

APRIL 2024

AJNR

VOLUME 45 • PP 361-531

# AJNR

## AMERICAN JOURNAL OF NEURORADIOLOGY

Official Journal ASNR • ASFNR • ASHNR • ASPNR • ASSR  
APRIL 2024 | VOLUME 45 | NUMBER 4 | WWW.AJNR.ORG

THE JOURNAL OF DIAGNOSTIC AND INTERVENTIONAL NEURORADIOLOGY

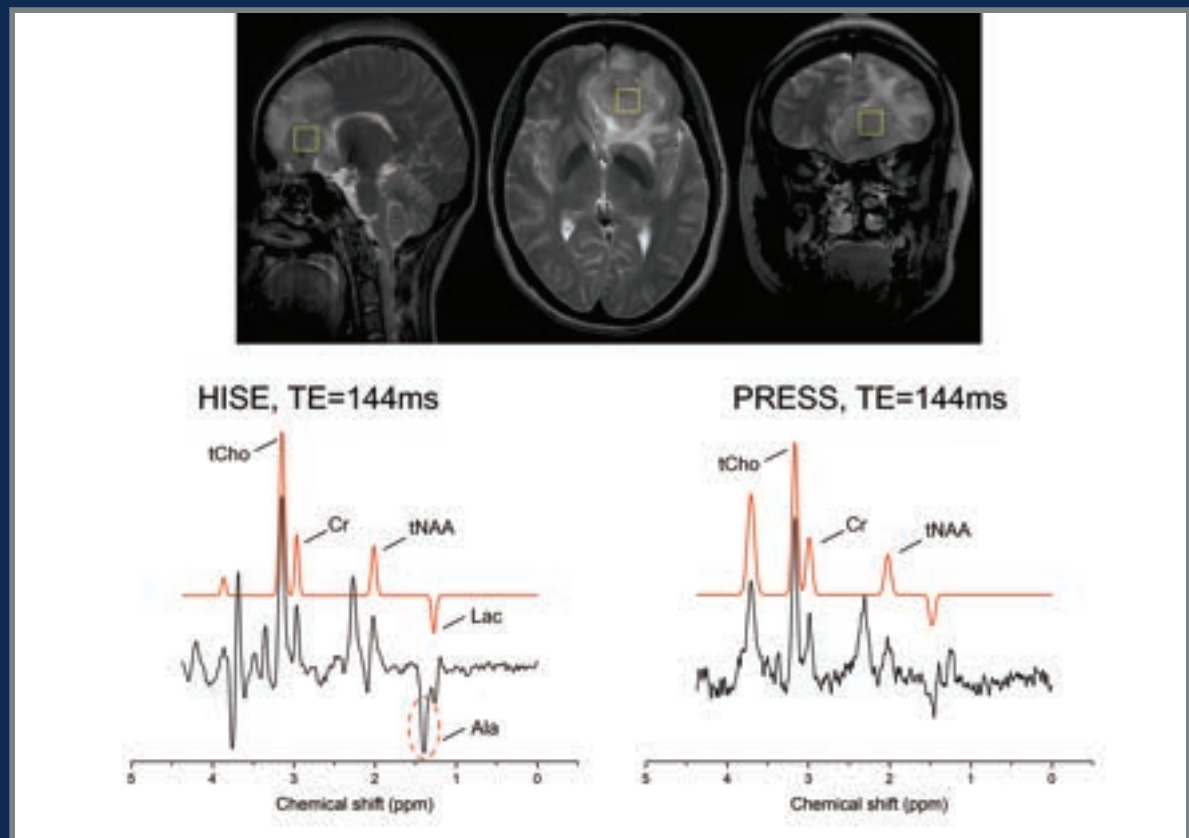
Bias in artificial intelligence

Microvascular brain damage in pseudoxanthoma elasticum

Deep learning approach to identify patients with CSF-venous fistula

New entity: high-grade astrocytoma with piloid features

MR neurography for postoperative peripheral trigeminal neuropathies

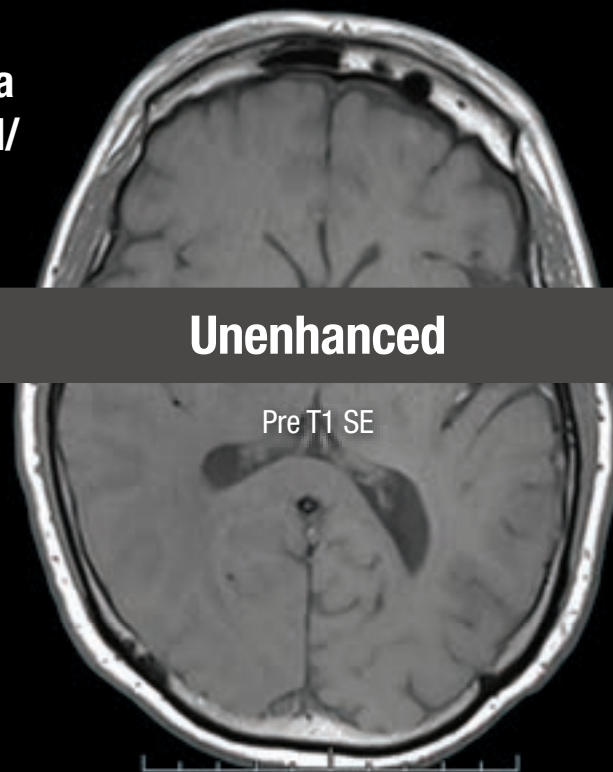


# THIS IS HALF Gd\*

**\*Effective contrast enhancement at half the gadolinium dose (0.05 mmol/kg) vs a macrocyclic GBCA at a dose of 0.1 mmol/kg in approved indications in the U.S.<sup>1-6†</sup>**

  
**Vueway®**  
(gadopiclenol) injection  
485.1 mg/mL

**NO COMPROMISE IN MRI FROM BRACCO,  
YOUR TRUSTED PARTNER**



†Phase III CNS Study Design (Study GDX-44-010): Intra-individual, crossover comparison of 0.05 mmol/kg VUEWAY (gadopiclenol) injection vs. 0.1 mmol/kg Gadavist® in MRI of the CNS. Patients with known or suspected CNS lesions. Three primary visualization endpoints (lesion border delineation, lesion internal morphology, degree of contrast enhancement). The CNS study included 256 patients with known or highly suspected CNS lesion(s) with a mean age of 57 years (range: 18-84 years), and 53% female patients.

Please see Brief Summary of Prescribing Information including Boxed Warning on adjacent page.

#### **VUEWAY® (gadopiclenol) solution for injection**

##### **Indications**

VUEWAY injection is indicated in adults and children aged 2 years and older for use with magnetic resonance imaging (MRI) to detect and visualize lesions with abnormal vascularity in:

- the central nervous system (brain, spine and surrounding tissues),
- the body (head and neck, thorax, abdomen, pelvis, and musculoskeletal system).

##### **IMPORTANT SAFETY INFORMATION**

**WARNING: RISK ASSOCIATED WITH INTRATHECAL USE and NEPHROGENIC SYSTEMIC FIBROSIS**

##### **Risk Associated with Intrathecal Use**

Intrathecal administration of gadolinium-based contrast agents (GBCAs) can cause serious adverse reactions including death, coma, encephalopathy, and seizures. VUEWAY is not approved for intrathecal use.

##### **NEPHROGENIC SYSTEMIC FIBROSIS**

Gadolinium-based contrast agents (GBCAs) increase the risk for NSF among patients with impaired elimination of the drugs. Avoid use of GBCAs in these patients unless the diagnostic information is essential and not available with non-contrasted MRI or other modalities. NSF may result in fatal or debilitating fibrosis affecting the skin, muscle and internal organs.

- The risk for NSF appears highest among patients with:

- **Chronic, severe kidney disease (GFR < 30 mL/min/1.73 m<sup>2</sup>), or**
- **Acute kidney injury.**
- **Screen patients for acute kidney injury and other conditions that may reduce renal function. For patients at risk for chronically reduced renal function (e.g. age > 60 years, hypertension, diabetes), estimate the glomerular filtration rate (GFR) through laboratory testing.**
- **For patients at highest risk for NSF, do not exceed the recommended VUEWAY dose and allow a sufficient period of time for elimination of the drug from the body prior to any re-administration.**

##### **Contraindications**

VUEWAY injection is contraindicated in patients with history of hypersensitivity reactions to VUEWAY.

##### **Warnings and Precautions**

There are **risks associated with intrathecal use** of GBCAs that can cause serious adverse reactions including death, coma, encephalopathy, and seizures. The safety and effectiveness of VUEWAY have not been established with intrathecal use and VUEWAY is not approved for intrathecal use.

Risk of **nephrogenic systemic fibrosis** is increased in patients using GBCA agents that have impaired elimination of the drugs, with the highest risk in patients with chronic, severe kidney disease as well as patients with acute kidney injury. Avoid use of GBCAs among these patients unless the diagnostic information is essential and not available with non-contrast MRI or other modalities.

**Hypersensitivity reactions**, including serious hypersensitivity reactions, could occur during use or shortly following VUEWAY administration. Assess all patients for any

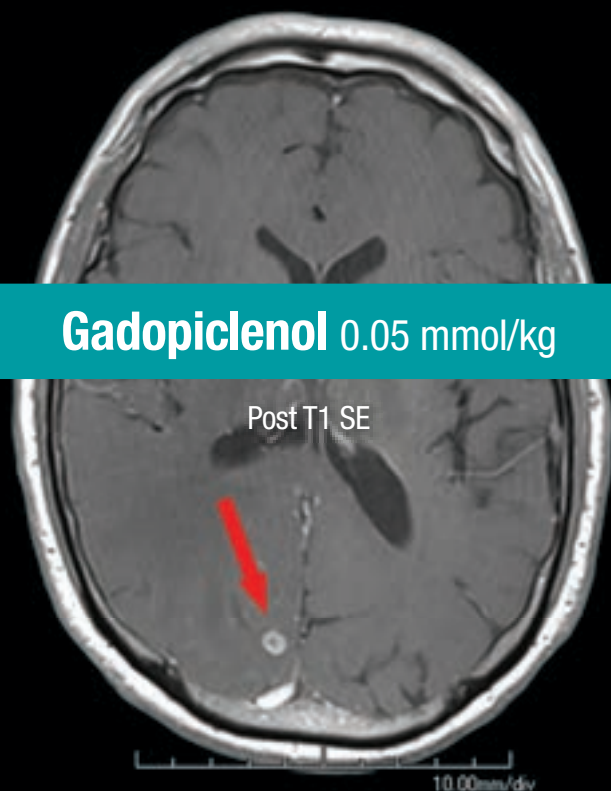


LIFE FROM INSIDE

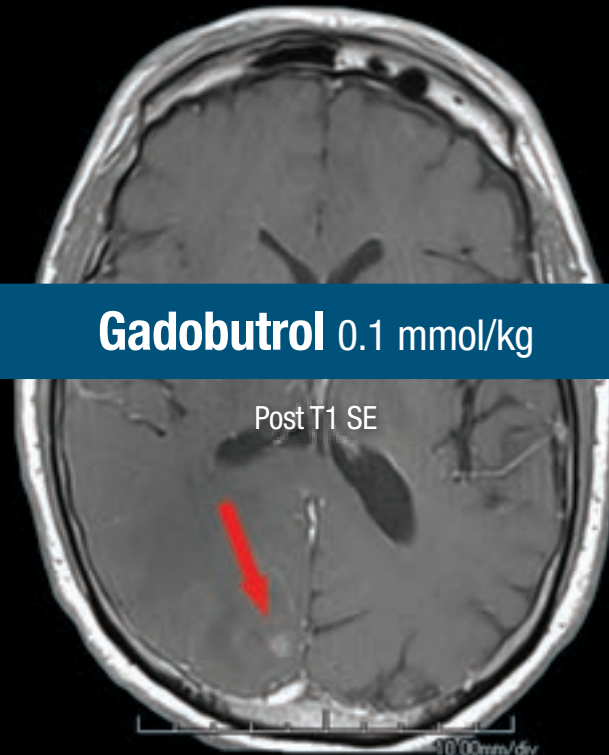
65-year-old man – 3.0 T Siemens  
Brain metastasis from lung adenocarcinoma<sup>7</sup>

SE = Spin Echo. These are representative images from reference studies; individual results may vary.

**Gadopichlenol** 0.05 mmol/kg



**Gadobutrol** 0.1 mmol/kg



history of a reaction to contrast media, bronchial asthma and/or allergic disorders, administer VUEWAY only in situations where trained personnel and therapies are promptly available for the treatment of hypersensitivity reactions, and observe patients for signs and symptoms of hypersensitivity reactions after administration.

**Gadolinium retention** can be for months or years in several organs after administration. The highest concentrations (nanomoles per gram of tissue) have been identified in the bone, followed by other organs (brain, skin, kidney, liver and spleen). Minimize repetitive GBCA imaging studies, particularly closely spaced studies, when possible.

**Acute kidney injury** requiring dialysis has occurred with the use of GBCAs in patients with chronically reduced renal function. The risk of acute kidney injury may increase with increasing dose of the contrast agent.

**Extravasation and injection site reactions** can occur with administration of VUEWAY. Ensure catheter and venous patency before the injection of VUEWAY.

VUEWAY may **impair the visualization of lesions** seen on non-contrast MRI. Therefore, caution should be exercised when VUEWAY MRI scans are interpreted without a companion non-contrast MRI scan.

The most common adverse reactions (incidence  $\geq 0.5\%$ ) are injection site pain (0.7%), and headache (0.7%).

**You are encouraged to report negative side effects of prescription drugs to the FDA. Visit [www.fda.gov/medwatch](http://www.fda.gov/medwatch) or call 1-800-FDA-1088.**

**Please see BRIEF SUMMARY of Prescribing Information for VUEWAY**

**(gadopichlenol) solution for injection including BOXED WARNING on Nephrogenic Systemic Fibrosis.**

Manufactured for Bracco Diagnostics Inc. by Liebel-Flarsheim Company LLC - Raleigh, NC, USA 27616.

VUEWAY is a trademark of Bracco Imaging S.p.A.

All other trademarks and registered trademarks are the property of their respective owners.

**References:** **1.** VUEWAY® (gadopichlenol) solution for injection, 485.1 mg/mL Full Prescribing Information and Patient Medication Guide. Monroe Twp., NJ: Bracco Diagnostics Inc.; January 2024. **2.** Robic C, Port M, Rousseaux O, et al. Physicochemical and pharmacokinetic profiles of gadopichlenol: a new macrocyclic gadolinium chelate with high T1 relaxivity. *Invest Radiol.* 2019 Aug;54:475-484. **3.** GADAVIST® (gadobutrol) Injection. Full Prescribing Information. Bayer HealthCare Pharmaceuticals Inc. Whippany, NJ; April 2022. **4.** DOTAREM® (gadoterate meglumine) Injection. Full Prescribing Information. Guerbet LLC. Princeton, NJ; April 2022. **5.** CLARISCAN™ (gadoterate meglumine) injection for intravenous use. Full Prescribing Information. GE Healthcare. Chicago, IL; February 2020. **6.** ProHance® (Gadoteridol) Injection, 279.3 mg/mL Full Prescribing Information and Patient Medication Guide. Monroe Twp., NJ: Bracco Diagnostics Inc.; January 2024. **7.** Loevner LA, Kolumban B, Hutóczy G, et al. Efficacy and safety of gadopichlenol for contrast-enhanced MRI of the central nervous system: the PICTURE randomized clinical trial. *Invest Radiol.* 2023 May;58(5):307-313.

Bracco Diagnostics Inc.  
259 Prospect Plains Road, Building H  
Monroe Township, NJ 08831 USA  
Phone: 609-514-2200  
Toll-Free: 1-877-272-2269 (U.S. only)  
Fax: 609-514-2446

© 2024 Bracco Diagnostics Inc.  
All Rights Reserved. US-VW-2300022 02/24



VISIT  
[VUEWAY.COM](http://VUEWAY.COM)  
FOR MORE  
INFORMATION

## VUEWAY®

(gadopiclenol) injection, for intravenous use

**BRIEF SUMMARY:** Please see package insert of full prescribing information.

### WARNING: RISK ASSOCIATED WITH INTRATHECAL USE and NEPHROGENIC SYSTEMIC FIBROSIS

#### Risk Associated with Intrathecal Use

Intrathecal administration of gadolinium-based contrast agents (GBCAs) can cause serious adverse reactions including death, coma, encephalopathy, and seizures. Vueway is not approved for intrathecal use [see Warnings and Precautions (5.1) in the full Prescribing Information].

#### Nephrogenic Systemic Fibrosis

Gadolinium-based contrast agents (GBCAs) increase the risk for NSF among patients with impaired elimination of the drugs. Avoid use of GBCAs in these patients unless the diagnostic information is essential and not available with non-contrast MRI or other modalities. NSF may result in fatal or debilitating fibrosis affecting the skin, muscle and internal organs.

- Patients for NSF appears highest among patients with:
  - Chronic, severe kidney disease (GFR <30 mL/min/1.73 m<sup>2</sup>), or
  - Acute kidney injury.
- Screen patients for acute kidney injury and other conditions that may reduce renal function. For patients at risk for chronically reduced renal function (e.g. age >60 years, hypertension, diabetes), estimate the glomerular filtration rate (GFR) through laboratory testing.
- For patients at highest risk for NSF, do not exceed the recommended VUEWAY dose and allow a sufficient period of time for elimination of the drug from the body prior to any re-administration [see Warnings and Precautions (5.2) in the full Prescribing Information].

### INDICATIONS AND USAGE

VUEWAY® (gadopiclenol) is a gadolinium-based contrast agent indicated in adult and pediatric patients aged 2 years and older for use with magnetic resonance imaging (MRI) to detect and visualize lesions with abnormal vascularity in:

- the central nervous system (brain, spine, and associated tissues),
- the body (head and neck, thorax, abdomen, pelvis, and musculoskeletal system).

### CONTRAINDICATIONS

VUEWAY is contraindicated in patients with history of hypersensitivity reactions to gadopidolol.

### WARNINGS AND PRECAUTIONS

**Risk Associated with Intrathecal Use** Intrathecal administration of GBCAs can cause serious adverse reactions including death, coma, encephalopathy, and seizures. The safety and effectiveness of VUEWAY have not been established with intrathecal use. VUEWAY is not approved for intrathecal use [see Dosage and Administration (2.1) in the full Prescribing Information].

