diseases.[30]

CONCLUSIONS

The addition of adrenoleukodystrophy to Newborn Screening has provided clinicians a novel, presymptomatic approach to disease. Understanding early lesion evolution is paramount for the accurate, early diagnosis of patients with progressive CALD. Our study indicates that measuring the volumetric changes in newly developing cerebral lesions across time can distinguish CALD phenotypes before symptom onset. When factored into the overall clinical presentation of a patient with a new brain lesion, our method may aid in the accurate prediction of patients eligible for therapy.

REFERENCES

1. Moser HW. **Adrenoleukodystrophy: phenotype, genetics, pathogenesis and therapy.** *Brain* 1997;120:1485–1508 CrossRef
2. Huffnagel IC, Laheji FK, Aziz-Bose R, et al. **The natural history of adrenal insufficiency in X-linked adrenoleukodystrophy: an international collaboration.** *J Clin Endocrinol Metab* 2019;104:118–26 CrossRef Medline
3. Moser HW, Loes DJ, Melhem ER, et al. **X-linked adrenoleukodystrophy: overview and prognosis as a function of age and brain magnetic resonance imaging abnormality—a study involving 372 patients.** *Neuropediatrics* 2000;31:227–39 CrossRef Medline

4. Liberato AP, Mallack EJ, Aziz-Bose R, et al. **MRI brain lesions in asymptomatic boys with X-linked adrenoleukodystrophy.** *Neurology* 2019;92:e1698–1708 CrossRef Medline

5. Loes DJ, Fatemi A, Melhem ER, et al. **Analysis of MRI patterns aids prediction of progression in X-linked adrenoleukodystrophy.** *Neurology* 2003;61:369–74 CrossRef Medline

6. Melhem ER, Loes DJ, Georgiades CS, et al. **X-linked adrenoleukodystrophy: the role of contrast-enhanced MR imaging in predicting disease progression.** *AJNR Am J Neuroradiol* 2000;21:839–44 Medline

7. Raymond GV, Aubourg P, Paker A, et al. **Survival and functional outcomes in boys with cerebral adrenoleukodystrophy with and without hematopoietic stem cell transplantation.** *Biol Blood Marrow Transplant* 2019;25:538–48 CrossRef Medline

8. Mallack EJ, van de Stadt S, Caruso PA, et al. **Clinical and radiographic course of arrested cerebral adrenoleukodystrophy.** *Neurology* 2020;94:e2499–2507 CrossRef Medline

9. Peters C, Charnas LR, DeFor T, et al. **Cerebral X-linked adrenoleukodystrophy: the University of Minnesota hematopoietic cell transplantation experience from 1991 to 2004.** *Blood* 2004;104:2153 CrossRef

10. Miller WP, Rothman SM, Nascene D, et al. **Outcomes after allogeneic hematopoietic cell transplantation for childhood cerebral adrenoleukodystrophy: the largest single-institution cohort report.** *Blood* 2011;118:1971–78 CrossRef Medline

11. Miller W, Bjoraker K, Charnas L, et al. **Hematopoietic cell transplantation (HCT) for adrenoleukodystrophy: ten-year experience with 62 patients.** *Ann Neurol* 2010;68:S121 CrossRef

12. Mahmood A, Raymond GV, Dubey P, et al. **Survival analysis of haematopoietic cell transplantation for childhood cerebral X-linked adrenoleukodystrophy: a comparison study.** *Lancet Neurol* 2007;6:687–92 CrossRef Medline

13. Pierpont EI, Nascene DR, Shanley R, et al. **Neurocognitive benchmarks following transplant for emerging cerebral adrenoleukodystrophy.** *Neurology* 2020;95:e591–600 CrossRef Medline

14. Vogel BH, Bradley SE, Adams DJ, et al. **Newborn screening for X-linked adrenoleukodystrophy in New York State: diagnostic protocol, surveillance protocol and treatment guidelines.** *Mol Genet Metab* 2015;114:599–603 CrossRef Medline

15. Loes DJ, Hite S, Moser H, et al. **Adrenoleukodystrophy: a scoring method for brain MR observations.** *AJNR Am J Neuroradiol* 1994;15:1761–66 Medline

16. Musolino PL, Rapalino O, Caruso P, et al. **Hypoperfusion predicts lesion progression in cerebral X-linked adrenoleukodystrophy.** *Brain* 2012;135:2676–83 CrossRef Medline

17. Pierpont EI, Eisengart JB, Shanley R, et al. **Neurocognitive trajectory of boys who received a hematopoietic stem cell transplant at an early stage of childhood cerebral adrenoleukodystrophy.** *JAMA Neurol* 2017;74:710–17 CrossRef Medline

18. Moser AB, Kreiter N, Bezman L, et al. **Plasma very long chain fatty acids in 3,000 peroxisome disease patients and 29,000 controls.** *Ann Neurol* 1999;45:100–10 CrossRef Medline

19. Eichler F, Duncan C, Musolino PL, et al. **Hematopoietic stem-cell gene therapy for cerebral adrenoleukodystrophy.** *N Engl J Med* 2017;377:1630–38 CrossRef Medline

20. Kikinis R, Pieper SD, Vosburgh Kg 3D. **Slicer a platform for subject-specific image analysis, visualization, and clinical support.** In: Jolesz FA, ed. *Intraoperative Imaging and Image-Guided Therapy.* Springer-Verlag New York; 2014:277–89

21. Egger J, Kapur T, Fedorov A, et al. **GBM volumetry using the 3D slicer medical image computing platform.** *Sci Rep* 2013;3:1364–67 CrossRef Medline

22. Mallack EJ, Turk BR, Yan H, et al. **MRI surveillance of boys with X-linked adrenoleukodystrophy identified by newborn screening: meta-analysis and consensus guidelines.** *J Inherit Metab Dis* 2021;44:728–39 CrossRef Medline

23. Shapiro E, Krivit W, Lockman L, et al. **Long-term effect of bone-marrow transplantation for childhood-onset cerebral X-linked adrenoleukodystrophy.** *Lancet* 2000;356:713–18 CrossRef Medline

24. Lebel C, Beaulieu C. **Longitudinal development of human brain wiring continues from childhood into adulthood.** *J Neurosci* 2011;31:10937–47 CrossRef Medline

25. Aubert-Broche B, Fonov V, Narayanan S, et al. Canadian Pediatric Demyelinating Disease Network. **Onset of multiple sclerosis before adulthood leads to failure of age-expected brain growth.** *Neurology* 2014;83:2140–46 CrossRef Medline

26. Tillema JM, Derks MG, Pouwels PJ, et al. **Volumetric MRI data correlate to disease severity in metachromatic leukodystrophy.** *Ann Clin Transl Neurol* 2015;2:932–40 CrossRef Medline

27. Sicotte NL, Voskuhl RR, Bouvier S, et al. **Comparison of multiple sclerosis lesions at 1.5 and 3.0 Tesla.** *Invest Radiol* 2003;38:423–27 CrossRef Medline

28. Kothari RU, Brott T, Broderick JP, et al. **The ABCs of measuring intracerebral hemorrhage volumes.** *Stroke* 1996;27:1304–05 CrossRef Medline

29. Chang K, Beers AL, Bai HX, et al. **Automatic assessment of glioma burden: a deep learning algorithm for fully automated volumetric and bidimensional measurement.** *Neuro Oncol* 2019;21:1412–22 CrossRef Medline

30. Wattjes MP, Rovira À, Miller D, et al. MAGNIMS study group. **Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis: establishing disease prognosis and monitoring patients.** *Nat Rev Neurol* 2015;11:597–606 CrossRef Medline

COVID-19-Associated Myelitis Involving the Dorsal and Lateral White Matter Tracts: A Case Series and Review of the Literature

H.Y. Huang, L.M. Shah, J.S. McNally, T. Sant, T.A. Hutchins, E.D. Goldstein, and M.E. Peckham

ABSTRACT

SUMMARY: Coronavirus disease 2019 (COVID-19) myelitis is a rare condition, most commonly presenting with nonenhancing central expansile cord T2 signal changes. A single case report has also described longitudinal involvement of the dorsal columns. We present 5 cases of COVID-19-associated myelitis with tract-specific involvement of the dorsal and lateral columns and discuss potential pathophysiologic pathways for this unique pattern.

ABBREVIATION: COVID-19 = coronavirus disease 2019

Coronavirus disease 2019 (COVID-19) is primarily known as a respiratory illness; however, neurologic responses to this infection have been increasingly described, including infectious encephalopathy, meningoencephalitis, Guillain-Barre syndrome, and stroke.[1] Myelitis has been reported multiple times as a post-infectious inflammatory reaction from COVID-19 infection. Previously documented cases have predominantly described central longitudinal T2 changes without corresponding enhancement,[2-12] with 2 cases of autoimmune myelitis (antimyelin oligodendrocyte glycoprotein-spectrum disorder and aquaporin-4 neuromyelitis optica) also found to be associated with COVID-19 infection.[13,14] At the more extreme ends of the spectrum, 1 report described T2-bright and centrally necrotic enhancing lesions,[15] with another case presenting with clinical symptoms of transverse myelitis with paraplegia below the T10 level but with a normal appearance of the spinal cord on clinical MR imaging.[16] A few cases have noted more tract-specific disease, with 1 case demonstrating ventral horn–predominant T2 hyperintensity and acute flaccid myelitis[17] and 1 interesting case demonstrating dorsal column–predominant T2 signal abnormality with progressive numbness in the feet and hands.[18]

Although most reports of COVID-19-associated myelitis have noted central-predominant T2 signal change, we describe 5 cases demonstrating lateral and dorsal column–specific disease. All patients were seen at our single academic institution, and were the only 5 cases of COVID-19-associated myelitis diagnosed and treated since the beginning of the pandemic in March 2020.

Case 1

Presentation. A 62-year-old woman without prior neurologic disease presented 10 days after a COVID-19 diagnosis with a 1-day history of rapidly ascending numbness with gait impairment and fecal incontinence. She previously experienced mild COVID-19 symptoms, including anosmia and cough, which had resolved several days before her hospitalization. Her examination findings were notable for diffuse hyperreflexia and reduced vibratory and pinprick sensations distally with sensory ataxia.

Laboratory Findings. Serum laboratory evaluation demonstrated elevated inflammatory markers including erythrocyte sedimentation rate and C-reactive protein, a ganglioside panel positive for ganglioside-monosialic acid antibodies, and CSF with elevated lymphocytes. CSF protein and glucose levels were normal. Further laboratory studies were unremarkable, including vitamins B3, B6, B12, copper, methylmalonic acid, vitamin E, and viral panels (Online Supplemental Data).