**Nephrogenic Systemic Fibrosis** Gadolinium-based contrast agents (GBCAs) increase the risk for nephrogenic systemic fibrosis (NSF) among patients with impaired elimination of the drugs. Avoid use of GBCAs among these patients unless the diagnostic information is essential and not available with non-contrast MRI or other modalities. The GBCA-associated NSF risk appears highest for patients with chronic, severe kidney disease (GFR <30 mL/min/1.73 m<sup>2</sup>) as well as patients with acute kidney injury. The risk appears lower for patients with chronic, moderate kidney disease (GFR 30-59 mL/min/1.73 m<sup>2</sup>) and little, if any, for patients with chronic, mild kidney disease (GFR 60-89 mL/min/1.73 m<sup>2</sup>). NSF may result in fatal or debilitating fibrosis affecting the skin, muscle, and internal organs. Report any diagnosis of NSF following VUEWAY administration to Bracco Diagnostics Inc. (1-800-257-5181) or FDA (1-800-FDA-1088 or [www.fda.gov/medwatch](http://www.fda.gov/medwatch)).

Screen patients for acute kidney injury and other conditions that may reduce renal function. Features of acute kidney injury consist of rapid (over hours to days) and usually reversible decrease in kidney function, commonly in the setting of surgery, severe infection, injury or drug-induced kidney toxicity. Serum creatinine levels and estimated GFR may not reliably assess renal function in the setting of acute kidney injury. For patients at risk for chronically reduced renal function (e.g., age >60 years, diabetes mellitus or chronic hypertension), estimate the GFR through laboratory testing.

Among the factors that may increase the risk for NSF are repeated or higher than recommended doses of a GBCA and the degree of renal impairment at the time of exposure. Record the specific GBCA and the dose administered to a patient. For patients at highest risk for NSF, do not exceed the recommended VUEWAY dose and allow a sufficient period of time for elimination of the drug prior to re-administration. For patients receiving hemodialysis, physicians may consider the prompt initiation of hemodialysis following the administration of a GBCA in order to enhance the contrast agent's elimination [see Use in Specific Populations (8.6) and Clinical Pharmacology (12.3) in the full Prescribing Information]. The usefulness of hemodialysis in the prevention of NSF is unknown.

**Hypersensitivity Reactions** With GBCAs, serious hypersensitivity reactions have occurred. In most cases, initial symptoms occurred within minutes of GBCA administration and resolved with prompt emergency treatment.

- Before VUEWAY administration, assess all patients for any history of a reaction to contrast media, bronchial asthma and/or allergic disorders. These patients may have an increased risk for a hypersensitivity reaction to VUEWAY.
- VUEWAY is contraindicated in patients with history of hypersensitivity reactions to VUEWAY [see Contraindications (4) in the full Prescribing Information].
- Administer VUEWAY only in situations where trained personnel and therapies are promptly available for the treatment of hypersensitivity reactions, including personnel trained in resuscitation.

- During and following VUEWAY administration, observe patients for signs and symptoms of hypersensitivity reactions.

**Gadolinium Retention** Gadolinium is retained for months or years in several organs. The highest concentrations (nanomoles per gram of tissue) have been identified in the bone, followed by other organs (e.g. brain, skin, kidney, liver, and spleen). The duration of retention also varies by tissue and is longest in bone. Linear GBCAs cause more retention than macrocyclic GBCAs. At equivalent doses, gadolinium retention varies among the linear agents with gadodiamide causing greater retention than other linear agents such as gadoxetate disodium, and gadobenate dimeglumine. Retention is lowest and similar among the macrocyclic GBCAs such as gadoterate meglumine, gadobutrol, gadoteridol, and gadopidolol. Consequences of gadolinium retention in the brain have not been established. Pathologic and clinical consequences of GBCA administration and retention in skin and other organs have been established in patients with impaired renal function [see Warnings and Precautions (5.2) in the full Prescribing Information]. There are rare reports of pathologic skin changes in patients with normal renal function. Adverse events involving multiple organ systems have been reported in patients with normal renal function without an established causal link to gadolinium.

While clinical consequences of gadolinium retention have not been established in patients with normal renal function, certain patients might be at higher risk. These include patients requiring multiple lifetime doses, pregnant and pediatric patients, and patients with inflammatory conditions. Consider the retention characteristics of the agent when choosing a GBCA for these patients. Minimize repetitive GBCA imaging studies, particularly closely spaced studies, when possible.

**Acute Kidney Injury** In patients with chronically reduced renal function, acute kidney injury requiring dialysis has occurred with the use of GBCAs. The risk of acute kidney injury may increase with increasing dose of the contrast agent. Do not exceed the recommended dose.

**Extravasation and Injection Site Reactions** Injection site reactions such as injection site pain have been reported in the clinical studies with VUEWAY [see Adverse Reactions (6.1) in the full Prescribing Information]. Extravasation during VUEWAY administration may result in tissue irritation [see Nonclinical Toxicology (13.2) in the full Prescribing Information]. Ensure catheter and venous patency before the injection of VUEWAY.

**Interference with Visualization of Lesions Visible with Non-Contrast MRI** As with any GBCA, VUEWAY may impair the visualization of lesions seen on non-contrast MRI. Therefore, caution should be exercised when VUEWAY MRI scans are interpreted without a companion non-contrast MRI scan.

### ADVERSE REACTIONS

The following serious adverse reactions are discussed elsewhere in labeling:

- Nephrogenic Systemic Fibrosis [see Warnings and Precautions (5.2) in the full Prescribing Information]
- Hypersensitivity Reactions [see Contraindications (4) and Warnings and Precautions (5.3) in the full Prescribing Information]

**Clinical Trials Experience** Because clinical trials are conducted under widely varying conditions, adverse reaction rates observed in the clinical trials of a drug cannot be directly compared to rates in the clinical trials of another drug and may not reflect the rates observed in clinical practice.

The safety of VUEWAY was evaluated in 1,047 patients who received VUEWAY at doses ranging from 0.025 mmol/kg (one half the recommended dose) to 0.3 mmol/kg (six times the recommended dose). A total of 708 patients received the recommended dose of 0.05 mmol/kg. Among patients who received the recommended dose, the average age was 51 years (range 2 years to 88 years) and 56% were female. The ethnic distribution was 79% White, 10% Asian, 7% American Indian or Alaska native, 2% Black, and 2% patients of other or unspecified ethnic groups.

Overall, approximately 4.7% of subjects receiving the labeled dose reported one or more adverse reactions.

Table 1 lists adverse reactions that occurred in >0.2% of patients who received 0.05 mmol/kg VUEWAY.

TABLE 1. ADVERSE REACTIONS REPORTED IN >0.2% OF PATIENTS RECEIVING VUEWAY IN CLINICAL TRIALS	
Adverse Reaction	VUEWAY 0.05 mmol/kg (n=708) (%)
Injection site pain	0.7
Headache	0.7
Nausea	0.4
Injection site warmth	0.4
Injection site coldness	0.3
Dizziness	0.3
Local swelling	0.3

Adverse reactions that occurred with a frequency ≤ 0.2% in patients who received 0.05 mmol/kg VUEWAY included: maculopapular rash, vomiting, worsened renal impairment, feeling hot, pyrexia, oral paresthesia, dysgeusia, diarrhea, pruritus, allergic dermatitis, erythema, injection site paresthesia, Cystatin C increase, and blood creatinine increase.

#### Adverse Reactions in Pediatric Patients

One study with a single dose of VUEWAY (0.05 mmol/kg) was conducted in 80 pediatric patients aged 2 years to 17 years, including 60 patients who underwent a central nervous system (CNS) MRI and 20 patients who underwent a body MRI. One adverse reaction (maculopapular rash of moderate severity) in one patient (1.3%) was reported in the CNS cohort.

#### USE IN SPECIFIC POPULATIONS

**Pregnancy Risk Summary** There are no available data on VUEWAY use in pregnant women to evaluate for a drug-associated risk of major birth defects, miscarriage or other adverse maternal or fetal outcomes. GBCAs cross the human placenta and result in fetal exposure and gadolinium retention. The available human data on GBCA exposure during pregnancy and adverse fetal outcomes are limited and inconclusive [see Data]. In animal reproduction studies, there were no adverse developmental effects observed in rats or rabbits with intravenous administration of VUEWAY during organogenesis [see Data]. Because of the potential risks of gadolinium to the fetus, use VUEWAY only if imaging is essential during pregnancy and cannot be delayed. The estimated background risk of major birth defects and miscarriage for the indicated population(s) are unknown. All pregnancies have a background risk of birth defect, loss, or other adverse outcomes. In the U.S. general population, the estimated background risk of major birth defects and miscarriage in clinically recognized pregnancies is 2% to 4% and 15% to 20% respectively. **Data Human Data Contrast enhancement** is visualized in the placenta and fetal tissues after maternal GBCA administration. Cohort studies and case reports on exposure to GBCAs during pregnancy have not reported a clear association between GBCAs and adverse

effects in the exposed neonates. However, a retrospective cohort study comparing pregnant women who had a GBCA MRI to pregnant women who did not have an MRI reported a higher occurrence of stillbirths and neonatal deaths in the group receiving GBCA MRI. Limitations of this study include a lack of comparison with non-contrast MRI and lack of information about the maternal indication for MRI. Overall, these data preclude a reliable evaluation of the potential risk of adverse fetal outcomes with the use of GBCAs in pregnancy.

**Animal Data Gadolinium Retention:** GBCAs administered to pregnant non-human primates (0.1 mmol/kg on gestational days 85 and 135) result in measurable gadolinium concentration in the offspring in bone, brain, skin, liver, kidney, and spleen for at least 7 months. GBCAs administered to pregnant mice (2 mmol/kg daily on gestational days 16 through 19) result in measurable gadolinium concentrations in the pups in bone, brain, kidney, liver, blood, muscle, and spleen at one-month postnatal age.

**Reproductive Toxicology:** Animal reproduction studies conducted with gadopidolol showed some signs of maternal toxicity at 10 mmol/kg and rabbits at 5 mmol/kg (corresponding to 52 times and 57 times the recommended human dose, respectively). This maternal toxicity was characterized in both species by swelling, decreased activity, and lower gestation weight gain and food consumption.

No effect on embryo-fetal development was observed in rats at 10 mmol/kg (corresponding to 52 times the recommended human dose). In rabbits, a lower mean fetal body weight was observed at 5 mmol/kg (corresponding to 57 times the recommended human dose) and this was attributed as a consequence of the lower gestation weight gain.

**Lactation Risk Summary** There are no data on the presence of gadopidolol in human milk, the effects on the breastfed infant, or the effects on milk production. However, published lactation data on other GBCAs indicate that 0.01% to 0.04% of the maternal gadolinium dose is excreted in breast milk. Additionally, there is limited GBCA gastrointestinal absorption in the breast-fed infant. Gadopidolol is present in rat milk. When a drug is present in animal milk, it is likely that the drug will be present in human milk [see Data]. The developmental and health benefits of breastfeeding should be considered along with the mother's clinical need for VUEWAY and any potential adverse effects on the breastfed infant from VUEWAY or from the underlying maternal condition. **Data** In lactating rats receiving single intravenous injection of [<sup>153</sup>Gd]-gadopiclenol, 0.3% and 0.2% of the total administered radioactivity was transferred to the pups via maternal milk at 6 hours and 24 hours after administration, respectively. Furthermore, in nursing rat pups, oral absorption of gadopidolol was 3.6%.

**Pediatric Use** The safety and effectiveness of VUEWAY for use with MRI to detect and visualize lesions with abnormal vascularity in the CNS (brain, spine, and associated tissues), and the body (head and neck, thorax, abdomen, pelvis, and musculoskeletal system) have been established in pediatric patients aged 2 years and older.

Use of VUEWAY in this age group is supported by evidence from adequate and well-controlled studies in adults with additional pharmacokinetic and safety data from an open-label, uncontrolled, multicenter, single dose study of VUEWAY (0.05 mmol/kg) in 80 pediatric patients aged 2 to 17 years. The 80 patients consisted of 60 patients who underwent a CNS MRI and 20 patients who underwent a body MRI [see Adverse Reactions (6.1) and Clinical Pharmacology (12.3) in the full Prescribing Information].

The safety and effectiveness of VUEWAY have not been established in pediatric patients younger than 2 years of age.

**Geriatric Use** Of the total number of VUEWAY-treated patients in clinical studies, 270 (26%) patients were 65 years of age and over, while 62 (6%) patients were 75 years of age and over. No overall differences in safety or efficacy were observed between these subjects and younger subjects.

This drug is known to be substantially excreted by the kidney, and the risk of adverse reactions to this drug may be greater in patients with impaired renal function. Because elderly patients are more likely to have decreased renal function, it may be useful to monitor renal function.

**Renal Impairment** In patients with renal impairment, the exposure of gadopidolol is increased compared to patients with normal renal function. This may increase the risk of adverse reactions such as nephrogenic systemic fibrosis (NSF). Avoid use of GBCAs among these patients unless the diagnostic information is essential and not available with non-contrast MRI or other modalities. No dose adjustment of VUEWAY is recommended for patients with renal impairment. VUEWAY can be removed from the body by hemodialysis [see Warnings and Precautions (5.2, 5.4, 5.5) and Clinical Pharmacology (12.3) in the full Prescribing Information].

### OVERDOSAGE

Among subjects who received a single 0.3 mmol/kg intravenous dose of gadopidolol (6 times the recommended dose of VUEWAY), headache and nausea were the most frequently reported adverse reactions. Gadopidolol can be removed from the body by hemodialysis [see Clinical Pharmacology (12.3) in the full Prescribing Information].

**PATIENT COUNSELING INFORMATION** Advise the patient to read the FDA-approved patient labeling (Medication Guide).

**Nephrogenic Systemic Fibrosis** Inform the patient that VUEWAY may increase the risk for NSF among patients with impaired elimination of the drugs and that NSF may result in fatal or debilitating fibrosis affecting the skin, muscle and internal organs.

Instruct the patients to contact their physician if they develop signs or symptoms of NSF following VUEWAY administration, such as burning, itching, swelling, scaling, hardening and tightening of the skin; red or dark patches on the skin; stiffness in joints with trouble moving, bending or straightening the arms, hands, legs or feet; pain in the hip bones or ribs; or muscle weakness [see Warnings and Precautions (5.2) in the full Prescribing Information].

**Gadolinium Retention** Advise patients that gadolinium is retained for months or years in brain, bone, skin, and other organs following VUEWAY administration even in patients with normal renal function. The clinical consequences of retention are unknown. Retention depends on multiple factors and is greater following administration of linear GBCAs than following administration of macrocyclic GBCAs [see Warnings and Precautions (5.4) in the full Prescribing Information].

**Injection Site Reactions** Inform the patient that VUEWAY may cause reactions along the venous injection site, such as mild and transient burning or pain or feeling of warmth or coldness at the injection site [see Warnings and Precautions (5.6) in the full Prescribing Information].

**Pregnancy** Advise pregnant women of the potential risk of fetal exposure to VUEWAY [see Use in Specific Populations (8.1) in the full Prescribing Information].

### Rx only

US Patent No. 10,973,934  
Manufactured for Bracco Diagnostics Inc. by Liebel-Flarsheim Company LLC  
- Raleigh, NC, USA 27616.  
Toll-Free: 1-877-272-2269 (U.S. only)  
Revised February 2024





# The ASNR Career Center

**The Go-To Job Site for Neuroradiology Employers and Job Seekers**

***For Job Seekers***

- Access to an expanded network of jobs via the National Healthcare Career Network
- Confidential resume posting
- Professional online profile

***For Employers***

- Employer resources to help you recruit top talent
- Multiple pricing options, including free Fellowship listings
- Resume search

**Start here: [careers.asnr.org](https://careers.asnr.org)**



## **Simplify** the MOC Process



## **Manage** your CME Credits Online

# **CMEgateway.org**

### **Available to Members of Participating Societies**

American Board of Radiology (ABR)  
American College of Radiology (ACR)  
American Roentgen Ray Society (ARRS)  
American Society of Neuroradiology (ASNR)  
Commission on Accreditation of Medical  
Physics Educational Programs, Inc. (CAMPEP)  
Radiological Society of North America (RSNA)  
Society of Interventional Radiology (SIR)  
SNM  
The Society for Pediatric Radiology (SPR)

### **It's Easy and Free!**

Log on to CME Gateway to:

- View or print reports of your CME credits from multiple societies from a single access point.
- Print an aggregated report or certificate from each participating organization.
- Link to SAMs and other tools to help with maintenance of certification.

### **American Board of Radiology (ABR) participation!**

By activating ABR in your organizational profile, your MOC-fulfilling CME and SAM credits can be transferred to your own personalized database on the ABR Web site.

### **Sign Up Today!**

go to [CMEgateway.org](http://CMEgateway.org)

# 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](http://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.



# AJNR

Dear *AJNR* Subscriber:

I hope this message finds you well. We've made significant enhancements to *AJNR*, including an expanded Editorial Board, an Academy of Reviewers, and a Junior Editorial Board. Now, I'm excited to share a crucial update about the *AJNR*'s future.

Towards the end of the year, *AJNR* will undergo a major transformation, transitioning to a fully electronic format and bidding farewell to traditional paper publication. This shift is driven by the need to manage rising costs associated with printing, shipping, and storage. Our goal is to achieve fiscal balance without imposing substantial increases in article processing charges.

The move to an electronic format brings numerous benefits. It opens the door to more personalized content, allowing readers to tailor their *AJNR* experience. The electronic format also enhances visual elements with more images and various media types, including videos and scrollable image stacks, especially valuable in neuroradiology.

Additionally, the electronic format allows for supplementary content, enriching the reader's experience with easy access to materials, commentaries, and related resources. We're committed to implementing enhanced searchability, catering to the diverse needs of neuroradiologists, researchers, trainees, and students.

While I have been a dedicated collector of *AJNR*'s physical copies for over two decades, recognizing the sadness in parting with this tradition, financial pragmatism compels us to embrace technological advancements. Despite this transition, we aim to maintain a sense of optimism and adaptability.

To involve our community in this journey, *AJNR* plans sessions to gather feedback, concerns, and innovative ideas. By fostering an open dialogue, we aim to shape our electronic platform in line with our audience's evolving expectations.

Please participate in these sessions that will be advertised on the *AJNR* website or please feel free to reach out directly to me ([mw.ajnr.eic@gmail.com](mailto:mw.ajnr.eic@gmail.com)) with any concern or suggestion.

Also, please immediately activate your individual subscription account and also consider registering for Publish Ahead of Print (including Preprint) and electronic Table of Contents, Case of the Week, and Keyword/Author Alerts. These links as well as additional information including upcoming Zoom sessions can be found at the *AJNR* website using the QR code below.

<https://www.ajnr.org/content/ajnr-update>

Thank you in advance for your engagement.

Sincerely,

Max Wintermark  
Editor-in-Chief



The graphic features the AJNR logo at the top, followed by the headline 'Moving to a fully Electronic Format by 2025'. Below this, it states 'Read the important announcement from Editor-in-Chief Dr. Max Wintermark in this month's editorial.' A box contains the following information:

Feedback and innovative ideas are encouraged through:	<ul style="list-style-type: none"><li>Online sessions</li><li>AJNR booth at ASNR24</li><li>Email <a href="mailto:mw.ajnr.eic@gmail.com">mw.ajnr.eic@gmail.com</a></li></ul>
---	---

A QR code is located to the right of the text, and an image of a hand holding a smartphone is at the bottom right.

# AJNR

## AMERICAN JOURNAL OF NEURORADIOLOGY

APRIL 2024  
VOLUME 45  
NUMBER 4  
WWW.AJNR.ORG

Publication Preview at [www.ajnr.org](http://www.ajnr.org) features articles released in advance of print.  
Visit [www.ajnrblog.org](http://www.ajnrblog.org) to comment on AJNR content and chat with colleagues.

### EDITORIALS

- 361 **Navigating Change: AJNR's Transition to an Electronic Format**  
*Max Wintermark, Editor-in-Chief*
- 362 **Introducing Our New AJNR Early Career, Women in Neuroradiology, and Global Neuroradiology Awards** *Nicholas S. Cho, et al.*
- 363 **Imaging Large Ischemic Strokes: Time for New Insight**  
*Marco Colasurdo, et al.*

### REVIEW ARTICLE

- 365 **Intraluminal Flow Diverter Design Primer for Neurointerventionalists** *Alexander A. Oliver, et al.*









NEUROINTERVENTION

### STATE OF PRACTICE

- 371 **Understanding Bias in Artificial Intelligence: A Practice Perspective** *Melissa A. Davis, et al.*

ARTIFICIAL INTELLIGENCE

### GENERAL CONTENTS

-  374 **Shift Volume Directly Impacts Neuroradiology Error Rate at a Large Academic Medical Center: The Case for Volume Limits**  
*Vladimir Ivanovic, et al.*
-   379 **Clinical Evaluation of a 2-Minute Ultrafast Brain MR Protocol for Evaluation of Acute Pathology in the Emergency and Inpatient Settings** *Min Lang, et al.*
-  386 **Increased Intracranial Arterial Pulsatility and Microvascular Brain Damage in Pseudoxanthoma Elasticum** *J.W. Bartstra, et al.*
-   393 **Hemorrhage Volume Drives Early Brain Injury and Outcome in Poor-Grade Aneurysmal SAH** *Pietro Panni, et al.*
-  400 **The Management of Persistent Distal Occlusions after Mechanical Thrombectomy and Thrombolysis: An Inter- and Intrarater Agreement Study** *W. Boisseau, et al.*
-  406 **A Clinical and Imaging Fused Deep Learning Model Matches Expert Clinician Prediction of 90-Day Stroke Outcomes** *Yongkai Liu, et al.*

HEALTH POLICIES/QUALITY IMPROVEMENT/EVIDENCE-BASED NEUROIMAGING

EMERGENCY NEURORADIOLOGY

NEUROVASCULAR/STROKE IMAGING

NEUROVASCULAR/STROKE IMAGING

NEUROVASCULAR/STROKE IMAGING

NEUROVASCULAR/STROKE IMAGING

AJNR (Am J Neuroradiol ISSN 0195-6108) is a journal published monthly, owned and published by the American Society of Neuroradiology (ASNR), 820 Jorie Boulevard, Oak Brook, IL 60523. Annual dues for the ASNR include approximately 19% for a journal subscription. The journal is printed by Intellicor Communications, 330 Eden Road, Lancaster, PA 17601; Periodicals postage paid at Oak Brook, IL and additional mailing offices. Printed in the U.S.A. POSTMASTER: Please send address changes to American Journal of Neuroradiology, P.O. Box 3000, Denville, NJ 07834, U.S.A. Subscription rates: nonmember \$475 (\$560 foreign) print and online, \$320 online only; institutions \$550 (\$625 foreign) print and basic online, \$1080 (\$1160 foreign) print and extended online, \$380 online only (basic), \$825 online only (extended); single copies are \$35 each (\$40 foreign). Indexed by PubMed/MEDLINE, BIOSIS Previews, Current Contents (Clinical Medicine and Life Sciences), EMBASE, Google Scholar, HighWire Press, Q-Sensei, RefSeek, Science Citation Index, SCI Expanded, ReadCube, and Semantic Scholar. Copyright © American Society of Neuroradiology.

-  412 Predictors of In-Stent Stenosis Following the Implantation of Pipeline Embolization Devices for the Treatment of Aneurysms Located at or beyond the Circle of Willis in the Anterior Circulation *Yajing Ma, et al.* **NEUROINTERVENTION**
-  418 Initial Experiences with the Trenza Embolization Device for the Treatment of Wide-Neck Intracranial Aneurysms: A 12-Patient Case Series *Rahul Raj, et al.* **NEUROINTERVENTION**
-  424 Evaluation of Conditions for the Development of Cryogenic Spinal Cord Injury Using a Canine Model: An Experimental Study on the Safety of Cryoablation for Metastatic Spinal Tumors *Motoya Kobayashi, et al.* **NEUROINTERVENTION**
- 432 Automated Detection of Cervical Spinal Stenosis and Cord Compression via Vision Transformer and Rules-Based Classification *David L. Payne, et al.* **SPINE IMAGING AND SPINE IMAGE-GUIDED INTERVENTIONS**
-  439 Identifying Patients with CSF-Venous Fistula Using Brain MRI: A Deep Learning Approach *Shahriar Faghani, et al.* **ARTIFICIAL INTELLIGENCE**
-  444 Compressed Sensitivity Encoding Artificial Intelligence Accelerates Brain Metastasis Imaging by Optimizing Image Quality and Reducing Scan Time *Mengmeng Wang, et al.* **ARTIFICIAL INTELLIGENCE**
-  453 Arterial Spin-Labeling and DSC Perfusion Metrics Improve Agreement in Neuroradiologists' Clinical Interpretations of Posttreatment High-Grade Glioma Surveillance MR Imaging—An Institutional Experience *Ghiam Yamin, et al.* **BRAIN TUMOR IMAGING**
- 461 Improved Detection of Target Metabolites in Brain Tumors with Intermediate TE, High SNR, and High Bandwidth Spin-Echo Sequence at 5T *Wenbo Sun, et al.* **BRAIN TUMOR IMAGING**
-  468 High-Grade Astrocytoma with Piloid Features: A Dual Institutional Review of Imaging Findings of a Novel Entity *Neetu Soni, et al.* **BRAIN TUMOR IMAGING**
-  475 Comparison of Volumetric and 2D Measurements and Longitudinal Trajectories in the Response Assessment of BRAF V600E-Mutant Pediatric Gliomas in the Pacific Pediatric Neuro-Oncology Consortium Clinical Trial *Divya Ramakrishnan, et al.* **PEDIATRIC NEUROIMAGING**
-  483 Expanding the Imaging Spectrum of Polymorphous Low-Grade Neuroepithelial Tumor of the Young in Children *Carmen Cerron-Vela, et al.* **PEDIATRIC NEUROIMAGING**
-  488 Prognostic Significance of Central Skull Base Remodeling in Chiari II Malformation *John T. Freiling, et al.* **PEDIATRIC NEUROIMAGING**
-  494 Neurofilament Light Chain Levels Interact with Neurodegenerative Patterns and Motor Neuron Dysfunction in Amyotrophic Lateral Sclerosis *Penelope Tilsley, et al.* **NEURODEGENERATIVE DISORDER IMAGING**
-  504 Visualizing the Habenula Using 3T High-Resolution MP2RAGE and QSM: A Preliminary Study *BingYang Bian, et al.* **NEUROIMAGING PHYSICS/FUNCTIONAL NEUROIMAGING/CT AND MRI TECHNOLOGY**
-  511 Does CISS MRI Reliably Depict the Endolymphatic Duct in Children with and without Vestibular Aqueduct Enlargement? *Olutayo I. Olubiyi, et al.* **HEAD AND NECK IMAGING**
-  518 Sinonasal (Schneiderian) Tumors in the Temporal Bone: Case Series and Systematic Review *Francis Deng, et al.* **HEAD AND NECK IMAGING**
- 525 Utility of MR Neurography for the Evaluation of Peripheral Trigeminal Neuropathies in the Postoperative Period *Tyler Thornton, et al.* **HEAD AND NECK IMAGING**

# **BOOK REVIEWS** *R.M. Quencer, Section Editor*

Please visit [www.ajnrblog.org](http://www.ajnrblog.org) to read and comment on Book Reviews.

At ultra-high fields, challenges emerge when measuring metabolites using <sup>1</sup>H-MRS. Sun, et al. showed that the high SNR and high bandwidth spin-echo (HISE) technique outperformed the clinical standard PRESS technique in detecting target metabolites of 23 pathologically confirmed brain tumors at 5T, particularly lactate and alanine. In this patient with an anterior cranial fossa atypical meningioma with invasion into the brain parenchyma, the HISE technique observed relatively strong Ala signal, while the PRESS technique could not differentiate between the Ala and Lac peaks.



Indicates Editor's Choices selection



Indicates Fellows' Journal Club selection



Indicates open access to non-subscribers at [www.ajnr.org](http://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



# AMERICAN JOURNAL OF NEURORADIOLOGY

Official Journal:

American Society of Neuroradiology  
American Society of Functional Neuroradiology  
American Society of Head and Neck Radiology  
American Society of Pediatric Neuroradiology  
American Society of Spine Radiology

APRIL 2024 • VOLUME 45 • NUMBER 4 • WWW.AJNR.ORG

## EDITOR-IN-CHIEF

**Max Wintermark, MD, MAS, FASFN, FICIS**

*Frank T. McGraw Memorial Chair in the Study of Cancer  
Professor and Chair of Neuroradiology  
The University of Texas MD Anderson Center*

## DEPUTY EDITOR

**Lubdha M. Shah, MD, MS**

Professor of Radiology and Director of Spine Imaging  
Department of Radiology and Imaging Sciences  
University of Utah

## SPECIAL ADVISORS TO THE EDITOR-IN-CHIEF

**Mauricio Castillo, MD, FACR**

MA Mauro Distinguished Professor of Radiology  
University of North Carolina-Chapel Hill

**Robert Quencer, MD**

Professor Emeritus  
Department of Radiology  
University of Miami

## ARTIFICIAL INTELLIGENCE

### Senior Editor

**Reza Forghani, MD, PhD**

Professor of Radiology & Artificial Intelligence  
Vice Chair of AI  
Director, Radiomics & Augmented Intelligence Laboratory (RAIL)  
Department of Radiology  
University of Florida College of Medicine

### Associate Editors

**Andreas Rauschecker, MD, PhD**

Assistant Professor-in-Residence  
Co-Executive Director and Clinical Director,  
Center for Intelligent Imaging (ci<sup>2</sup>)  
Department of Radiology & Biomedical Imaging  
University of California, San Francisco

**Sam (Seyedmehdi) Payabvash, MD**

Assistant Professor of Radiology  
Yale School of Medicine  
Connecticut

## BRAIN TUMOR IMAGING

### Senior Editor

**Ben Ellingson, PhD**

Professor and Director of MRI Research  
Director, UCLA Brain Tumor Imaging Laboratory  
Department of Radiological Sciences  
David Geffen School of Medicine  
University of California, Los Angeles

### Associate Editors

**Ali Nabavizadeh, MD**

Assistant Professor of Radiology  
Division of Neuroradiology  
University of Pennsylvania

**Mark S. Shiroishi, MD, MS, FASFN**

Assistant Professor, Division of Neuroradiology,  
Department of Radiology  
Director of Neuro-Oncology Imaging - USC Brain Tumor Center  
Chief of Pediatric Neuroradiology - Los Angeles General Medical Center  
Affiliated Faculty - USC Imaging Genetics Center  
Mark and Mary Stevens Neuroimaging and Informatics Institute  
Keck School of Medicine of USC  
University of Southern California

## EMERGENCY NEURORADIOLOGY

### Senior Editor

**Karen Buch, MD**

Assistant Professor of Radiology  
Massachusetts General Hospital

### Associate Editors

**Melissa A. Davis, MD, MBA**

Vice Chair of Medical Informatics  
Associate Professor  
Department of Radiology and Biomedical Imaging  
Yale School of Medicine  
Connecticut

**Jason Talbott, MD, PhD**

Associate Professor, Neuroradiology Section  
Department of Radiology and Biomedical Imaging  
University of California, San Francisco and  
Zuckerberg San Francisco General Hospital

## HEAD AND NECK IMAGING

### Senior Editor

**Amy Juliano, MD**

Associate Professor of Radiology  
Massachusetts Eye and Ear  
Harvard Medical School

### Associate Editors

**Burce Ozgen, MD**

Clinical Professor of Radiology  
University of Illinois at Chicago

**David Zander, MD**

Assistant Professor, Radiology  
University of Colorado School of Medicine

## HEALTH POLICIES/QUALITY IMPROVEMENT/ EVIDENCE-BASED NEUROIMAGING

### Senior Editor

**Nadja Kadom, MD, FACR, FAAP**

Professor, Department of Radiology and Imaging Sciences  
Emory University School of Medicine  
Pediatric Neuroradiologist, Department of Radiology  
Children's Healthcare of Atlanta

## Associate Editors

**Melissa M. Chen, MD**

Associate Professor,  
Department of Neuroradiology, Division of Diagnostic Imaging  
The University of Texas MD Anderson Center

**Ajay Malhotra, MBBS, MD, MMM**

Professor of Radiology and Biomedical Imaging and Neurosurgery  
Yale School of Medicine  
Connecticut

## MOLECULAR NEUROIMAGING/NUCLEAR MEDICINE

### Senior Editor

**Ana M. Franceschi, MD, PhD**

Associate Professor of Radiology  
Donald and Barbara Zucker School of Medicine at Hofstra/Northwell  
The Feinstein Institutes for Medical Research  
Neuro-PET Imaging  
Lenox Hill Hospital  
New York

### Associate Editors

**Nadya Pyatigorskaya, MD, PhD**

Neuroradiology Department, Pitié-Salpêtrière Hospital  
Researcher, Brain Institute (ICM)  
France

**Marc Daniel Benayoun, PhD, MD**

Assistant Professor of Radiology  
Co-Section Head of Nuclear Medicine  
Medical Director of Nuclear Medicine and PET  
Vice Chair of Radiation Drug Research Committee  
Atrium Wake Forest Health  
North Carolina

## NEURODEGENERATIVE DISORDER IMAGING

### Senior Editor

**Gloria Chiang, MD**

Associate Professor, Co-Director of the Brain Health Imaging Institute  
Department of Radiology  
Weill Cornell Medicine/NewYork-Presbyterian Hospital

### Associate Editors

**Fang Frank Yu, MD**

Assistant Professor of Radiology  
Division of Neuroradiology, Department of Radiology  
Advanced Imaging Research Center  
University of Texas Southwestern Medical Center



**Priya Rajagopalan, MBBS, MPH**

Assistant Professor of Radiology, Division of  
Neuroradiology  
Associate Program Director, Neuroradiology  
Fellowship  
Medical Director, Center for Image Acquisition  
Mark and Mary Stevens Neuroimaging and  
Informatics Institute  
Keck School of Medicine, University of Southern  
California

**NEUROIMAGING PHYSICS/FUNCTIONAL  
NEUROIMAGING/CT AND MRI TECHNOLOGY****Senior Editor****Hongyu An, DSc**

Professor, Mallinckrodt Institute of Radiology  
Neurology, Biomedical Engineering, Electrical and  
Systems Engineering, Division of Biology and  
Biomedical Sciences  
Director, Biomedical Magnetic Resonance Center  
Associate Director, Center for Clinical Imaging Research  
Washington University in St. Louis

**Associate Editors****Timothy J. Carroll, PhD**

Professor, Department of Radiology  
University of Chicago

**Hugo de Jong, PhD**

Professor of Medical Physics Radiology and  
Nuclear Medicine  
UMC Utrecht  
the Netherlands

**NEUROINTERVENTION****Senior Editor****Steven Hetts, MD, FACP**

Co-Chief, NeuroEndovascular Surgery Service Line  
Chief of Interventional Neuroradiology, Mission  
Bay Hospitals  
Professor of Radiology, Biomedical Imaging, and  
Neurological Surgery  
University of California, San Francisco

**Associate Editors****Kristine Blackham, MD**

Associate Professor  
Diagnostic and Interventional Neuroradiology  
Clinic for Radiology and Nuclear Medicine  
University Hospital of Basel

**Maksim Shapiro, MD**

Clinical Associate Professor  
Departments of Radiology, Neurosurgery, and  
Neurology  
Division of Neurointerventional Radiology  
NYU Langone Health and Bellevue NYCH +  
Hospitals  
New York

**NEUROPSYCHIATRIC IMAGING****Senior Editor****Jody Tanabe, MD**

Professor  
Chief of Neuroradiology  
University of Colorado-Anschutz Medical Campus

**Associate Editors****John-Paul J. Yu, MD, PhD**

Assistant Professor of Radiology, Psychiatry, and  
Biomedical Engineering  
University of Wisconsin School of Medicine and  
Public Health

**J. Eric Schmitt, MD, PhD**

Assistant Professor of Radiology and Psychiatry  
Division of Neuroradiology  
Perelman School of Medicine, University of  
Pennsylvania

**NEUROVASCULAR/STROKE IMAGING****Senior Editor****Ajay Gupta, MD, MS**

Professor and Chair, Department of Radiology  
Columbia University  
New York

**Associate Editors****Shalini Amukotuwa, MB BS, PhD,  
FRANZCR**

Head of Neuroradiology and Director of MRI,  
Monash Health  
Associate Professor of Radiology, Monash University  
Australia

**Mahmud Mossa-Basha, MD**

Professor of Radiology, Neurology and Electrical  
Engineering  
Vice Chair of Clinical Research and Clinical  
Transformation  
Co-Director of the Research Vascular Imaging Lab  
University of Washington School of Medicine