Imaging Findings. Thoracic spine MR imaging showed patchy areas of T2 hyperintensity as well as patchy areas of enhancement throughout the thoracic cord, predominantly involving the lateral cord and dorsal columns (Fig 1). There was also ventral and

FIG 1. Axial T2 (*A*), and T1 postcontrast (*B*) images demonstrate T2 hyperintensity involving the posterior columns and patchy enhancement predominantly in the posterior and lateral columns (*arrow* pointing to posterior column in *A*, and *arrow* to lateral column in *B*). Sagittal T2 (*C*) and postcontrast T1 (*D*) images demonstrate corresponding signal changes (*arrows*).

dorsal nerve root enhancement along the cauda equina. MR imaging findings of the brain were unremarkable, with no enhancing lesions or other pathology. Electromyography and nerve conduction studies were obtained 12 days after the onset of symptoms, which did not show electrodiagnostic evidence of a peripheral polyneuropathy.

Management. Treatment with plasma exchange and oral prednisone was initiated with improvement in her clinical symptoms. She was discharged with a month-long oral prednisone taper. At the 4-month follow-up visit, her previous clinical symptoms had largely resolved.

Case 2

Presentation. A 79-year-old man with a history of hypertension, hyperlipidemia, prior back surgery, and chronic low back pain requiring a walker for ambulation at baseline presented with a 2-month history of worsening lower extremity weakness and fecal incontinence. He had been diagnosed with COVID-19 infection approximately 2 weeks before symptom onset. His examination was notable for paraparesis with minimal right-ankle dorsiflexion and plantar flexion, areflexia, and increased tone in his lower extremities, as well as reduced vibration and proprioception in a length-dependent pattern.

Laboratory Findings. Serum laboratory evaluation demonstrated elevated inflammatory markers (erythrocyte sedimentation rate and C-reactive protein), but CSF studies showed a normal white blood cell count and protein and glucose levels. Further laboratory studies showed normal vitamins B1, B12, vitamin E, folate, copper, monoclonal protein, and methylmalonic acid levels. He was found to be heterozygous for the methylenetetrahydrofolate reductase (*MTHFR*) genes (*C665T* and *A1286C*) (Online Supplemental Data).

Imaging Findings. MR imaging demonstrated ventral medullary and left-lateral/dorsal-lateral cord signal abnormality from C2

through C3 along with patchy lateral and dorsal column involvement of the mid and lower cervical cord and thoracic cord with associated diffusion restriction (Fig 2). There was no associated enhancement. Electromyography and nerve conduction studies showed signs of active denervation in the proximal-greater-than-distal lower extremities.

Management. He was treated with 5 days of intravenous immunoglobulin and 5 sessions of plasma exchange, unfortunately without notable improvement in symptoms. He was then discharged to inpatient rehabilitation with a rehabilitation course complicated by urinary tract infection and aspiration pneumonitis. After extensive inpatient rehabilitation, he demonstrated improvement in his functional status and was discharged home.

Case 3

Presentation. A 40-year-old woman with a medical history including kidney transplant secondary to focal segmental glomerulosclerosis on chronic immunosuppression, latent tuberculosis, and diabetes mellitus type 2 presented with new and rapid onset of low back pain, lower extremity weakness, and sensory changes. She developed these symptoms within a few days of testing positive for COVID-19. Her examination findings were notable for hyperreflexia in the lower extremities and decreased strength in the proximal-greater-than-distal lower extremities with preserved sensation.