**PEDIATRIC NEUROIMAGING****Senior Editor****Caroline D. Robson, MBChB**

Division Chief & Endowed Chair, Neuroradiology  
Director, Head & Neck Imaging  
Department of Radiology  
Department of Otolaryngology  
Boston Children's Hospital  
Harvard Medical School

**Associate Editors****Anna Trofimova, MD, PhD**

Assistant Professor, Radiology and Imaging  
Sciences, Emory University  
Pediatric Neuroradiologist, Children's Healthcare  
of Atlanta

**Matthew Whitehead, MD**

Department of Radiology, Division of  
Neuroradiology  
Children's Hospital of Philadelphia  
Associate Professor of Radiology  
Perelman School of Medicine, University of  
Pennsylvania

**SPINE IMAGING AND SPINE IMAGE-GUIDED  
INTERVENTIONS****Senior Editor****J. Levi Chazen, MD**

Associate Professor, Neuroradiology  
Director, Spine Imaging  
Hospital for Special Surgery  
Weill Cornell Medicine  
New York

**Associate Editors****Jennifer McCarty, MD**

UTHealth Houston

**Vinil Shah, MD**

Associate Professor of Radiology Neuroradiology  
Division Chief  
University of California, San Francisco

**ULTRA-HIGH-FIELD MRI/IMAGING OF  
EPILEPSY/DEMYELINATING DISEASES/  
INFLAMMATION/INFECTION****Senior Editor****Erik Middlebrooks, MD**

Professor of Radiology  
Mayo Clinic Florida

**Associate Editors****Susie Y. Huang, MD, PhD**

Associate Professor of Radiology, Harvard  
Medical School  
Associate Chair, Faculty Affairs, Department of  
Radiology  
Director of Translational Neuro MR Imaging &  
Connectomics, Athinoula A. Martinos Center for  
Biomedical Imaging  
Massachusetts General Hospital

**Girish Bathla, MD, FRCR**

Associate Professor, Neuroradiology  
Mayo Clinic  
Minnesota

**OUTREACH AND EDUCATION****Senior Editor****Lea Alhilali, MD**

Radiology Partners, HonorHealth Research  
Institute  
Arizona

**DIGITAL MEDIA AND ENGAGEMENT****Senior Editor****Kevin Hsu, MD**

Clinical Assistant Professor  
Department of Radiology  
NYU Grossman School of Medicine  
New York

**Case Collection Editors****Matylda Machnowska, BMedSc, MD,  
FRCSC, ABR**

Assistant Professor of Radiology, University of  
Toronto  
Neuroradiologist, Sunnybrook Health Sciences  
Centre

**Anvita Pauranik, MD**

Clinical Assistant Professor of Radiology  
BC Children's Hospital  
University of British Columbia

**Sandy Cheng-Yu Chen, MD**

Chair, Translational Imaging Research Center  
Taipei Medical University Hospital  
Vice President of Taipei Medical University

**Social Media Editor****Kimberly Seifert, MD, MS**

Stanford University  
California

**Assistant Social Media Editors****Ani Hoxha, MD**

Mother Theresa UHC  
Albania

**Maxwell Opoku, MD**

Novosibirsk State Research University  
Russia

#### Podcast Editor

##### **Kevin Hiatt, MD**

Assistant Professor, Radiology  
Wake Forest University School of Medicine  
North Carolina

#### Deputy Podcast Editor

##### **George K. Vilanilam, MD**

PGY5 Resident Physician, Department of Radiology  
University of Arkansas for Medical Sciences

#### STATISTICS

##### Senior Editor

##### **Bryan A. Comstock, MS**

Senior Biostatistician, Department of Biostatistics  
University of Washington

#### EDITORIAL FELLOWS

##### **Alexandre Boutet, MD, PhD**

Neuroradiologist  
Joint Department of Medical Imaging  
University of Toronto

##### **Nicholas S. Cho, MD/PhD Candidate**

University of California, Los Angeles

##### **Burak Berksu Ozkara, MD**

Research Fellow  
Department of Neuroradiology  
The University of Texas MD Anderson Center

#### BOOK REVIEW EDITOR

##### **Robert Quencer, MD**

Professor Emeritus  
Department of Radiology  
University of Miami

---

*Founding Editor*  
**Juan M. Taveras**

*Editors Emeriti*  
Mauricio Castillo, Robert I. Grossman,  
Michael S. Huckman, Robert M. Quencer,  
Jeffrey S. Ross

---

*Managing Editor*  
**Karen Halm**  
*Assistant Managing Editor*  
**Laura Wilhelm**  
*Executive Director, ASNR*  
**Mary Beth Hepp**

---

*Digital Publications Senior Manager*  
**Gwen Travis**

# Navigating Change: *AJNR*'s Transition to an Electronic Format

Dear *AJNR* Reader, Author, and/or Reviewer,

I trust this letter finds you well and that you have appreciated the recent changes we have introduced in the *American Journal of Neuroradiology* (*AJNR*), including an expanded Editorial Board with comprehensive expertise, an Academy of Reviewers, a streamlined submission process, a shortened turnaround review time for articles submitted, and a shorter time from acceptance to publication.

I am writing to share an important update regarding the future of the *American Journal of Neuroradiology*.

Toward the end of the year, *AJNR* will undergo a significant transformation, embracing a fully electronic format and bidding farewell to its traditional paper publication, like many radiology journals have done before us. This strategic shift is driven by the need to manage escalating costs associated with printing, shipping, and storage. Our goal is to achieve fiscal equilibrium without imposing substantial increases in article-processing charges for our valued authors.

The transition to an electronic milieu heralds numerous advantages for both authors and readers. Specifically, it introduces the exciting prospect of increased and more personalized content, allowing readers to tailor their *AJNR* experience to their individual preferences. The electronic format will also facilitate a richer and more immersive visual experience by accommodating more images and incorporating a broader array of media types, including videos. Readers may even have the ability to scroll through stacks of images, a particularly valuable feature in the field of neuroradiology.

Moreover, the electronic format enables the addition of secondary content, enhancing the reader's experience by providing

seamless access to supplementary materials, commentaries, and related resources. Our aim is to offer a comprehensive understanding of *AJNR* content. We are also committed to implementing an enhanced searchability feature in our revamped electronic journal, catering to the diverse needs of practicing neuroradiologists, researchers, trainees, and students.

As an avid enthusiast of printed books and a devoted collector of physical copies of *AJNR* for more than 2 decades, I am touched by a sense of sadness at the realization that the cherished tradition of receiving *AJNR* in my mailbox will come to an end. Undoubtedly, I anticipate that many of our printed journal readers will resonate with these sentiments. Nevertheless, financial pragmatism impels us to embrace the technological advancements that represent the future of the publishing industry and to navigate this transition with optimism and adaptability.

In the spirit of this transition, *AJNR* recognizes the importance of engaging our readership, authors, reviewers, and all stakeholders. Plans are underway to organize sessions aimed at gathering valuable feedback, concerns, and innovative ideas. Through fostering an open dialogue, *AJNR* aspires to sculpt its electronic platform to align with the evolving expectations of our audience.

In embracing the digital evolution, *AJNR* not only adjusts to the current academic publishing landscape and embraces sustainability goals but also moves toward a more dynamic and interactive future in neuroradiologic literature. I extend a warm invitation to our readers, authors, and reviewers to partake in this journey, contributing your insights to redefine the boundaries of scholarly communication in the realm of neuroradiology.

Thank you in advance for your engagement.

Sincerely,

Max Wintermark, Editor-in-Chief

# Introducing Our New *AJNR* Early Career, Women in Neuroradiology, and Global Neuroradiology Awards

Nicholas S. Cho, Burak Berkus Ozkara, Alexandre Boutet, Karen Halm, Laura Wilhelm, Lubdha Shah, and Max Wintermark

There is much to celebrate at the *American Journal of Neuroradiology* (*AJNR*). Thanks to the strong efforts of our recently appointed Editors<sup>1</sup> and dedicated reviewer pool, the *AJNR* has now reduced its median time from submission to first decision after peer review to an astounding 17 days. We have also launched our Academy of Reviewer resources<sup>2</sup> to aid our reviewers in the peer review process. Moreover, we published our first electronic “preprints” of accepted manuscripts thanks to the adoption of our *AJNR* manuscript templates by the contributing authors.<sup>3</sup> These preprints are indexed and citable 1–3 business days after acceptance to facilitate rapid dissemination of our articles. Today, we also have the opportunity to recognize some of the finest articles published in the *AJNR* with the 2024 *AJNR* journal award nominees.

We have a long-standing annual tradition at the *AJNR* of awarding the *AJNR* Lucien Levy Best Research Article Award to an original research article with the highest impact on our field. This award is named after the late Dr Levy, MD, PhD, Senior Editor for the *AJNR* from 2007 to 2013.<sup>4</sup> We eagerly await continuing this tradition and awarding the 2024 recipient at ASNR24 this May in Las Vegas. In addition, we are also pleased to announce the creation of 3 new, additional awards to celebrate the high-impact research articles specifically from our diverse authorship: 1) *AJNR* Early Career Award, 2) *AJNR* Women in Neuroradiology Award, and 3) *AJNR* Global Neuroradiology Award.

We believe that our *AJNR* accolades should celebrate the accomplishments of the multitude of voices within our neuroradiology community. For our *AJNR* Early Career Award, we want to recognize our trainees and young faculty because they are the future of our field. This recognition will empower them at the start of their careers in their path towards becoming future leaders in neuroradiology. Our *AJNR* Women in Neuroradiology Award acknowledges the leading female voices in our field. Because women remain underrepresented in neuroradiology

across all levels, spanning from the trainee level (~20% are women)<sup>5</sup> to academic leadership ranks (~13% are women),<sup>6</sup> their achievements are often overlooked. Last, with the *AJNR* Global Neuroradiology Award, we want to recognize our colleagues beyond North America as well as our growing number of international medical graduates<sup>5</sup> and visiting scholars who come to our country to make great contributions to our field.

Importantly, we also want to hear all your voices when selecting the 2024 *AJNR* journal awards. We will now include a public voting process for our 4 journal awards recognizing the *AJNR* articles with high impact on our field. The nominees for all awards and instructions to vote can be found at the *AJNR* website: [ajnr.org/awards](http://ajnr.org/awards). We will announce the winning articles at ASNR24 this May at our *AJNR* booth in the exhibit hall. We encourage you to visit our *AJNR* booth to celebrate our awardees and meet our new Editorial Board to share any ideas you may have for the *AJNR*.

In closing, we would like to thank the nominated authors for their great contributions, the reviewers for their valuable time, and you for your loyal readership. Congratulations to all the nominees, and we look forward to awarding our 2024 *AJNR* Lucien Levy Best Research Article Award as well as our inaugural *AJNR* Early Career, Women in Neuroradiology, and Global Neuroradiology Awards at the ASNR24 Annual Meeting in Las Vegas this May!

## REFERENCES

1. Wintermark M. A new era dawns at *AJNR*: revamping the editorial board and streamlining the review process. *AJNR Am J Neuroradiol* 2023;44:1241 CrossRef Medline
2. Wintermark M. Introducing *AJNR* SCANTastic and the *AJNR* Academy of Reviewers. *AJNR Am J Neuroradiol* 2023;44:1357 CrossRef
3. Wintermark M. Streamlining the manuscript submission process at the *American Journal of Neuroradiology*: paving the way for efficiency and reducing hurdles for authors. *AJNR Am J Neuroradiol* 2024;45:249 CrossRef Medline
4. Tu RK, Bhojwani R, Taheri R, et al. Lucien Maurice Levy, MD, PhD. *Radiology* 2015;275:623 CrossRef Medline
5. Prajapati P, Wu X, Bajaj S, et al. Trends of diversity in neuroradiology trainees in United States 2015–2022. *AJNR Am J Neuroradiol* 2023;44:1009–11 CrossRef Medline
6. Ahmadi M, Khurshid K, Sanelli PC, et al. Influences for gender disparity in academic neuroradiology. *AJNR Am J Neuroradiol* 2018;39:18–23 CrossRef Medline



# Imaging Large Ischemic Strokes: Time for New Insight

Marco Colasurdo, Huanwen Chen, and Dheeraj Gandhi

Six randomized, multicenter trials demonstrated a clinical benefit for patients with low ASPECTS regardless of the amount of the penumbra on advanced perfusion imaging. These findings not only question the utility of neuroimaging and advanced techniques to identify infarcted tissue, but they also profoundly challenge the foundation of our understanding of hypoperfusion and its reversibility. Now more than ever, there is tremendous need for better neuroimaging tools to reliably identify patients with large ischemic strokes who will benefit from endovascular treatment.

Recently, 6 large, multicenter, randomized controlled trials<sup>1-6</sup> have concluded their investigations on the efficacy and safety of endovascular thrombectomy (EVT) for patients presenting with large ischemic strokes seen on initial neuroimaging. Overall, the 6 trials demonstrated consistent clinical benefit for patients with a low ASPECTS of 3–5, with subgroup analyses also showing the benefit of EVT for even lower ASPECTS (0–2) and large infarct volumes seen on perfusion imaging or MR imaging (>200 mL), regardless of the amount of stroke penumbra. These results seem to suggest that re-establishing blood flow to areas of the brain labeled as “ischemic core” (which was previously thought to be infarcted) may yield clinically significant benefit, and they have major implications for the role of neuroimaging during acute stroke triage and management.

Historically, neuroimaging played a critical role in the understanding of ischemic brain diseases. CT revealed the brain as a digital craniotome, and the detailed tissue resolution provided by MR imaging yielded great insight into stroke pathophysiology. When it became clear that patients with strokes have limited time before tissue ischemia progresses to infarct and neuronal death, neuroimaging techniques were used to provide real-time assessment of tissue viability. To this end, the identification of early ischemic changes on noncontrast CT was proposed, and the quasi-quantitative ASPECTS system was developed. Subsequently, quantitative analyses using CT perfusion and MR imaging made possible the identification of already-infarcted tissue (also termed the ischemic core) and symptomatic-but-salvageable ischemic tissue (also termed the “stroke penumbra”), and recent advances in artificial intelligence ushered these advanced imaging tools into routine clinical use. Some hospital systems also devised specialized protocols to allow rapid MR imaging, which provides additional tissue resolution to quantify the approximate age and extent of ischemic lesions during acute stroke triage. These advanced imaging modalities were part of the main inclusion and exclusion criteria for numerous landmark stroke trials, and they laid the foundation for the expanding use of both IV thrombolysis (via the Efficacy and Safety of the MRI-Based Thrombolysis in Wake-Up Stroke [WAKE-UP] and Extending the Time for Thrombolysis in

Emergency Neurological Deficits [EXTEND] trials) and EVT (via the Solitaire With the Intention For Thrombectomy as PRiMary Endovascular Treatment for Acute Ischemic Stroke [SWIFT PRIME], Extending the Time for Thrombolysis in Emergency Neurological Deficits-Intra-Arterial [EXTEND-IA], Clinical Mismatch in the Triage of Wake Up and Late Presenting Strokes Undergoing Neurointervention with Trevo [DAWN], and Endovascular Therapy Following Imaging Evaluation for Ischemic Stroke 3 [DEFUSE 3] trials).

The utility of advanced neuroimaging during acute stroke triage was largely predicated on the notion that stroke therapy does not always lead to successful recanalization of occluded vessels, with substantial hemorrhagic and procedural risks. Thus, it was thought that aggressive treatment for larger ischemic cores and smaller penumbras may expose patients to minimal benefit and excessive harm, particularly for EVT. The first endovascular treatment for acute intracranial occlusions can be dated to the 1980s, mostly using intra-arterial thrombus-dissolving medications. For about 25 years, these procedures were mostly unsuccessful until the first mechanical thrombectomy devices were introduced (Merci retriever, Concentric Medical; and Penumbra Separator, Penumbra). Since 2015, technical advancements from stent retrievers to direct aspiration catheters to balloon guide catheters have been concatenating incredibly quickly. Research behind these devices moves so quickly that the US FDA had to establish guidelines specifically regulating preclinical and clinical studies for neurothrombectomy tools, first in 2007,<sup>7</sup> then updated in 2016.<sup>8</sup>

The remarkable advances in neurothrombectomy devices culminated in high rates of successful recanalization and low rates of hemorrhagic transformations in the recently published trials on large-core thrombectomies, and the consistent efficacy signal observed across the 6 trials essentially upends the original value proposition of advanced neuroimaging during acute stroke triage. Thus, we are now entering an era of crisis in which we are witnessing a reverse trend in which the so-called “advanced imaging” techniques are set aside in favor of noncontrast head CT or even direct-to-thrombectomy protocols.

Despite the proved benefit of EVT in multiple trials, not all patients will benefit from reperfusion. The pathophysiologic umbra-penumbra dogma was originally proposed by Astrup et al,<sup>9</sup> in which there is a certain volume of tissue irreversibly damaged defined as “core” and a certain volume of tissue that is oligemic defined as “penumbra.” This artificial construct was a valuable model during the development of acute stroke reperfusion treatments. Volumetric analyses of core and penumbra have been reliably associated with patient prognosis following revascularization, suggesting that the umbra-penumbra dogma likely holds true in a general sense. However, this model is likely an oversimplification of true pathophysiology, which likely involves many more nuances in the dynamic macrovascular and microvascular environment during acute ischemic stroke. Thus, while the umbra-penumbra dogma can provide valuable information on neuroprognosis following EVT, it is likely insufficient for predicting the treatment effect

of EVT on each individual patient and may not be as well-suited for selecting patients for EVT treatment as previously believed. Our foothold at this time is that we know that at a certain point, ischemic brain tissue will fully infarct if blood flow is not restored. Thus, a major challenge for treatment decisions in the future will be to better identify when tissue is truly no longer viable and therefore will not benefit from reperfusion. We must concede that current neuroimaging tools for acute stroke triage (ASPECTS, CT perfusion, and DWI) are inconsistent and largely unable to accurately predict the extent of unsalvageable stroke damage.

The high rates of poor outcomes in the EVT arms of the recently published trials for large strokes (close to 80% death or dependency) suggest that treating all patients with large-vessel occlusion strokes regardless of tissue status is a crude strategy at best, and doing so will inevitably strain the existing neuroendovascular workforce and jeopardize precious health care resources for patients more likely to benefit from treatment. Thus, there remains tremendous foundational research opportunities for better neuroradiologic technologies to further optimize stroke triage, focusing on the need to devise new imaging constructs specifically with the aim of predicting treatment effect and improving patient outcomes. It is, therefore, our belief that the recently published trial results should not be interpreted as a setback for advanced stroke imaging, but instead, they should be viewed as a call for further development and refinement of neuroimaging tools to rapidly assess the viability of ischemic stroke tissue.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

1. Yoshimura S, Sakai N, Yamagami H, et al. **Endovascular therapy for acute stroke with a large ischemic region.** *N Engl J Med* 2022;386:1303–13 CrossRef Medline
2. Zaidat OO, Kasab SA, Sheth S, et al. **TESLA trial: rationale, protocol, and design.** *Stroke: Vascular and Interventional Neurology* 2023;3:e000787 CrossRef
3. Bendszus M, Fiehler J, Subtil F, et al; TENSION Investigators. **Endovascular thrombectomy for acute ischaemic stroke with established large infarct: multicentre, open-label, randomised trial.** *Lancet* 2023;402:1753–63 CrossRef Medline
4. Huo X, Ma G, Tong X, et al; ANGEL-ASPECT Investigators. **Trial of endovascular therapy for acute ischemic stroke with large infarct.** *N Engl J Med* 2023;388:1272–83 CrossRef Medline
5. Sarraj A, Hassan AE, Abraham MG, et al; SELECT2 Investigators. **Trial of endovascular thrombectomy for large ischemic strokes.** *N Engl J Med* 2023;388:1259–71 CrossRef Medline
6. Costalat V, Lapergue B, Albucher JF, et al; LASTE Trial Investigators. **Evaluation of acute mechanical revascularization in large stroke (ASPECTS <5) and large vessel occlusion within 7 h of last-seen-well: the LASTE multicenter, randomized, clinical trial protocol.** *Int J Stroke* 2023;19:114–19 CrossRef Medline
7. Center for Devices and Radiological Health. **Pre-Clinical and Clinical Studies for Neurothrombectomy Devices.** March 19, 2020. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/pre-clinical-and-clinical-studies-neurothrombectomy-devices>. Accessed October 23, 2023
8. CFR. **Code of Federal Regulations Title 21.** <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?fr=882.5600>. Accessed October 23, 2023
9. Astrup J, Symon L, Branston NM, et al. **Cortical evoked potential and extracellular K<sup>+</sup> and H<sup>+</sup> at critical levels of brain ischemia.** *Stroke* 1977;8:51–57 CrossRef Medline

# Intraluminal Flow Diverter Design Primer for Neurointerventionalists

Alexander A. Oliver, Yigit Can Senol, Cem Bilgin, Jeremy E. Schaffer, Ramanathan Kadirvel, David F. Kallmes, and John M. Wainwright

## ABSTRACT

**SUMMARY:** The clinical use of flow diverters for the treatment of intracranial aneurysms has rapidly grown. Consequently, the market and technology for these devices has also grown. Clinical performance characteristics of the flow diverter are well-known to the clinician. However, the engineering design principles behind how these devices achieve ideal clinical performance are less understood. This primer will summarize flow diverter design parameters for neurointerventionalists with the aim of promoting collaboration between clinicians and engineers.

**ABBREVIATIONS:** DFT = drawn filled tubing; FD = flow diverter; ID = inner diameter; PPI = pics-per-inch

During the past decade, flow diverters (FDs) have grown into a standard endovascular approach for the treatment of intracranial aneurysms. FDs have demonstrated excellent safety profiles and aneurysm occlusion rates in the clinic.<sup>1</sup> Consequently, several FDs have achieved market approval, and many more devices are in development.<sup>2</sup> While the clinical performance characteristics of FDs are well-known to the clinician, how the FD engineer achieves successful device performance is less well-understood. This primer will outline the basic FD design parameters available to the engineer and describe how their selection influences the characteristics of device performance for the practicing interventionalist.

## Ideal FD

The generic term “flow diverter” describes devices that cover the aneurysm/parent artery interface to achieve aneurysm occlusion. Initially after placement, the device likely causes some degree of “diversion” of flow into the parent artery that otherwise would have entered the aneurysm, and the term itself is nonspecific and likely no more valid than terms such as “flow disruptor.”

However, with time, tissue will grow over the aneurysm neck and the aneurysm will become occluded, and ultimately, the end goal is that all flow will be diverted away from the aneurysm sac. Furthermore, there now are both intraluminal and intrasaccular flow diversion devices; this review will focus exclusively on the former type.

In general, FDs are braided or woven stents deployed in the parent artery over the neck of the aneurysm. These devices aim to divert most of the blood flow past the aneurysm, resulting in blood stagnation and coagulation within the aneurysm sac. This diversion of flow with resultant intra-aneurysmal blood stasis theoretically forms a thrombus plug within the aneurysm cavity, hindering blood flow into the aneurysm. In the meantime, the device serves as a scaffold for endothelialization and neointimal growth over the aneurysm neck, ultimately occluding the aneurysm from blood flow.<sup>3</sup> Overall, the ideal FD maximizes ease of deployment and aneurysm occlusion while minimizing the risk of complications.

Several interrelated FD performance characteristics contribute to clinical efficacy. The ideal FD maximizes stagnation of blood flow within the aneurysm sac as well as rapid endothelialization and neointimal growth over the aneurysm neck.<sup>3</sup> The major design goals are listed here briefly and then discussed in detail below in relation to the engineering principles underlying design choices. First, optimized FD apposition to the parent artery wall after delivery benefits both saccular blood stagnation and endothelialization.<sup>4,5</sup> Additionally, tight FD wall apposition can reduce the risk of thrombus formation between the wires of the device and the arterial wall, which could lead to the blockage of perforator arteries and stenosis. Better FD wall apposition can be achieved by improving the flexibility of the device<sup>6</sup> and, in some

Received August 8, 2023; accepted after revision October 29.

From the Department of Biomedical Engineering and Physiology (A.A.O., D.F.K.), Mayo Clinic Graduate School of Biomedical Sciences, Rochester, Minnesota; Departments of Radiology (A.A.O., C.B., R.K., D.F.K.), Neurologic Surgery (Y.C.S., R.K.), Mayo Clinic, Rochester, Minnesota; Fort Wayne Metals (J.E.S.), Fort Wayne, Indiana; MIVI Neuroscience (J.M.W.), Eden Prairie, Minnesota; and The Henry Samueli School of Engineering (J.M.W.), University of California, Irvine, California.

This work was partially funded by National Institutes of Health grant Nos. R01 NS076491 and R21 NS128199. Alexander Oliver is supported by American Heart Association grant No. 23PRE1012781.

Please address correspondence to Alexander A. Oliver, MD, Mayo Clinic, Radiology, 200 1st St SW, Rochester, MN 55905-0002; e-mail: Oliver.Alexander@mayo.edu  
<https://dx.doi.org/10.3174/ajnr.A8076>

		Flow Diverter Performance Characteristics						
		Endothelialization Across Aneurysm	Wall Apposition	Opening Force <sup>a</sup>	Risk of Thromboembolism	Risk of Perforator Artery Occlusion	Catheter ID	Delivery Force
Flow Diverter Design Parameters	↑Elastic Modulus	↔	↔ <sup>b</sup>	↑	↔	↔	↑	↑
	↑Wire Diameter	↓	↔ <sup>b</sup>	↑	↑	↑	↑	↑
	↑Number of Wires	↑	↔	↑	↑	↑	↑	↑
	↑Pics per Inch	↑	↑	↑	↑	↑	↑	↑
	Oversizing	↔	↑	↑	↔	↔	↑	↑
	↑Number of Layers	↑	↓	↑	↑	↑	↑	↔

**FIG 1.** The relationship between FD design parameters and resulting performance characteristics. *Green and red arrows* indicate a beneficial or detrimental relationship, respectively, in terms of device performance and safety. *Black horizontal arrows* indicate either an independent or multifactorial relationship between the design parameters and performance characteristics. *A*, Improved opening force assumes that the wires within the braid do not plastically deform after being crimped in their microcatheters, allowing complete device re-expansion. *B*, Wall apposition is complex and depends on many variables, including the tortuosity of the parent vessel.

cases, the radial opening force (or chronic outward force) of the device, which is the force exerted by the device expanding outward on the artery wall.<sup>7</sup> In straighter vessels, increased opening force theoretically pushes the wires tighter to the vessel wall. However, in tortuous anatomies, device flexibility may be more advantageous than radial opening force for maintaining wall apposition. Therefore, wall apposition is ultimately the result of numerous factors. Second, the ideal FD minimizes the risk of complications such as thromboembolism induced within the lumen of the device or the occlusion of perforator arteries adjacent to the aneurysm. Third, the ideal FD eases delivery for the interventionalist.

FDs that can be deployed through smaller inner diameter (ID) microcatheters are advantageous for advancing the microcatheter through the tortuous cerebrovasculature and may allow their advancement to more distal aneurysms. The current benchmark for FD delivery microcatheters is an ID of  $\leq 0.027$  inches,<sup>2</sup> and the ideal FD should be deployed through a microcatheter this size or smaller. Reduced delivery force, which we are defining as the force required to track the FD through the lumen of the microcatheter, also eases delivery for the interventionalist. The ideal FD achieves suitable wall apposition as easily as possible for the interventionalist.<sup>8,9</sup> Last, the FD needs to achieve a suitable degree of radiopacity to allow fluoroscopically guided delivery without compromising its MR imaging safety or inducing excessive CT or MR imaging artifacts, which may preclude noninvasive follow-up imaging.

### FD Design Parameters

Generally, FDs are constructed by braiding a set of 36–96 wires around a mandrel and heat-treating them to lock in the shape.

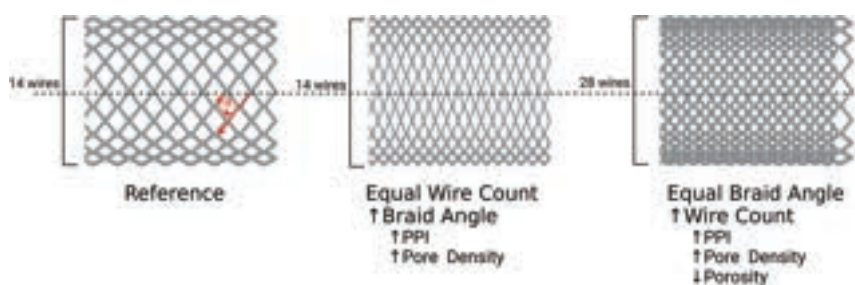
The mandrel is then removed, leaving a metal mesh tube that is cut to the desired length. The device may then be further processed by securing the loose wire ends, adding radiopaque markers, and/or treating the device surface. Finally, the device is loaded into its delivery system. While the general construction of FDs is simple in concept, there are many FD design parameters, including the wire diameter, the number of wires in the braid, the selection of wire material, and others, which all contribute to device performance. The selection of 1 design parameter may positively impact one FD performance characteristic while negatively impacting another. The FD engineer is tasked with selecting parameters to find a balance that optimizes the performance characteristics that the treating interventionalists find most important. The relationships between FD design parameters and effects on performance are summarized in Fig 1.

### FD Braid Geometry

There are many FD braid geometry design parameters available for the engineer to tune. The first is the diameter of the individual wires within the braid. Increasing the wire diameter increases the overall strength of the device, improving the opening force and radial crush resistance.<sup>10,11</sup> However, this improved opening force comes at the cost of increased delivery force and may require a larger ID microcatheter for delivery. A larger wire diameter may be associated with delayed endothelialization and the risk of thromboembolism, as is the case for much larger coronary stent struts.<sup>12</sup> However, this relationship is yet to be extensively studied in the FD application. Market-approved FDs typically have wire diameters in the  $\sim 18$ – $35\ \mu\text{m}$  ( $\sim 0.0007$ – $0.001$  inch) range.<sup>2</sup>

Another design parameter is the number of wires within the braid. As with the wire diameter, increasing the number of wires





**FIG 2.** Illustration of the relationship between the braid angle (defined by the red  $\theta$ ) and the number of wires or PPI and pore density. This figure was made with BioRender (<https://www.biorender.com/>).

within the braid will increase the opening force<sup>13</sup> of the FD at the cost of delivery force, potentially requiring a larger ID microcatheter. For example, the larger diameters of the Pipeline Vantage (Medtronic) contain 64 wires and are deployed through a 0.027-inch ID microcatheter, whereas the smaller diameters contain 48 wires and can be deployed through a 0.021-inch ID microcatheter.<sup>14</sup> Incorporating more wires into the braid has been shown to be beneficial for promoting endothelialization across the aneurysm neck.<sup>15</sup> However, more wires may also be associated with a greater risk of perforator artery occlusion and thromboembolism.<sup>16,17</sup> Market-approved FDs typically contain 48–64 wires.<sup>2</sup> A similar FD design parameter is pics-per-inch (PPI), which is defined as the number of wire crossings per inch along the length of the device. PPI can be increased by increasing the number of wires in the braid and can also be adjusted by altering the angle at which the wires are braided, as illustrated in Fig 2.<sup>11,13</sup> When one views the FD from the longitudinal view, if the braid angle is defined as the angle between the wires and the axial axis of the stent (as defined in Fig 2), increasing the braid angle increases the PPI.<sup>11,13</sup> A higher PPI results in FDs with increased opening force and can improve device conformity, ultimately resulting in improved wall apposition.<sup>11,13</sup> However, the increased density of wires associated with increased PPI may also increase the delivery force and require a larger-ID microcatheter. Similar to increasing the number of wires within the braid, a higher PPI promotes aneurysm occlusion and endothelialization across the aneurysm neck but may increase the risk of perforator artery occlusion and thromboembolism.<sup>16–18</sup>

Oversizing of FDs is another design parameter or option available to the interventionalist when selecting a device. Devices may be constructed so that their nominal diameter, resulting from the mandrel diameter around which the wires were braided, is larger than the diameter of their intended parent artery. Oversizing was initially motivated to improve wall apposition and opening force. However, in vitro fluid dynamics studies have demonstrated that oversizing causes alterations to device geometry such as the angle between the wires and pore or cell size, resulting in higher blood flow within the aneurysm sac.<sup>19</sup> Consequently, oversizing has been correlated with lower rates of aneurysm occlusion in the clinic.<sup>20</sup> Foreshortening of the device also needs to be considered when selecting an FD size. FDs that are overexpanded in the radial direction will end up becoming shorter. Conversely, FDs will become longer than

anticipated when deployed in arteries with much smaller diameters than the device.<sup>21</sup>

Multiple layers of wire braids may also be used when constructing FDs. Increasing the number of layers theoretically increases the opening force but may require larger-ID microcatheters. The primary motivation for adding layers is to “artificially” increase the PPI and improve the flow diversion function by depositing more wires over the aneurysm neck. Multiple layers may result in increased blood stagnation

within the aneurysm sac and more rapid tissue growth across the device and aneurysm neck.<sup>22</sup> However, stacked wires increase their extension into the lumen of the parent vessel, worsening the risk of thromboembolism and the rate at which wires apposed to the artery wall become endothelialized.<sup>16,23</sup> Furthermore, additional layers may create malapposition between the layers or more disruption of the endothelium through additional device deployments. The Flow-Redirection Endoluminal Device (FRED) (MicroVention) has a unique design with a larger wire weave on the outside, similar to that of the LVIS Stent (MicroVention), and the braided FD on the inside.<sup>2</sup> On a related note, the use of multiple single-layer devices deployed telescopically over the aneurysm neck has also been shown to increase the rate of thromboembolic complications.<sup>24</sup>

Porosity, pore density, and the percentage of metal coverage are classic FD metrics that are ultimately a result of its selected design parameters.<sup>25</sup> Porosity is defined as the percentage area of the device wall not covered by wires. Conversely, percentage metal coverage is defined as the percentage area of the device wall covered by wires; both are less specific because thicker wires or more wire density can result in higher-percentage metal coverage or lower porosity.<sup>25</sup> Pore density is defined as the number of open cells or pores within the FD wall per area and is more specific.<sup>25</sup> These characteristics will determine how well the device diverts blood flow away from the aneurysm and stagnates blood within it.<sup>26</sup> Increasing the number of wires within the braid, PPI, and number of layers will decrease porosity and increase pore density and percentage metal coverage.<sup>11,13</sup> Another consideration is the curvature of the parent vessel. For a given device, when deployed in a curved segment, the porosity will be greater in regions of the FD near the outer edge of the curve.<sup>21</sup> Porosity and pore density of market-approved FDs are typically in the range of 50%–70% and 13–30 pores/mm<sup>2</sup>, respectively.<sup>27</sup>

### Material Selection

The material selected to construct the wires has a notable impact on overall FD performance due to a variety of often interrelated constitutive properties. As an illustration, consider 1 material property taken in the design of FD appliances: the elastic modulus (also known as the Young modulus), which is how much the material resists elastic deformation per unit of strain or simply considered as elastic stiffness. The connection of the elastic

**Table 1: Properties of base materials used in market-approved FDs**

Primary FD Materials			
Alloy	Elastic Modulus	Elastic Limit	Market-Approved FD Examples
Cobalt-nickel-chromium	~120–240 GPa <sup>a</sup>	~ 0.5 to 1.5% <sup>a</sup>	PED, PED Flex, PED Shield, PED Vantage, Surpass Streamline (Stryker), Surpass Evolve <sup>2</sup>
Nitinol	~ 40–10 GPa <sup>a</sup>	~ 6%–12% <sup>a</sup>	FRED, FRED Jr, FRED X, Silk, Silk+, p64, Tubridge <sup>2</sup>

<sup>a</sup> Use due consideration in applying values for modeling because the Young modulus and elastic limit (pseudoelastic for nitinol) can vary widely due to process-structure texture and strength outcomes. Bookends given here are realistic possible outcomes dependent on the supplier production path and target outcomes.

modulus to overall device stiffness and performance is not straightforward because of other influencing factors. For example, structural stiffness depends more on wire diameter and braid parameters than the elastic modulus. Furthermore, other phenomena, such as local yielding, and elastic strain limits can dramatically change the actual device response, depending on boundary conditions.