Laboratory Findings. Serum laboratory evaluation demonstrated an elevated inflammatory marker (erythrocyte sedimentation rate with normal C-reactive protein levels). CSF studies showed a normal white blood cell count and protein and glucose levels, but 11 unique oligoclonal bands. Findings for evaluation of serum antibodies including myelin oligodendrocyte glycoprotein aquaporin-4 antibody immunoglobulin G for demyelinating disease were negative. Other laboratory studies included normal vitamin B12, B1, B6, D, copper, and folate levels (Online Supplemental Data). In the setting of her immunosuppressive
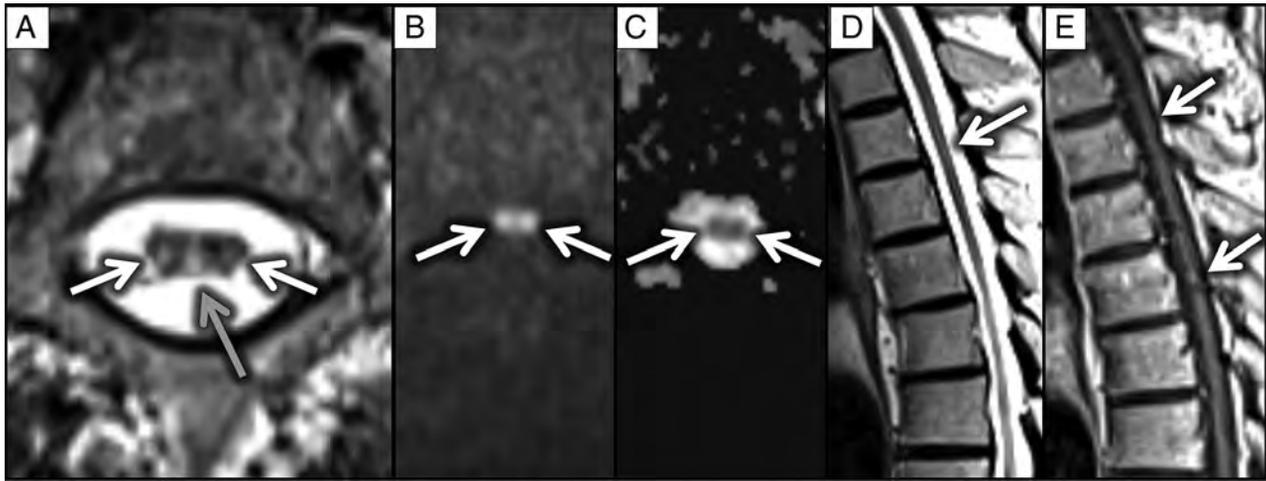
FIG 2. Axial T2 (*A*) and diffusion-weighted (*B*) images demonstrate T2 hyperintensity involving the posterior (*gray arrow*) and lateral columns (*white arrows*) with corresponding DWI hyperintensity and ADC hypointensity (*C*) confirming restriction (*white arrows*). Findings were less conspicuous on sagittal T2 (*D*) and postcontrast T1 (*E*) images, with only faint signal changes present (*white arrows*).
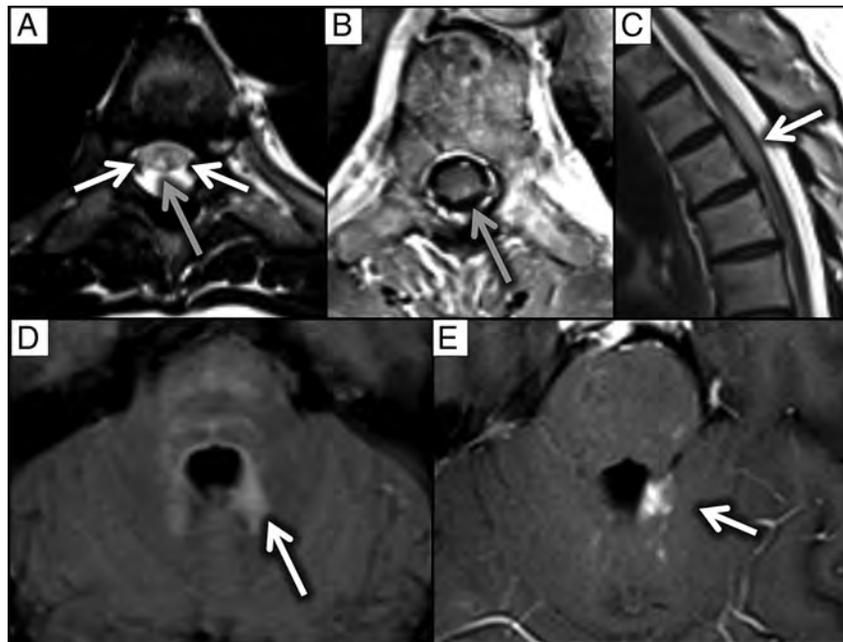
FIG 3. Axial T2 (*A*) and T1 postcontrast (*B*) images demonstrate T2 hyperintensity involving the posterior columns (*gray arrow*), with faint involvement of the lateral columns (*white arrows*, *A*). There was corresponding scattered patchy enhancement in these regions (*gray arrow*, *B*). T2 hyperintensity and expansion of the upper thoracic cord was well-demonstrated on sagittal views (*white arrow*, *C*). Brain MR imaging shows T2/FLAIR hyperintensity and corresponding enhancement surrounding the fourth ventricle (*white arrows*, *D* and *E*).

state, extensive infectious work-up was performed, which did not show evidence of active infection. Electromyography and nerve conduction studies showed no evidence of neuropathy or myopathy.

Imaging Findings. MR imaging of the brain showed T2 and FLAIR hyperintensities with enhancement in the midbrain, pons, and medulla. MR imaging of the spine showed multifocal T2/

STIR hyperintensities and enhancement predominantly involving the dorsal and lateral columns within the cervical and thoracic spinal cord (Fig 3). Intracranial MR vessel wall imaging showed no arterial narrowing or abnormal enhancement along the vessel walls.

Management. She was given a 3-day course of high-dose intravenous methylprednisolone with subjective improvement in her symptoms. She was discharged home with an increased dose of oral prednisone along with her other immunosuppressive medications. Unfortunately, during the course of several months, she developed worsening symptoms, including increased weakness, gait instability, severe constipation, and sensory changes. Repeat MR imaging demonstrated worsened T2-hyperintense enhancing lesions at the level of T3–T4 as well as new enhancing brain lesions with enhancement most pronounced in the basal ganglia, pons, and cerebellum. Further investigation regarding other etiologies of her encephalomyelitis, including unremarkable brain biopsy findings, have been unrevealing thus far, and her working diagnosis remains COVID-19-associated encephalomyelitis.

Case 4

Presentation. A 62-year-old man with a history of end-stage renal disease on hemodialysis, diabetes mellitus type 2, hypertension, and hyperlipidemia presented with back pain, fever, cough,
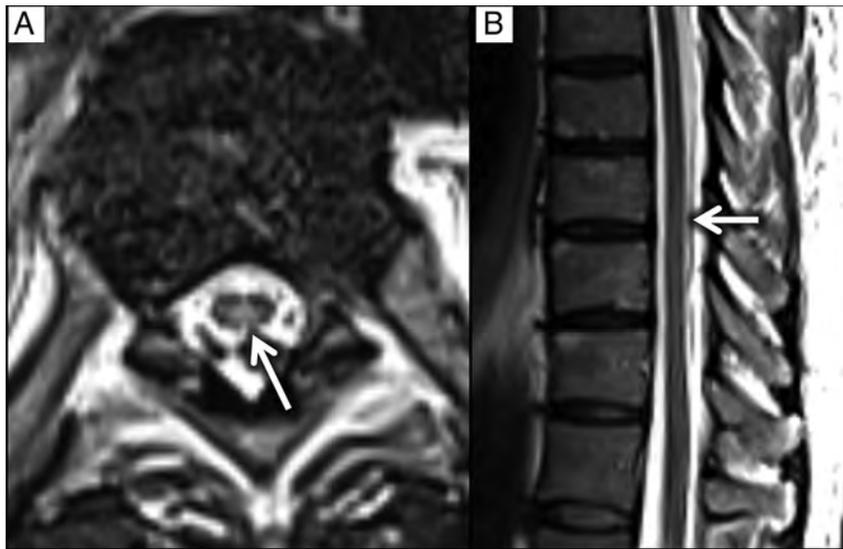
FIG 4. Axial T2 image through the thoracic cord demonstrates T2 hyperintensity involving the posterior columns (*white arrow*). Subtle hyperintensity could be seen in this region on sagittal T2 images (*white arrow*).