Materials with a higher elastic modulus require a greater amount of energy to deform but will then release a greater amount of energy when springing back to their original shape (imagine a stiff spring verses a Slinky). Conversely, materials with a high elastic modulus usually show lower elastic strain limits, which can reduce size recovery, or the diameter the device is able to re-expand after being crimped into a catheter. Consequently, FDs composed of wires with a higher elastic modulus may exhibit a higher opening force or lose force before apposing because of inadequate size recovery and may require larger-ID microcatheters to compensate. Cobalt-nickel-chromium alloys have been optimized to increase the elastic modulus and, therefore, are a commonly selected FD wire material. The Pipeline Embolization Device (PED; Medtronic) and the Surpass Evolve (Stryker) are composed primarily of cobalt alloy wires.<sup>2</sup>

Another critical material property is the elastic strain limit, which is defined as how much the material can be deformed before experiencing a permanent alteration to its shape. In the FD application, a higher elastic strain limit allows devices to be crimped to smaller diameters and bent around tighter curves while still maintaining the ability to self-expand back to their original intended deployment shape. This feature is advantageous for navigating to more distal aneurysms where the parent arteries are smaller in diameter and more tortuous.<sup>28</sup> Nickel titanium (nitinol) alloys have been optimized to increase the elastic strain limit via phase transformation and its associated superelasticity and, therefore, are also a commonly used FD wire material. FRED and p64 (phenox) are composed primarily of nitinol wires.<sup>2</sup>

While cobalt-nickel-chromium and nitinol represent excellent material selections based on their mechanical properties, they offer little radiopacity. Materials with relatively high radiodensity such as tantalum, platinum, or platinum-tungsten alloys are used to provide greater visibility than materials such as cobalt-nickel-chromium or nitinol to aid in fluoroscopy-guided delivery. All market-approved FDs feature some sort of radiopaque elements within their construction. For example, the PED and Surpass Evolve feature platinum-tungsten wires within their braid, whereas the FRED features tantalum radiopaque markers attached at either end of the device as well as unifying tantalum wires between the layers.<sup>2</sup> Another approach to convey radiopacity is the use of drawn filled tubing (DFT) wires. DFT wires contain an inner core wire, typically composed of radiopaque platinum or

tantalum, surrounded by an outer layer of nitinol or a cobalt-nickel-chromium alloy. FDs constructed from DFT wires aim to provide uniform radiopacity along the entire device length, while still taking advantage of the favorable mechanical properties of nitinol or cobalt-nickel-chromium alloys. Silk Vista Baby (Balt) and DERIVO FD (Acandis) are examples of CE-approved FDs constructed from nitinol outer/platinum inner DFT wires.<sup>2</sup> The Pipeline Vantage is constructed from cobalt alloy outer/platinum inner DFT wires.<sup>14</sup> Tables 1 and 2 summarize the material properties and examples of market-approved FDs constructed by the conventional FD materials.

### Surface Modifications

Modifying the surface of FDs is another approach to improve device performance. The surface of the wires may be electrically, chemically, or mechanically altered or coated to produce a beneficial effect. Surface modifications aim to reduce thrombogenicity, increase endothelialization, and/or act as lubricants to reduce delivery force. Several market-approved FDs have specific surface modifications intended to improve clinical efficacy, and FD surface modifications remain a hot topic of research.

Newer iterations of the PED feature Shield Technology (Medtronic), which is a phosphorylcholine layer deposited on the wire surfaces intended to reduce thrombogenicity and increase endothelialization. In vitro studies have demonstrated reduced thrombogenicity relative to an uncoated PED.<sup>23</sup> In vivo studies have demonstrated lower thrombogenicity and a more rapid endothelialization due to the coating.<sup>14,29</sup> The DERIVO FD wires are treated to form surface oxides and oxynitrides, termed BlueXide, which is intended to reduce friction/delivery force and thrombogenicity.<sup>30</sup> The newer iteration of FRED, FRED X, features a poly (2-methoxyethyl acrylate) coating, which has been shown to reduce thrombogenicity in vitro.<sup>31</sup> A newer iteration of the p48, p48\_HPC (phenox), features an antithrombogenic glycan-based hydrophilic coating that has demonstrated promising results in a 5-patient cohort of unruptured aneurysms treated with single antiplatelet therapy.<sup>32</sup>

### Future Directions

As the clinical use of FDs has become more popular, the intensity with which FD technology is being researched and developed has also become more popular. The growing body of clinical FD performance data can be leveraged in the design of next-generation devices. Computational hemodynamic approaches are being applied to better understand the optimal FD design parameters for saccular blood stagnation and aneurysm occlusion.<sup>33</sup> Active surface coatings are being engineered to impart therapeutic effects. For example, nitric oxide-releasing coatings show promise in reducing thrombogenicity and increasing endothelialization

**Table 2: Properties of radiopaque materials used in market-approved FDs**

Radiopaque Materials			
Alloy	Radiopacity (Density) <sup>a</sup>	Market-Approved FD Examples	Radiopaque Element Type
Platinum-tungsten	~19.75 g/cm <sup>3</sup>	PED, PED Flex, PED Shield, Surpass Streamline, Surpass Evolve	12 Platinum-tungsten wires incorporated into the overall braid <sup>2</sup>
Tantalum	16.65 g/cm <sup>3</sup>	FRED, FRED Jr, FRED X	Interwoven tantalum strands between braided layers and 4 markers on flared ends of device <sup>2</sup>
Platinum	21.45 g/cm <sup>3</sup>	Silk, Silk+, p64, Tubridge	Silk features 4 platinum wire strands with flared ends. Silk+ features 8 platinum wires and 4 platinum coils at device ends. <sup>2</sup> The p64 features 2 platinum wires wrapped along the length of the device and 8 markers at the device end. <sup>36</sup> Tubridge features 2 platinum strands <sup>2</sup>
Nitinol outer, platinum inner DFT	~6 g/cm <sup>3</sup> for nitinol; 21.45 g/cm <sup>3</sup> for platinum	Silk Vista, Silk Vista Baby, p48MW (phenox), p48_HPC, DERIVO	48 Braided DFT wires. DERIVO features 3 additional markers on the device end <sup>2</sup>
Cobalt-nickel-chromium outer, platinum inner DFT	~8.5 g/cm <sup>3</sup> for cobalt-nickel-chromium; 21.45 g/cm <sup>3</sup> for platinum	Pipeline Vantage	48 or 64 Braided DFT wires <sup>14</sup>

<sup>a</sup> Density is given as a rough proxy for radiopacity in lieu of the lack of corroborative data for these materials and geometries by a consistent approach. Future studies, outside the present scope, may focus on quantitative radiopacity comparison as a function of material, geometry, filtering materials, and imaging parameters. Such a study could be used to aid in the design of future devices.

of vascular scaffolds.<sup>34</sup> Heparin or other anticoagulant-containing coatings may also reduce thrombogenicity.<sup>35</sup> Proteins may be attached to device surfaces to facilitate endothelial cell capture and proliferation.<sup>35</sup> Bioresorbable FDs are also under development.<sup>27</sup> These devices are intended to dissolve after healing the aneurysm, mitigating or eliminating complications associated with the permanent presence of conventional FDs.

## CONCLUSIONS

Most practicing interventionalists gain their understanding of FD performance on the basis of their clinical experience. Choosing an FD may be related to personal preference and the availability of FDs at a given medical center. Furthermore, certain devices have only achieved market approval in specific geographic regions. For example, the PED, Surpass, and FRED are the 3 FDs currently available in the United States; Silk, p64, and DERIVO are from European companies with CE approval, and Tubridge (MicroPort) has National Medical Products Administration approval in China.<sup>2</sup> The geographic availability influences the popularity of certain devices in certain regions. Each FD features a unique design resulting in its own set of advantages and disadvantages. A more comprehensive understanding of the relationship between engineering design principles and clinical performance may help practicing interventionalists navigate the rapidly growing FD market. Furthermore, we believe this improved understanding can promote collaboration between clinicians and engineers, ultimately accelerating the development of FD technology.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

- Brinjikji W, Murad MH, Lanzino G, et al. **Endovascular treatment of intracranial aneurysms with flow diverters: a meta-analysis.** *Stroke* 2013;44:442–47 CrossRef Medline
- Dandapat S, Mendez-Ruiz A, Martínez-Galdámez M, et al. **Review of current intracranial aneurysm flow diversion technology and clinical use.** *J Neurointerv Surg* 2021;13:54–62 CrossRef Medline
- Ravindran K, Casabella AM, Cebal J, et al. **Mechanism of action and biology of flow diverters in the treatment of intracranial aneurysms.** *Neurosurgery* 2020;86:S13–19 CrossRef Medline
- Rouchaud A, Ramana C, Brinjikji W, et al. **Wall apposition is a key factor for aneurysm occlusion after flow diversion: a histologic evaluation in 41 rabbits.** *AJNR Am J Neuroradiol* 2016;37:2087–91 CrossRef Medline
- Aquarius R, de Korte A, Smits D, et al. **The importance of wall apposition in flow diverters.** *Neurosurgery* 2019;84:804–10 CrossRef Medline
- Ma J, You Z, Byrne J, et al. **Design and mechanical properties of a novel cerebral flow diverter stent.** *Ann Biomed Eng* 2014;42:960–70 CrossRef Medline
- Dabir D, Feisst A, Thomas D, et al. **Physical properties of venous stents: an experimental comparison.** *Cardiovasc Intervent Radiol* 2018;41:942–50 CrossRef Medline
- Kühn AL, Wakhloo AK, Gounis MJ, et al. **Use of self-expanding stents for better intracranial flow diverter wall apposition.** *Interv Neuroradiol* 2017;23:129–36 CrossRef Medline
- Kühn AL, Rodrigues KM, Wakhloo AK, et al. **Endovascular techniques for achievement of better flow diverter wall apposition.** *Interv Neuroradiol* 2019;25:344–47 CrossRef Medline
- Ni XY, Pan CW, Gangadhara Prusty B. **Numerical investigations of the mechanical properties of a braided non-vascular stent design using finite element method.** *Comput Methods Biomech Biomed Engin* 2015;18:1117–25 CrossRef Medline
- Suzuki T, Takao H, Fujimura S, et al. **Relationships between geometrical parameters and mechanical properties for a helical braided flow diverter stent.** *Technol Health Care* 2017;25:611–23 CrossRef Medline
- Kastrati A, Mehilli J, Dirschinger J, et al. **Intracoronary stenting and angiographic results: strut thickness effect on restenosis outcome (ISAR-STEREO) trial.** *Circulation* 2001;103:2816–21 CrossRef Medline
- Kim JH, Kang TJ, Yu WR. **Mechanical modeling of self-expandable stent fabricated using braiding technology.** *J Biomech* 2008;41:3202–12 CrossRef Medline
- Starke RM, Thompson J, Pagani A, et al. **Preclinical safety and efficacy evaluation of the Pipeline Vantage Embolization Device with Shield Technology.** *J Neurointerv Surg* 2020;12:981–86 CrossRef Medline
- Kallmes DF, Ding Y, Dai D, et al. **A second-generation, endoluminal, flow-disrupting device for treatment of saccular aneurysms.** *AJNR Am J Neuroradiol* 2009;30:1153–58 CrossRef Medline

16. Rogers C, Edelman ER. **Endovascular stent design dictates experimental restenosis and thrombosis.** *Circulation* 1995;91:2995–3001 CrossRef Medline
17. Hong B, Wang K, Huang Q, et al. **Effects of metal coverage rate of flow diversion device on neointimal growth at side branch ostium and stented artery: an animal experiment in rabbit abdominal aorta.** *Neuroradiology* 2012;54:849–55 CrossRef Medline
18. Sadasivan C, Cesar L, Seong J, et al. **An original flow diversion device for the treatment of intracranial aneurysms: evaluation in the rabbit elastase-induced model.** *Stroke* 2009;40:952–58 CrossRef Medline
19. Mut F, Cebal J. **Effects of flow-diverting device oversizing on hemodynamics alteration in cerebral aneurysms.** *AJNR Am J Neuroradiol* 2012;33:2010–16 CrossRef Medline
20. Tong X, Han M, Wu Z, et al. **Effects of different stent size selection on Pipeline embolization device treatment of intracranial aneurysms.** *Ther Adv Neurol Disord* 2023;16:17562864231151475 CrossRef Medline
21. Shapiro M, Raz E, Becske T, et al. **Variable porosity of the Pipeline embolization device in straight and curved vessels: a guide for optimal deployment strategy.** *AJNR Am J Neuroradiol* 2014;35:727–33 CrossRef Medline
22. Ding Y, Tieu T, Kallmes D. **Experimental testing of a new generation of flow diverters in sidewall aneurysms in rabbits.** *AJNR Am J Neuroradiol* 2015;36:732–36 CrossRef Medline
23. Girdhar G, Andersen A, Pangerl E, et al. **Thrombogenicity assessment of Pipeline Flex, Pipeline Shield, and FRED flow diverters in an in vitro human blood physiological flow loop model.** *J Biomedical Materials Res* 2018;106:3195–202 CrossRef
24. Tan LA, Keigher KM, Munich SA, et al. **Thromboembolic complications with Pipeline Embolization Device placement: impact of procedure time, number of stents and pre-procedure P2Y12 reaction unit (PRU) value.** *J Neurointerv Surg* 2015;7:217–21 CrossRef Medline
25. Lieber BB, Sadasivan C. **Endoluminal scaffolds for vascular reconstruction and exclusion of aneurysms from the cerebral circulation.** *Stroke* 2010;41:S21–25 CrossRef Medline
26. Dholakia R, Sadasivan C, Fiorella DJ, et al. **Hemodynamics of flow diverters.** *J Biomech Eng* 2017;139 CrossRef Medline
27. Oliver AA, Carlson KD, Bilgin C, et al. **Bioresorbable flow diverters for the treatment of intracranial aneurysms: review of current literature and future directions.** *J Neurointerv Surg* 2023;15:178–82 CrossRef Medline
28. Chiu AH, Phillips TJ. **Future directions of flow diverter therapy.** *Neurosurgery* 2020;86:S106–16 CrossRef Medline
29. Matsuda Y, Jang DK, Chung J, et al. **Preliminary outcomes of single antiplatelet therapy for surface-modified flow diverters in an animal model: analysis of neointimal development and thrombus formation using OCT.** *J Neurointerv Surg* 2019;11:74–77 CrossRef Medline
30. Akgul E, Onan HB, Akpınar S, et al. **The DERIVO embolization device in the treatment of intracranial aneurysms: short-and mid-term results.** *World Neurosurg* 2016;95:229–40 CrossRef Medline
31. Yoshizawa K, Kobayashi H, Kaneki A, et al. **Poly (2-methoxyethyl acrylate) (PMEA) improves the thromboresistance of FRED flow diverters: a thrombogenic evaluation of flow diverters with human blood under flow conditions.** *J Neurointerv Surg* 2023;15:1001–06 CrossRef Medline
32. Bhogal P, Bleise C, Chudyk J, et al. **The p48\_HPC antithrombogenic flow diverter: initial human experience using single antiplatelet therapy.** *J Int Med Res* 2020;48:300060519879580 CrossRef Medline
33. Xu L, Zhao B, Liu X, et al. **Computational methods applied to analyze the hemodynamic effects of flow-diverter devices in the treatment of cerebral aneurysms: current status and future directions.** *Med Nov Technol Devices* 2019;3:100018 CrossRef
34. Rao J, Pan Bei H, Yang Y, et al. **Nitric oxide-producing cardiovascular stent coatings for prevention of thrombosis and restenosis.** *Front Bioeng Biotechnol* 2020;8:578 CrossRef Medline
35. Liu P, Liu Y, Li P, et al. **Rosuvastatin-and heparin-loaded poly (l-lactide-co-caprolactone) nanofiber aneurysm stent promotes endothelialization via vascular endothelial growth factor type A modulation.** *ACS Appl Mater Interfaces* 2018;10:41012–18 CrossRef Medline
36. Bhogal P, Moreno RM, Ganslandt O, et al. **Use of flow diverters in the treatment of unruptured saccular aneurysms of the anterior cerebral artery.** *J Neurointerv Surg* 2017;9:283–89 CrossRef Medline



# Understanding Bias in Artificial Intelligence: A Practice Perspective

● Melissa A. Davis, ● Ona Wu, ● Ichiro Ikuta, ● John E. Jordan, ● Michele H. Johnson, and ● Edward Quigley

## ABSTRACT

**SUMMARY:** In the fall of 2021, several experts in this space delivered a Webinar hosted by the American Society of Neuroradiology (ASNR) Diversity and Inclusion Committee, focused on expanding the understanding of bias in artificial intelligence, with a health equity lens, and provided key concepts for neuroradiologists to approach the evaluation of these tools. In this perspective, we distill key parts of this discussion, including understanding why this topic is important to neuroradiologists and lending insight on how neuroradiologists can develop a framework to assess health equity–related bias in artificial intelligence tools. In addition, we provide examples of clinical workflow implementation of these tools so that we can begin to see how artificial intelligence tools will impact discourse on equitable radiologic care. As continuous learners, we must be engaged in new and rapidly evolving technologies that emerge in our field. The Diversity and Inclusion Committee of the ASNR has addressed this subject matter through its programming content revolving around health equity in neuroradiologic advances.

**ABBREVIATIONS:** AI = artificial intelligence; ASNR = American Society of Neuroradiology; TAT = turnaround time

Many artificial intelligence (AI) tools currently in clinical practice involve neuroimaging, including tools for detection, acquisition, and segmentation. It is important for neuroradiologists to evaluate these tools for clinical efficacy and safety, including how the use of these tools will impact patient care and workflow. There is ample literature to help neuroradiologists understand the basic principles and technology of AI and how to approach the evaluation and validation of AI tools. Although the original literature focused on the scientific development process continues to evolve, there is increasing interest in the potential biases of these types of learning applications.<sup>1</sup>

Health disparities in neurologic diseases are well-characterized and cross many sociodemographic variables, including race, socioeconomic status, and insurance status. The effects on population health are highlighted through the study of social determinants of health, which can serve as key drivers of health disparities.<sup>1</sup> Such disparities can have negative compounding effects on the health care continuum and, ultimately, patient

outcomes. Understanding AI through the lens of health equity is necessary to recognize bias that might be introduced in AI algorithms and to mitigate biases that can occur.

In the fall of 2021, several experts in this space delivered a Webinar hosted by the American Society of Neuroradiology (ASNR) Diversity and Inclusion Committee, focused on expanding the understanding of bias in AI and provided key concepts for neuroradiologists to approach evaluation of these tools. In this perspective, we distill key insights from the dynamic discussion that ensued. The source Webinar is available as enduring content on the ASNR Education Connection ([https://www.pathlms.com/asnr/courses/56243/video\\_presentations/268414#](https://www.pathlms.com/asnr/courses/56243/video_presentations/268414#)).

## Why Is It Important for Neuroradiologists to Care about Bias in AI?

As long as AI algorithms are relegated to the role of cognitive assistant, algorithmic bias might not be a pressing issue for neuroradiologists. Once AI models are used to predict outcomes, manage care, or order workflow, potential bias in the collection or labeled training data needs to be considered. One can imagine that under-representation of populations in the training data can lead to inaccurate predictions of outcomes. One example is an application using algorithms to segment out the “core” infarct on CT perfusion imaging.<sup>2</sup> If the core infarct volume is larger than a certain threshold, some studies have suggested that there will be no benefit to the patient from endovascular treatment and therefore the patient should be excluded.<sup>3</sup> However, subsequent

Received August 29, 2023; accepted after revision October 17.

From Yale University (M.A.D., M.H.J.), New Haven, Connecticut; Massachusetts General Hospital (O.W.), Charlestown, Massachusetts; Mayo Clinic Arizona, Department of Radiology (I.I.), Phoenix, Arizona; Stanford University School of Medicine (J.E.J.), Stanford, California; and University of Utah (E.Q.), Salt Lake City, Utah.

Please address correspondence to Melissa A. Davis, MD, MBA, Yale University, 330 Cedar Street, New Haven, CT 06510; e-mail: [Melissa.a.davis@yale.edu](mailto:Melissa.a.davis@yale.edu)

<http://dx.doi.org/10.3174/ajnr.A8070>

studies have shown that core infarcts predicted by acute CTP fail to manifest on follow-up noncontrast CT after successful recanalization, especially when patients are treated early.<sup>2</sup> In this situation, incorrect prediction of the algorithm of a large-volume infarction might have precluded a beneficial treatment if treatment was guided purely by a computer algorithm.

There is increasing interest in using generative algorithms for many clinical neuroimaging applications, ranging from increasing resolution on low-resolution data to being able to convert one technique to another, such as a CT-to-MR imaging conversion. The synthetics are very effective and can even recreate susceptibility artifacts. The accuracy of such transformations depends highly on the training data. If there is bias in the data, inaccurate transformations can arise. This issue is highlighted by the Face-Depixelizer example (Tg-bomze/Face-Depixelizer; <https://github.com/tg-bomze/Face-Depixelizer>). Like methods that convert low-resolution images to high-resolution images, the Face-Depixelizer can take a compressed low-resolution image and generate an image with quality similar to that of the original image. This technique holds great promise for applications, ranging from data storage to streaming. However, it was quickly discovered that the Depixelizer failed for nonwhite faces, which were incorrectly transformed into faces with white features.<sup>4</sup>

### **Tools to Approach Understanding Bias**

By integrating fairness into the machine learning lifecycle, we can mitigate unfairness. The machine learning lifecycle can be simplified to the model-development phase, deployment phase, and feedback reverting to the development phase to refine the model performance. By defining fairness requirements and by involving diverse stakeholders at the model-development stage, potential bias can be mitigated. Sources of bias include definition of the task, data set construction, and cost-function for the training algorithm. Poor task and cost-function definitions can lead inadvertently to racially biased machine learning algorithms. Imbalances in the training data can lead to underdiagnoses in the under-represented data set. For models that focus on positive or negative predictive values, the prevalence of the disease in the training and validation data should match real-world clinical distribution. A detailed discussion on the consequences of data mishandling is found in a review by Rouzrokh et al.<sup>5</sup>

One method to potentially mitigate problems using machine learning algorithms for clinical decision-making is explainability. For example, to explain the classification of hemorrhage subtypes,<sup>6</sup> attention maps were used to highlight which features had the greatest weights in the decision of the algorithm. The authors showed that regions corresponding to SAH and intraventricular hemorrhage were appropriately highlighted. In contrast, there is the risk of data leakage in which features indirectly associated with disease prevalence are incorrectly interpreted by the model to be the disease itself. This was demonstrated in a pneumonia classification program that focused on metallic tokens as the most relevant feature in patients with pneumonia instead of the lungs because the tokens could be used to identify which hospital provided the data, ie, the hospital more likely to have pneumonia

cases.<sup>7</sup> Although the algorithm nominally had high accuracy with its training and testing data, the algorithm did not actually learn the true, relevant features in the lungs and would likely fail when deployed at other hospital centers.

We must always consider such effects in daily practice and the influence of these technologies on our decision-making, impacting patient care.<sup>8</sup> Consider the following examples.

### **Examination Triage**

AI algorithms may aid in the detection of emergent findings such as acute intracranial hemorrhage and cervical spine fractures. O'Neill et al<sup>9</sup> showed that simply marking an examination as having an emergent finding does not affect report turnaround time (TAT), but grouping all marked examinations at the top of the reading list does result in improved TAT. However, we might take a step back and look at the bias of an AI system and how that might propagate unfairness. As an example, acute intracranial hemorrhage detection was initially thought to have high sensitivity, specificity, positive predictive value, negative predictive value, and accuracy. Yet, subsequent clinical implementations of the same algorithm at other institutions showed a positive predictive value of 81.3%, with decreased performance for older patients.<sup>10</sup> If an older patient with an acute intracranial hemorrhage and a false-negative AI result waited longer for his or her scan to be read due to a triage system set up around AI results, that scenario would not be equitable.

### **Examination Scheduling**

Maximizing examinations performed in a day results in maximal profits for a radiology practice. Identifying delays, potential “no-shows,” and potential times for additional examinations could optimize scheduling and potentially improve patient satisfaction. Panykh et al<sup>11</sup> showed that while several AI models have been used to predict examination times/delays, these models all decline in quality with time. Various factors could affect the model such as seasonal/migrant workers, a sudden influx of refugees, major industry closure with layoffs and loss of health care benefits, or a global pandemic. These factors could be addressed with the implementation of continuous-learning AI that repeatedly uses updated data input from hospital information systems to dynamically address fluctuations in population health and identify barriers to health care that can be addressed, such as lack of dependable transportation, lack of health insurance, lack of child-care, or an inability to pay.

## **DISCUSSION**

In this perspective, we have introduced resources to aid the neuroradiologist in learning and contemplating the intersection of AI and health equity. By leveraging examples of clinical workflow implementation of these tools, we can begin to see how AI tools will impact discourse on equitable radiologic care. As end users of these tools, we are responsible for understanding potential pitfalls and implicit biases that may affect our ability as physicians and neuroradiologists to deliver equitable high-quality care to our patients.

As continuous learners, we must be engaged especially as new and rapidly evolving technologies emerge in our field. AI is the

newest of these advances, and there is an urgent need to remain focused on health equity within radiology as we begin to leverage and automate these rapidly evolving tools. The Diversity and Inclusion Committee of the ASNR has taken on this task in collaboration with the Computer Science and Informatics Committee and the Artificial Intelligence Committee. Through this programming content, learners can access in-depth discussions regarding health equity in neuroradiologic advances.

## CONCLUSIONS

Since 2020, the ASNR Diversity and Inclusion Committee has hosted Webinars spanning medical-social objectives to core science objectives. These types of Webinars allow neuroradiologists to engage in digestible content like bias in AI. The Webinar discussed in this article focused on the intersection of health equity and bias, with the goal of introducing imaging experts to these concepts in meaningful ways.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

1. Davis MA, Lim N, Jordan J, et al. **Imaging artificial intelligence: a framework for radiologists to address health equity, from the AJR Special Series on DEI.** *AJR Am J Roentgenol* 2023;221:302–08 CrossRef Medline
2. Martins N, Aires A, Mendez B, et al. **Ghost infarct core and admission computed tomography perfusion: redefining the role of neuroimaging in acute ischemic stroke.** *Interv Neurol* 2018;7:513–21 CrossRef Medline
3. Albers GW, Marks MP, Kemp S, et al; DEFUSE 3 Investigators. **Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging.** *N Engl J Med* 2018;378:708–18 CrossRef Medline
4. Menon S, Damian A, Hu S, et al. **Pulse: self-supervised photo upsampling via latent space exploration of generative models.** *arXiv* 2020 CrossRef <https://arxiv.org/abs/2003.03808>. Accessed June 23, 2023
5. Rouzrokh P, Khosravi B, Faghani S, et al. **Mitigating bias in radiology machine learning, 1: data handling.** *Radiol Artif Intell* 2022;4:e210290 CrossRef Medline
6. Lee H, Yune S, Mansouri M, et al. **An explainable deep-learning algorithm for the detection of acute intracranial haemorrhage from small datasets.** *Nat Biomed Eng* 2019;3:173–82 CrossRef Medline
7. Zech JR, Badgeley MA, Liu M, et al. **Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study.** *PLoS Med* 2018;15:e1002683 CrossRef Medline
8. Filippi CG, Stein JM, Wang Z, et al. **Ethical considerations and fairness in the use of artificial intelligence for neuroradiology.** *AJNR Am J Neuroradiol* 2023;44:1242–48 CrossRef Medline
9. O'Neill TJ, Xi Y, Stehel E, et al. **Reprioritization of the reading workload using artificial intelligence has a beneficial effect on the turnaround time for interpretation of head CT with intracranial hemorrhage.** *Radiol Artif Intell* 2021;3:e200024 CrossRef Medline
10. Voter AF, Meram E, Garrett JW, et al. **Diagnostic accuracy and failure mode analysis of a deep learning algorithm for the detection of intracranial hemorrhage.** *J Am Coll Radiol* 2021;18:1143–52 CrossRef Medline
11. Pianykh OS, Langa G, Dewey M, et al. **Continuous learning AI in radiology: implementation principles and early applications.** *Radiology* 2020;297:6–14 CrossRef Medline

# Shift Volume Directly Impacts Neuroradiology Error Rate at a Large Academic Medical Center: The Case for Volume Limits

Vladimir Ivanovic, Kenneth Broadhead, Yu-Ming Chang, John F Hamer, Ryan Beck, Lotfi Hacein-Bey, and Lihong Qi

## ABSTRACT

**BACKGROUND AND PURPOSE:** Unlike in Europe and Japan, guidelines or recommendations from specialized radiological societies on workflow management and adaptive intervention to reduce error rates are currently lacking in the United States. This study of neuroradiologic reads at a large US academic medical center, which may hopefully contribute to this discussion, found a direct relationship between error rate and shift volume.

**MATERIALS AND METHODS:** CT and MR imaging reports from our institution's Neuroradiology Quality Assurance database (years 2014–2020) were searched for attending physician errors. Data were collected on shift volume specific error rates per 1000 interpreted studies and RADPEER scores. Optimal cutoff points for 2, 3 and 4 groups of shift volumes were computed along with subgroups' error rates.

**RESULTS:** A total of 643 errors were found, 91.7% of which were clinically significant (RADPEER 2b, 3b). The overall error rate (errors/1000 examinations) was 2.36. The best single shift volume cutoff point generated 2 groups:  $\leq 26$  studies (error rate 1.59) and  $> 26$  studies (2.58; OR: 1.63,  $P < .001$ ). The best 2 shift volume cutoff points generated 3 shift volume groups:  $\leq 19$  (1.34), 20–28 (1.88; OR: 1.4,  $P = .1$ ) and  $\geq 29$  (2.6; OR: 1.94,  $P < .001$ ). The best 3 shift volume cutoff points generated 4 groups:  $\leq 24$  (1.59), 25–66 (2.44; OR: 1.54,  $P < .001$ ), 67–90 (3.03; OR: 1.91,  $P < .001$ ), and  $\geq 91$  (2.07; OR: 1.30,  $P = .25$ ). The group with shift volume  $\geq 91$  had a limited sample size.

**CONCLUSIONS:** Lower shift volumes yielded significantly lower error rates. The lowest error rates were observed with shift volumes that were limited to 19–26 studies. Error rates at shift volumes between 67–90 studies were 226% higher, compared with the error rate at shift volumes of  $\leq 19$  studies.

**ABBREVIATIONS:** ICC = intraclass correlation coefficient; QA = quality assurance; VIF = variance inflation factor

Medical errors are the third leading cause of death in the United States.<sup>1</sup> Many industries have implemented system-wide measures and technological advancements to decrease errors and fatalities. Between 1973 and 2020, the number of commercial aviation fatalities has decreased by 97%.<sup>2</sup> Anesthesia-related mortality in the United States has decreased by 97% from 1952 to 2005.<sup>3</sup> Reported rates of major errors in radiology range between 2%–6%.<sup>4–7</sup> Error rates in diagnostic radiology have remained relatively constant since the time they were first studied in 1949.<sup>8–14</sup>

Studies of both general radiology<sup>15,16</sup> and neuroradiology<sup>17</sup> have found higher shift volumes to be associated with diagnostic errors. Larger scale studies evaluating radiology error rates as a function of specific shift volume are lacking. The identification of specific shift volume ranges that minimize error rates could lead to corrective interventions. Our aims were to extend the knowledge gained from our prior work by using the same data set<sup>17–22</sup> to evaluate shift volume specific error rates by neuroradiologists at a large academic medical center and to identify the shift volume ranges with the lowest error rates.

## MATERIALS AND METHODS


### Study Setting

Institutional review board approval with a waiver of informed consent was obtained for this retrospective study. Our department's neuroradiology quality assurance (QA) database was searched for attending neuroradiologist errors between 2014–2020, which includes cases collected during the review of all addended reports and a review of computer selected studies during a clinical service, as previously described.<sup>20</sup> Whenever assigned to a

Received October 6, 2023; accepted after revision December 18.

From the Department of Radiology, Section of Neuroradiology (V.I., J.F.H., R.B.), Medical College of Wisconsin, Milwaukee, Wisconsin; Department of Statistics (K.B.), Colorado State University, Fort Collins, Colorado; Department of Radiology, Section of Neuroradiology (Y.-M.C.), Beth Israel Deaconess Medical Center, Boston, Massachusetts; Department of Radiology, Section of Neuroradiology (L.H.-B.), University of California Davis Medical Center, Sacramento, California; and Department of Public Health Sciences (L.Q.), School of Medicine, University of California Davis, Davis, California.

Please address correspondence to Vladimir Ivanovic, MD, Medical College of Wisconsin, Department of Radiology, 8701 Watertown Plank Rd, Milwaukee, WI 53226; e-mail: Vladimir.ivanovic@yahoo.com

 Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8119>

clinical service (all shifts, all radiologists), the neuroradiologist is presented with 3 randomly computer selected CT or MR imaging examinations, and the neuroradiologist is instructed to review and assign a score (1, 2a, 2b, 3a, 3b) to each by using the American College of Radiology RADPEER system.<sup>23</sup> All RADPEER scores of 2 and 3 are further reviewed by either 2 additional attending neuroradiologists or the entire neuroradiology division during a quarterly QA conference to reach a consensus RADPEER score. Each reviewed examination is entered into the QA database along with its associated RADPEER score. Diagnostic error rates are expressed as the number of errors per 1000 interpreted examinations that are shift volume specific.

All 16 current or former members of the neuroradiology division are full-time employees who have completed neuroradiology fellowships. Division members have similar yearly productivity, read similar neuroimaging case mixes and rotate with similar frequency among general neuroradiology nonspecialized clinical services: emergency/inpatient (9-hour day shift), spine (9-hour day shift), outpatient (9-hour day shift), evening weekday call (6-hour evening shift), and weekend call (9–12-hour day shift). Attending read a similar mix of adult and pediatric CT and MR imaging neuroradiology studies, which are read in chronological order on either a single common outpatient study worklist or a dedicated emergency/inpatient worklist. Broadly speaking, approximately 40% of the studies were dictated by trainees.

### Data Collection

Data on shift volumes on the days an error was documented, the number of errors per shift volume, the total number of studies interpreted for each shift volume and demographic information were extracted from the radiology departmental database. Diagnostic error rates are expressed as the number of errors per 1000 interpreted examinations for each specific shift volume (for example, if there were 10 errors detected on all shifts of 40 studies, and there were 1000 shifts during which 40 studies were interpreted, then the error rate for a shift volume of 40 would be  $10 / [40 \times 1000]$ ). The shift volume was defined as the total number of CT and MR imaging examinations that the neuroradiologist interpreted during the shift (eg, if 6 radiologists are working on a given day, this would constitute 6 total shifts for that day). We excluded shifts with 10 or fewer interpreted studies per day as outliers; many of these likely represent reports generated on administrative or off days, possibly to catch up with overflow work from the day before. For examinations comprising multiple unique accession numbers (eg, head MR imaging-MRA), each accession number was considered to be a single examination for purposes of determining shift volume.

### Statistical Methods

Descriptive statistics (mean and SD for age; count and percentages for age) were obtained for the error group.

Optimal cutoff points categorizing examinations into one of several risk groups, using shift volume level as the diagnostic variable, were computed by using ordinary logistic regression.<sup>24</sup> Given the wide range of daily shift volume productivity within our practice, we wanted to determine more than one possible cutoff point and error rate differences for the resultant groups. Each possible

combination of cutoff points, categorizing examinations into 2, 3 or 4 separate risk groups, were computed. The optimal cutoff values are defined as the ones with the most significant overall likelihood ratio test statistics, giving the most significant overall set of risk groups for a specified number of cutoff points. Individual odds ratios and *P* values comparing each group with the group with the lowest shift volume (group 1) within each split are also computed, but these were not used to determine optimality. The error rate per 1000 studies was computed for each group within a split.

To assess the impact of potential correlation among examinations read by the same radiologist, the intraclass correlation, measured by the intraclass correlation coefficient (ICC), and the resulting variance inflation factor (VIF) were examined. In the presence of high intraclass correlation or VIFs, the use of ordinary logistic regression is inappropriate. As covariate information on the attending radiologist was unavailable for all examinations, the ICC and VIF were computed by using a subset of the data analyzed in a previous study.<sup>25</sup> ICC values near 1 indicate high correlation, whereas values near 0 indicate low correlation. High VIFs indicate that variances of estimated parameters are higher than they would be in the absence of intraclass correlation, whereas values near 1 indicate little inflation. The ICC, its 95% confidence interval, and the VIFs were computed by using resampling techniques.<sup>26</sup> All statistical analyses were performed in R, version 4.1.1.

### RESULTS

There were 643 patients within the error group; 349 were men (54.3%), 293 were women (45.6%), and 1 (0.2%) patient's sex was not recorded. The mean patient age was 49.8 years (SD, 25).