and chills. Symptoms started 2 days after receiving the COVID-19 vaccine, and he subsequently tested positive for COVID-19 infection. His examination was notable for mildly decreased strength in the bilateral upper extremities with intact strength in the lower extremities without noted sensory deficits.

Laboratory Findings. Serum evaluation demonstrated elevated inflammatory markers including erythrocyte sedimentation rate and C-reactive protein. Vitamin B12 and folate levels were normal (Online Supplemental Data).

Imaging Findings. MR imaging of the spine showed intramedullary T2 hyperintensities throughout the cervical and thoracic cord, predominantly involving the lateral and dorsal columns (Fig 4).

Management. He was discharged with medications for pain management, but no immunomodulating therapies were prescribed.

Case 5

Presentation. A 64-year-old man with no prior neurologic history on zinc supplementation during the COVID-19 pandemic, presented with a 3-month history of lower extremity paresthesia and gait imbalance, which had a rapid onset 2 days after diagnosis of COVID-19 infection. Physical examination findings were notable for lower extremity hyperreflexia and reduced vibratory and proprioceptive sensations in a length-dependent pattern with sensory ataxia and preservation of strength.

Laboratory Findings. Serum laboratory evaluation demonstrated an elevated zinc level and an undetectable copper level (Online Supplemental Data).

Imaging Findings. Findings of MR imaging of the brain were unremarkable, with no enhancing lesions or other pathology. MR imaging of the spine showed diffuse T2-hyperintense signal

abnormality predominantly involving the dorsal columns throughout the length of the spinal cord (Fig 5). Electromyography and nerve conduction studies demonstrated mild axonal neuropathy.

Management. He underwent oral copper supplementation. On repeat evaluation 3 months later, he reported no subjective improvement in his symptoms. Testing for serum copper levels was repeated at that time with findings within normal limits.

DISCUSSION

Hypothesized mechanisms for neurologic sequelae of COVID-19 have been ascribed to a systemic inflammatory response or sequela of critical illness.[19] While multiple reports have described associated nonspecific central cord edema, few have demonstrated tract-specific disease in the setting of this virus.[18]

Lateral and dorsal column involvement is relatively atypical in viral myelitis, which usually demonstrates a propensity for central gray matter regions of the cord.[20] While infectious posterior column involvement has been seen in HIV,[21] and posterior and lateral column involvement in human T-cell lymphotropic virus type 1-associated myelopathy,[22] involvement of these regions is more commonly seen in the noninfectious setting of subacute combined degeneration.[20] Subacute combined degeneration is caused by a disturbance in the methylation pathway,[23] which can stem from metabolic causes such as vitamin B12[24] and copper deficiency, vitamin E deficiency,[25,26] and toxic causes such as excess nitrous oxide (secondarily causes qualitative B12 deficiency),[27] and intrathecal methotrexate.[28] Metabolic and nutritional factors have been found to have key roles in supporting the immune system, specifically in the setting of COVID-19,[29] with some evidence supporting vitamin B12 supplementation in the fight against this pathogen.[30]

It can only be speculated that COVID-19 may affect the methylation cycle, which, in turn, may expose deficiencies in the immune system that are supported by this pathway. This finding may be supported by dorsal and sometimes lateral column involvement in these cases as well as in the 2 cases that demonstrated an underlying propensity for methylation cycle abnormalities. In case 2, the patient was heterozygous for mutation of the *MTHFR* gene, which has been found to strongly predispose to subacute combined degeneration.[31] Most interesting, the *MTHFR* C677T polymorphism has also previously been found to predispose to a more severe course of COVID-19, with vitamin supplementation recommended in individuals with this mutation.[32] In case 5, the patient had an underlying copper deficiency, likely resulting from zinc toxicity from supplementation. Although this may be considered a confounder for the patient's myelitis findings, clinical symptoms of rapid onset are in line
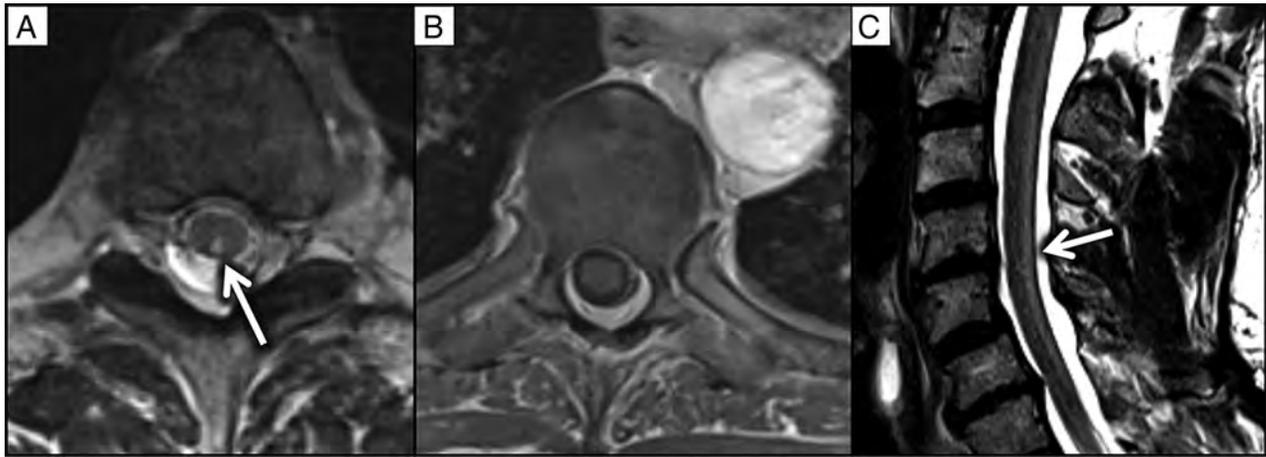
FIG 5. Axial T2 (*A*) and T1 postcontrast (*B*) images through the thoracic cord demonstrate T2 hyperintensity involving the posterior columns (*white arrow*), with no corresponding enhancement. A sagittal T2 image in the cervical spine (*C*) demonstrates hyperintensity along the dorsal aspect of the cord (*white arrow*).

with the other reported COVID-19-associated cases, and these symptoms did not improve with reversal of the copper deficiency.

Most interesting, none of the 5 COVID-19-associated cases of myelitis diagnosed and treated at our institution demonstrated the predominantly central expansile T2-signal hyperintensity reported most commonly throughout the literature.[2-12] This finding may be related to regional differences in viral strains or environmental factors. Additionally, it may be that this tract-specific appearance, only rarely described in the literature, is not always clinically attributed to the patient's COVID-19 status.

Work-up and management of these 5 patients varied widely. These differences can be attributed to inpatient-versus-outpatient evaluation as well as different primary services overseeing their care. Ideally, further investigation regarding the mechanism and etiology of COVID-19 myelitis would provide a more comprehensive and consistent evaluation of these patients.

CONCLUSIONS

Imaging features of dorsal and lateral white matter tract viral/postviral involvement further support a vitamin deficiency/nutritional component to the severity of COVID-19 symptoms. A link between the pathophysiology of COVID-19 and the methylation pathway is possible.

REFERENCES

1. Whittaker A, Anson M, Harky A. **Neurological manifestations of COVID-19: a systematic review and current update.** *Acta Neurol Scand* 2020;142:14–22 CrossRef Medline
2. Munz M, Wessendorf S, Koretsis G, et al. **Acute transverse myelitis after COVID-19 pneumonia.** *J Neurol* 2020;267:2196–97 CrossRef Medline
3. Chow CC, Magnussen J, Ip J, et al. **Acute transverse myelitis in COVID-19 infection.** *BMJ Case Rep* 2020;13:e236720 CrossRef Medline
4. Chakraborty U, Chandra A, Ray AK, et al. **COVID-19-associated acute transverse myelitis: a rare entity.** *BMJ Case Rep* 2020;13:e238668 CrossRef Medline
5. Kaur H, Mason JA, Bajracharya M, et al. **Transverse myelitis in a child with COVID-19.** *Pediatr Neurol* 2020;112:5–6 CrossRef Medline
6. AlKetbi R, AlNuaimi D, AlMulla M, et al. **Acute myelitis as a neurological complication of Covid-19: a case report and MRI findings.** *Radiol Case Rep* 2020;15:1591–95 CrossRef Medline
7. Baghbanian SM, Namazi F. **Post COVID-19 longitudinally extensive transverse myelitis (LETM): a case report.** *Acta Neurol Belg* 2020 Sept 18. [Epub ahead of print] CrossRef Medline
8. Fumery T, Baudar C, Ossemann M, et al. **Longitudinally extensive transverse myelitis following acute COVID-19 infection.** *Mult Scler Relat Disord* 2021;48:102723 CrossRef Medline
9. Valiuddin H, Skwirsk B, Paz-Arabo P. **Acute transverse myelitis associated with SARS-CoV-2: a case-report.** *Brain Behav Immun Health* 2020;5:100091 CrossRef Medline
10. Khedr EM, Karim AA, Soliman RK. **Case report: acute spinal cord myelopathy in patients with COVID-19.** *Front Neurol* 2020;11:610648 CrossRef Medline
11. Garg RK, Paliwal VK, Gupta A. **Spinal cord involvement in COVID-19: a review.** *J Spinal Cord Med* 2021 Mar 11. [Epub ahead of print] CrossRef Medline
12. Rodríguez de Antonio LA, González-Suárez I, Fernández-Barriuso I, et al. **Para-infectious anti-GD2/GD3 IgM myelitis during the Covid-19 pandemic: case report and literature review.** *Mult Scler Relat Disord* 2021;49:102783 CrossRef Medline
13. Zhou S, Jones-Lopez EC, Soneji DJ, et al. **Myelin oligodendrocyte glycoprotein antibody-associated optic neuritis and myelitis in COVID-19.** *J Neuroophthalmol* 2020;40:398–402 CrossRef Medline
14. Shaw VC, Chander G, Puttanna A. **Neuromyelitis optica spectrum disorder secondary to COVID-19.** *Br J Hosp Med (Lond)* 2020;81:1–3 CrossRef Medline
15. Sotoca J, Rodríguez-Álvarez Y. **COVID-19-associated acute necrotizing myelitis.** *Neurol Neuroimmunol Neuroinflamm* 2020;7:e803 CrossRef Medline
16. Zachariadis A, Tulbu A, Strambo D, et al. **Transverse myelitis related to COVID-19 infection.** *J Neurol* 2020;267:3459–61 CrossRef Medline
17. Abdelhady M, Elsotouhy A, Vattoth S. **Acute flaccid myelitis in COVID-19.** *BJR Case Rep* 2020;6:20200098 CrossRef Medline
18. Kim JW, Abdullayev N, Neuneier J, et al. **Post-COVID-19 encephalomyelitis.** *Neurol Res Pract* 2021;3:18 CrossRef Medline
19. Ellul MA, Benjamin L, Singh B, et al. **Neurological associations of COVID-19.** *Lancet Neurol* 2020;19:767–83 CrossRef Medline

20. Kranz PG, Amrhein TJ. **Imaging approach to myelopathy: acute, subacute, and chronic.** *Radiol Clin North Am* 2019;57:257–79 CrossRef Medline

21. Di Rocco A. **Diseases of the spinal cord in human immunodeficiency virus infection.** *Semin Neurol* 1999;19:151–55 CrossRef Medline

22. Umehara F, Nose H, Saito M, et al. **Abnormalities of spinal magnetic resonance images implicate clinical variability in human T-cell lymphotropic virus type I-associated myelopathy.** *J Neurovirol* 2007;13:260–67 CrossRef Medline

23. Winston GP, Jaiser SR. **Copper deficiency myelopathy and subacute combined degeneration of the cord: why is the phenotype so similar?** *Med Hypotheses* 2008;71:229–36 CrossRef Medline

24. Cao J, Su ZY, Xu SB, et al. **Subacute combined degeneration: a retrospective study of 68 cases with short-term follow-up.** *Eur Neurol* 2018;79:247–55 CrossRef Medline

25. Kumar N, Gross JB Jr, Ahlskog JE. **Copper deficiency myelopathy produces a clinical picture like subacute combined degeneration.** *Neurology* 2004;63:33–39 CrossRef Medline

26. Rowin J, Lewis SL. **Copper deficiency myeloneuropathy and pancytopenia secondary to overuse of zinc supplementation.** *J Neurol Neurosurg Psychiatry* 2005;76:750–51 CrossRef Medline

27. Pema PJ, Horak HA, Wyatt RH. **Myelopathy caused by nitrous oxide toxicity.** *AJNR Am J Neuroradiol* 1998;19:894–96 Medline

28. Yi Y, Kang HJ, Shin HY, et al. **Progressive myelopathy mimicking subacute combined degeneration after intrathecal chemotherapy.** *J Child Neurol* 2015;30:246–49 CrossRef Medline

29. Calder PC. **Nutrition, immunity and COVID-19.** *BMJ Nutr Prev Health* 2020;3:74–92 CrossRef Medline

30. Wee AK. **COVID-19's toll on the elderly and those with diabetes mellitus: is vitamin B12 deficiency an accomplice?** *Med Hypotheses* 2021;146:110374 CrossRef Medline

31. Zhang X, Hou C, Liu P, et al. **Methylenetetrahydrofolate reductase (MTHFR) C677T polymorphism and subacute combined degeneration: revealing a genetic predisposition.** *Front Neurol* 2018;9:1162 CrossRef Medline

32. Karst M, Hollenhorst J, Achenbach J. **Life-threatening course in coronavirus disease 2019 (COVID-19): is there a link to methylenetetrahydrofolic acid reductase (MTHFR) polymorphism and hyperhomocysteinemia?** *Med Hypotheses* 2020;144:110234 CrossRef Medline

Are Retropharyngeal Parathyroid Adenomas Associated with Retropharyngeal Carotid Arteries?

We read with interest the article by Steinl et al[1] discussing the simultaneous detection of retropharyngeal carotid arteries (CAs) and parathyroid adenomas in 5 patients. At our institution, we also perform 4D-CT for hyperparathyroidism and routinely highlight the presence of a retropharyngeal CA in the report to alert the operating surgeon. We agree with the authors that such a description is vital in preventing catastrophic hemorrhage, either during intubation or parathyroidectomy.

However, unlike the authors, we assess the position of the CA in relation to an uncovertebral joint and label the artery as retropharyngeal only when it lies medial to the joint. Our evaluation of the CA position is based on the methodology used by Koreckij et al.[2] The classification used by Pfeiffer and Ridder[3] involves taking 1 measurement on each side at the level of the nasopharynx, oropharynx, and hypopharynx, amounting to a total of at least 6 values for both CAs and potentially even more if the distances are measured for both the common and internal carotid arteries. Being affiliated with an academic institute, we find that using the uncovertebral joint is more reproducible across varying levels of training, ranging from the first-year resident to the advanced neuroimaging fellow.

Regarding the association between retropharyngeal parathyroid adenomas and medially displaced CAs, there are several confounding factors that can potentially render the association spurious. All patients discussed in the report of Steinl et al[1] were women, and 4 of 6 were older than 60 years of age. Studies have shown a higher prevalence of retropharyngeal CAs in both women and the older population. Koreckij et al[2] found a statistically significant correlation between female sex and retropharyngeal carotid deviation. Likewise, most patients studied by Pfeiffer and Ridder[3] were older than 60 years of age. Some authors even question whether medial deviation of the CAs is entirely age-related.[4] Atherosclerosis and hypertension are also important associations in retropharyngeal carotid deviation.[4] The status of atherosclerotic and hypertensive disease in the 5 cases discussed by Steinl et al[1] is unknown. Coincidentally, the retropharynx is also a common location for ectopic superior parathyroid adenomas.

To summarize, although we agree that there may still be a theoretic embryologic link between retropharyngeal CAs and parathyroid adenomas, we believe that the reader should be aware of the important confounding factors listed above. We also suggest using the uncovertebral joint as an anatomic landmark in routine clinical practice for its ease, speed, and reproducibility.

REFERENCES

1. Steinl GK, Yeh R, McManus CM, et al. **Variations in the course of the carotid arteries in patients with retropharyngeal parathyroid adenomas.** *AJNR Am J Neuroradiol* 2021;42:749–52 CrossRef Medline
2. Koreckij J, Alvi H, Gibly R, et al. **Incidence and risk factors of the retropharyngeal carotid artery on cervical magnetic resonance imaging.** *Spine (Phila Pa 1976)* 2013;38:E109–12 CrossRef Medline
3. Pfeiffer J, Ridder GJ. **A clinical classification system for aberrant internal carotid arteries.** *Laryngoscope* 2008;118:1931–36 CrossRef Medline
4. Del Corso L, Moruzzo D, Conte B, et al. **Tortuosity, kinking, and coiling of the carotid artery: expression of atherosclerosis or aging?** *Angiology* 1998;49:361–71 CrossRef Medline

🄳G.H. Kadam
🄳S.G. Dua
Section of Neuroimaging
Department of Radiology
Rush University Medical Center
Chicago, Illinois

REPLY:

We want to thank Drs Kadam and Dua for their thoughtful and interesting commentary on our article in a Letter to the Editor. We acknowledge that are different methods to evaluate the carotid artery (CA) position and define anatomic variations. The method used by the authors of the Letter to assess the CA position in relation to the uncovertebral joint by Koreckij et al[1] is, in fact, more reproducible, faster, and simpler than the classification by Pfeiffer et al[2] that we used in our article. Differences between the 2 methods are probably less scientific in nature than they are due to differences in perspective and operative approach between the orthopedic spine surgeons (Koreckij group) and head and neck surgeons (Pfeiffer group). Thus, Koreckij et al propose using the uncovertebral joint, a landmark in the spine, as the anatomic point of reference, whereas Pfeiffer proposes the pharyngeal wall, a landmark in the neck. On additional further review, 4 of 5 (80%) of our patients would qualify as having a retropharyngeal CA based on using the uncovertebral joint. Because we are a group of endocrine surgeons and a radiologist specializing in parathyroid disease, we used the classification based on the pharyngeal wall because this is more easily assessed in our surgical field than the uncovertebral joint.

We agree with the authors that age and sex are potential confounders that cannot be fully analyzed in a case series of 5 patients. However, the average patient for primary hyperparathyroidism is a 62-year-old woman (affecting women 3 times more often than men); therefore, a case series of 5 women older than 60 years of age is not unique in this patient cohort. Additionally, although posteriorly located upper adenomas are rather common, accounting for 45% of all single gland adenomas, they are often located in the paraesophageal space or tracheoesophageal groove.[3] True, ectopic retroesophageal adenomas are rarer, accounting for only 6.7% of all parathyroid adenomas, the second most common ectopic location (the first is the cervical thymus) to the cervical thymus.[4] We agree that to definitively analyze whether there is a true association between kissing carotids and retroesophageal adenomas, a case-control study comparing the incidence of retroesophageal adenomas in all patients with primary hyperparathyroidism with that in those patients with kissing carotids is required. Unfortunately, given the overall low incidence, it is unlikely that a study can be conducted with enough power to determine the association.

Thus, the intention of our article was not to actually prove this association, but to highlight this potential association between kissing carotids and retroesophageal adenomas that we noted among these 5 patients. Even though the association cannot be proved statistically, we believe that given the much higher risk of morbidity with re-operative parathyroidectomy, any guidance that can facilitate successful localization of ectopic parathyroid adenomas before the first exploration can be helpful.

REFERENCES

1. Koreckij J, Alvi H, Gibly R, et al. **Incidence and risk factors of the retropharyngeal carotid artery on cervical magnetic resonance imaging.** *Spine (Phila Pa)* 1976;38:E109–12 CrossRef Medline
2. Pfeiffer J, Ridder GJ. **A clinical classification system for aberrant internal carotid arteries.** *Laryngoscope* 2008;118:1931–36 CrossRef Medline
3. Harari A, Mitmaker E, Grogan RH, et al. **Primary hyperparathyroidism patients with positive preoperative sestamibi scan and negative ultrasound are more likely to have posteriorly located upper gland adenomas (PLUGs).** *Ann Surg Oncol* 2011;18:1717–22 CrossRef Medline
4. Roy M, Mazeh H, Chen H, et al. **Incidence and localization of ectopic parathyroid adenomas in previously unexplored patients.** *World J Surg* 2013;37:102–06 CrossRef Medline

J.H. Kuo
Section of Endocrine Surgery,
Columbia University, New York, New York
R. Yeh, on behalf of all the authors
Department of Radiology,
Memorial Sloan Kettering Cancer Center, New York, New York

In the articles "Antiplatelet Management for Stent-Assisted Coiling and Flow Diversion of Ruptured Intracranial Aneurysms: A DELPHI Consensus Statement" (*AJNR Am J Neuroradiol* 2020;41:1856–62) and "Considerations for Antiplatelet Management of Carotid Stent Placement in the Setting of Mechanical Thrombectomy: A Delphi Consensus Statement" (*AJNR Am J Neuroradiol* 2020;41:2274–79), the dosage parameter of eptifibatide in Table 1 should have been mcg/kg/min.

The authors regret the errors.