The estimated ICC among attending radiologists was low (0.001; 95% CI, 0–0.079), as was the estimated VIF (1.099), suggesting minimal correlation between examinations and minimal variance inflation. Six attending radiologists had small cluster sizes (having read fewer than 10 examinations). A sensitivity analysis of the ICC and VIF estimates was performed by rerunning the analysis without the small clusters. Similar estimates of the ICC (0.0005; 95% CI, 0–0.0796) and VIF (1.073) were obtained. Because of the low correlation and VIF, the method for determining optimal risk groups by using ordinary logistic regression was deemed acceptable.

A total of 643 errors were detected during the study period; 590 (91.7%) were clinically significant (RADPEER 2b, 3b). The overall error rate (per 1000 interpreted examinations) was found to be 2.36. Figure 1 presents a detailed view of the trend in error rate as a function of shift volume compared with the overall error rate. The Table presents the risk groups by shift volume for 1, 2 and 3 cutoff points. All group splits were found to be highly significant by using the likelihood ratio test ( $P < .001$ ).

For a single cutoff point (search done for each shift volume level), group 1 consists of examinations read during shift volumes of  $\leq 26$  (error rate 1.59), whereas group 2 consists of examinations read during shift volumes of  $\geq 27$  (error rate 2.58). The odds of an examination in group 2 containing an error were found to be 63% higher than examinations in group 1 (OR: 1.63,  $P < .001$ ).

For 2 cutoff points (search done by fours), group 1 consists of examinations read during shift volumes of  $\leq 19$  (error rate 1.34), group 2 consists of examinations read during shift volumes



between 20 and 28 (error rate 1.88), and group 3 consists of examinations read during shift volumes of  $\geq 29$  (error rate 2.6). The odds of an examination in group 2 containing an error were found to be 40% higher than examinations in group 1 (OR: 1.4,  $P = .1$ ), whereas the odds of examinations in group 3 were found to be 94% higher than examinations in group 1 (OR 1.94,  $P < .001$ ).

For 3 cutoff points (search done by tens), group 1 consists of examinations read during shift volumes of  $\leq 24$  (error rate 1.59), group 2 consists of examinations read during shift volumes between 25 and 66 (error rate 2.44), group 3 consists of examinations read during shift volumes between 67 and 90 (error rate 3.03), and group 4 consists of examinations read during shift volumes of  $\geq 91$  (error rate 2.07). The odds of an examination in group 2 containing an error were found to be 54% higher than examinations in group 1 (OR: 1.54,  $P < .001$ ), the odds of an examination in group 3 containing an error were found to be 91% higher than examinations in group 1 (OR: 1.91,  $P < .001$ ), and the odds of an examination in group 4 containing an error were found to be 30% higher than examinations in group 1 (OR 1.30,  $P = .25$ ).

## DISCUSSION

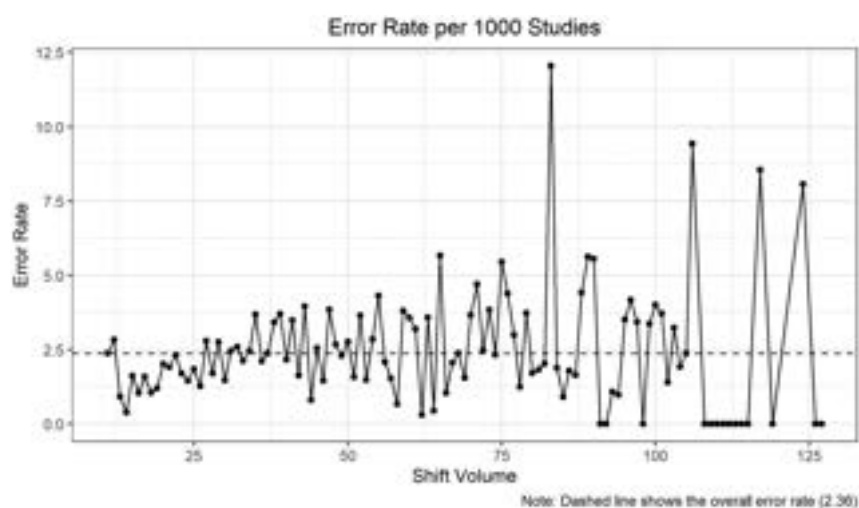
In this study, we have found that neuroradiology errors are significantly lower on low shift volumes compared with high shift

volumes. Based on our data, the best outcomes are seen when shift volumes are  $\leq 19$ –26 studies. Our data are in line with the findings of a recent nationwide survey of 42 academic neuroradiology division chiefs that suggested a mean optimal shift volume of between 25–33 studies.<sup>27</sup> Our group error rate was 2.36; if our shift volume could be limited to 19 studies, error rates could potentially decrease by 57%. The highest error rates were observed at shift volumes of between 67–90 studies, which was 226% higher than the lowest error rate, which was observed at shift volumes of  $\leq 19$  studies. Error rates at high shift volumes might be difficult to justify, given this degree of difference. Approximately 74 million CT and 39 million MR imaging studies are performed in the United States per year,<sup>28</sup> and decreasing the mean error rates might have a significant impact from a public health perspective.<sup>1</sup> Medicolegal outcomes of errors are not insignificant. An analysis of 2019 malpractice claims in diagnostic and interventional neuroradiology found a median plaintiff award of \$2,877,847 and a median settlement amount of \$1,950,000.<sup>29</sup> Missing aneurysms on CT or MR imaging may lead to plaintiff awards of \$4–43 million.<sup>30</sup>

For 3 cutoff points, the odds ratio for the group 4 risk group decreases, seemingly suggesting that it is safer than groups 2 and 3. However, this is most likely due to comparatively limited sample

sizes for the shift volume levels comprising that group. Figure 2 presents a plot of the sample size for each shift volume; shift volumes of 91 and greater can be seen to have markedly lower sample sizes than those of the lower shift volumes. Furthermore, many shift volumes after this threshold are so small that no error has yet been observed at these volume levels. Together with the overall trends seen in Fig 1, we do not have any reason to suspect that risk group 4 is any safer than groups 2 and 3. We conclude the decreased odds ratio for group 4 is the result of sample size limitations.

Prior literature has highlighted a broad range of potential strategies for reducing diagnostic errors within general and specialized practices, including interpretation



**FIG 1.** Detailed view of the trends in error rate as a function of shift volume level against the overall error rate.

### Optimal risk groups

	Group 1 (Error Rate) <sup>g</sup>	Group 2 (Error Rate)	Group 3 (Error Rate)	Group 4 (Error Rate)	Odds Ratio 1 (P Value)	Odds Ratio 2 (P Value)	Odds Ratio 3 (P Value)	LRT P Value <sup>h</sup>
1 cut point <sup>a</sup>	26 and below (1.59)	27+ (2.58)	—	—	1.63 (<.001) <sup>d</sup>	—	—	<.001
2 cut points <sup>b</sup>	19 and below (1.34)	20–28 (1.88)	29+ (2.6)	—	1.40 (.10) <sup>e</sup>	1.94 (<.001) <sup>e</sup>	—	<.001
3 cut points <sup>c</sup>	24 and below (1.59)	25–66 (2.44)	67–90 (3.03)	91+ (2.07)	1.54 (<.001) <sup>f</sup>	1.91 (<.001) <sup>f</sup>	1.30 (.25) <sup>f,i</sup>	<.001

**Note:**—LRT indicates likelihood ratio test.

<sup>a</sup> Search done for each shift volume level.

<sup>b</sup> Search done by fours.

<sup>c</sup> Search done by tens.

<sup>d</sup> Interpretation: The odds of containing an error are 63% higher for studies in group 2, compared with studies in group 1.

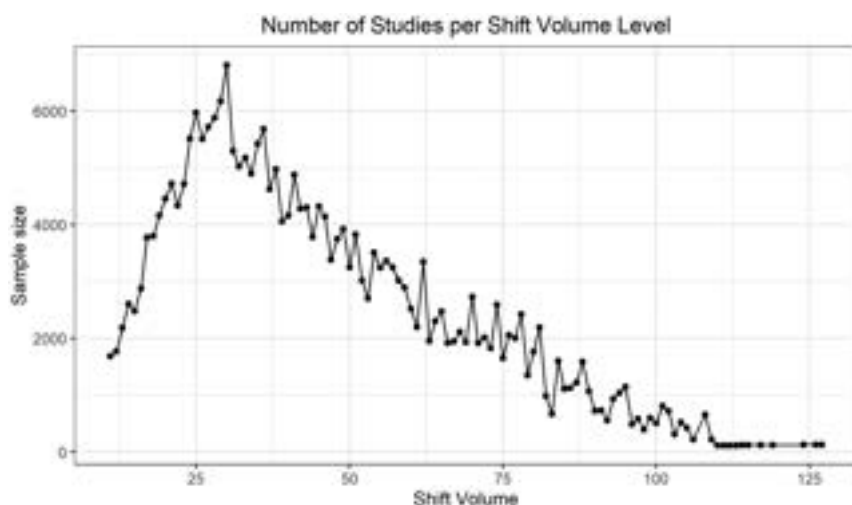
<sup>e</sup> Interpretation: The odds of containing an error are 40% higher for studies in group 2, compared with studies in group 1; the odds of containing an error are 94% higher for studies in group 3, compared with studies in group 1.

<sup>f</sup> Interpretation: The odds of containing an error are 54% higher for studies in group 2, compared studies in group 1; the odds of containing an error are 91% higher for studies in group 3, compared with studies in group 1; the odds of containing an error are 30% higher for studies in group 4, compared with studies in group 1.

<sup>g</sup> Error rate per 1000 exams.

<sup>h</sup> Optimal cutoff points are defined as the ones with the most significant (likelihood ratio test) split.

<sup>i</sup> Odds ratio 3 decreases, seemingly suggesting that group 4 is safer (than 2); however, this is due to sample size limitations in group 4.



**FIG 2.** Sample size for each shift volume level.

by specialty and subspecialty radiologists,<sup>9,10,31,32</sup> lower shift volumes,<sup>15</sup> limits on shift lengths to less than 10 hours,<sup>15,33</sup> reductions in noninterpretive tasks during clinical shifts,<sup>34</sup> radiologist participation at multispecialty tumor boards,<sup>20</sup> awareness of most common types of misses,<sup>18,19,21</sup> and reduced interpretation speed.<sup>35</sup> Subspecialty interpretations of neuroradiologic studies can change clinical care up to 41% of the time when compared with those of general radiologists or nonspecialized neuroradiologist interpretations.<sup>9,32</sup> Radiologist workloads have been steadily increasing over recent decades.<sup>36,37</sup> If we can develop more uniform policies and hospital business models in which shift volume limits and subspecialized interpretations are incentivized to act synergistically, as has been successfully done at many institutions across the country, the error reduction rates might approximate those of the airline industry or anesthesia departments. In 2012, the Royal College of Radiologists issued a national guideline recommending that radiologists interpret up to 2 complex CT or MR imaging examinations and up to 6 CT or MR imaging examinations, overall, per hour.<sup>38</sup> In 2022, the Japanese College of Radiology issued a national guideline recommending that radiologists interpret up to 4 examinations per hour.<sup>39</sup> The American College of Radiology has not issued official guidelines aimed at reducing error rates in the United States.

The results of this study can be used to guide future strategies. In our institution, the initial step has been to reduce the effect of extreme shift volumes by adjusting workflow without increasing shift frequency. Flexibility toward a labor division based on seniority level and experience with certain studies is encouraged as part of the process. Promoting educational and research efforts, multidisciplinary consultations, and peer-review discussions allows us to maintain our focus on quality and deepening our expertise. Mid- and long-term goals of limiting shift volume within the 19–26 range should then facilitate adjustments to the traditional business models and should thereby result in increased job attractiveness and more effective support to cohesive and motivated teams while maintaining profitability.

The limitations of our study include its single-center retrospective design. Second, though our study contains a large number of detected errors compared with other publications, only a fraction

of potential errors was detected, given our methods of data collection. Third, all radiologists were fellowship-trained neuroradiologists at an academic medical center. Thus, our findings may not apply for general radiology practice, for neuro-radiologists at community practices, or for academic neuroradiologists who have significantly different workflows than ours. Ideally, each institution should conduct an analysis of its own data. Fourth, we did not include confounders in our analysis. Based on prior work, neuroradiologists' years of experience, trainee participation, and emergency interpretation setting did not have an impact on diagnostic errors, and shift volume was shown to be the dominant variable that influences outcomes.<sup>25</sup> Shift

length was not evaluated as a risk factor for diagnostic error, and the same volume threshold for a 6-hour shift should be different in comparison with that of a 9-hour or 12-hour shift; our shifts were generally shorter than 9 hours, whereas prior literature suggests a significant increase in errors beyond a threshold of 10 hours.<sup>15,33</sup> Fifth, the analysis did not account for examination complexity (number of findings or diagnoses per study), which would be virtually impossible to evaluate, given the >250,000 studies performed during the study period. It might be reasonable to assume that higher volume shifts could have a higher proportion of negative or simple studies (negative CT head) and a lower proportion of complex MR imaging studies. Sixth, sub-analyses were not performed with respect to specific RADPEER scores (2a, 2b, 3a, or 3b). Further, prospective research is needed to determine shift volume safety cutoff points at the group and individual levels, using shift volume as the only variable and holding all possible confounders constant. Seventh, the lack of full covariate information on attending radiologists limited the analytic approaches that could be employed to handle potential intraclass correlation. Future studies should give thought to handling potential correlation though direct modeling.

## CONCLUSIONS

Limiting the number of studies per shift to a low range, as opposed to high, unregulated volumes, should allow for the prevention of a significant number of diagnostic neuroradiology errors. Our study of shift volumes at a busy multispecialty, tertiary, large referral academic medical center suggests that neuroradiologic shift volumes should be limited to a 19–26 range per shift.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

1. Makary MA, Daniel M. **Medical error—the third leading cause of death in the US.** *BMJ* 2016;353:i2139 CrossRef Medline
2. The flight safety foundation's aviation safety network. Fatalities per million passengers. <https://aviation-safety.net/statistics/>. Accessed September 23, 2023

3. Li G, Warner M, Lang BH, et al. **Epidemiology of anesthesia-related mortality in the United States, 1999–2005.** *Anesthesiology* 2009;110:759–65 CrossRef Medline
4. Wu MZ, McInnes MDF, Macdonald DB, et al. **CT in adults: Systematic review and meta-analysis of interpretation discrepancy rates.** *Radiology* 2014;270:717–35 CrossRef Medline
5. Babiarz LS, Yousem DM. **Quality control in neuroradiology: discrepancies in image interpretation among academic neuroradiologists.** *AJNR Am J Neuroradiol* 2012;33:37–42 CrossRef Medline
6. Borgstede JP, Lewis RS, Bhargavan M, et al. **RADPEER quality assurance program: a multifacility study of interpretive disagreement rates.** *J Am Coll Radiol* 2004;1:59–65 CrossRef
7. Kim YW, Mansfield LT. **Fool me twice: delayed diagnoses in radiology with emphasis on perpetuated errors.** *AJR Am J Roentgenol* 2014;202:465–70 CrossRef Medline
8. Garland LH. **On the scientific evaluation of diagnostic procedures.** *Radiology* 1949;52:309–28 CrossRef Medline
9. Loevner LA, Sonners AI, Schulman BJ, et al. **Reinterpretation of cross-sectional images in patients with head and neck cancer in the setting of a multidisciplinary cancer center.** *AJNR Am J Neuroradiol* 2002;23:1622–26 Medline
10. Rosenkrantz AB, Duszak R, Babb JS, et al. **Discrepancy rates and clinical impact of imaging secondary interpretations: a systematic review and meta-analysis.** *J Am Coll Radiology* 2018;15:1222–31 CrossRef Medline
11. Kabadi SJ, Krishnaraj A. **Strategies for improving the value of the radiology report: a retrospective analysis of errors in formally over-read studies.** *J Am Coll Radiol* 2017;14:459–66 CrossRef Medline
12. Revesz G, Kundel HL. **Psychophysical studies of detection errors in chest radiology.** *Radiology* 1977;123:559–62 CrossRef Medline
13. Quekel LGBA, Kessels AGH, Goei R, et al. **Miss rate of lung cancer on the chest radiograph in clinical practice.** *Chest* 1999;115:720–24 CrossRef Medline
14. Abujudeh HH, Boland GW, Kaewlai R, et al. **Abdominal and pelvic computed tomography (CT) interpretation: discrepancy rates among experienced radiologists.** *Eur Radiology* 2010;20:1952–57 CrossRef Medline
15. Hanna TN, Lamoureux C, Krupinski EA, et al. **Effect of shift, schedule, and volume on interpretive accuracy: a retrospective analysis of 2.9 million radiologic examinations.** *Radiology* 2018;287:205–12 CrossRef Medline
16. Kasalak Ö, Alnahwi H, Toxopeus R, et al. **Work overload and diagnostic errors in radiology.** *Eur J Radiology* 2023;167:111032 CrossRef Medline
17. Ivanovic V, Paydar A, Chang Y-M, et al. **Impact of shift volume on neuroradiology diagnostic errors at a large tertiary academic center.** *Acad Radiology* 2023;30:1584–88 CrossRef Medline
18. Vong S, Chang J, Assadsangabi R, et al. **Analysis of perceptual errors in skull-base pathology.** *Neuroradiol J* 2023;36:515–23 CrossRef Medline
19. Biddle G, Assadsangabi R, Broadhead K, et al. **Diagnostic errors in cerebrovascular pathology: retrospective analysis of a neuroradiology database at a large tertiary academic medical center.** *AJNR Am J Neuroradiol* 2022;43:1271–78 CrossRef Medline
20. Ivanovic V, Assadsangabi R, Hacein-Bey L, et al. **Neuroradiology diagnostic errors at a tertiary academic centre: effect of participation in tumour boards and physician experience.** *Clin Radiology* 2022;77:607–12 CrossRef Medline
21. Ferguson A, Assadsangabi R, Chang J, et al. **Analysis of misses in imaging of head and neck pathology by attending neuroradiologists at a single tertiary academic medical centre.** *Clin Radiology* 2021;76:786.e9–e13 CrossRef Medline
22. Ivanovic V, Broadhead K, Beck R, et al. **Factors associated with neuroradiologic diagnostic errors at a large tertiary-care academic medical center: a case-control study.** *AJR Am J Roentgenol* 2023;221:355–62 CrossRef Medline
23. Goldberg-Stein S, Frigini LA, Long S, et al. **ACR RADPEER committee white paper with 2016 updates: Revised scoring system, new classifications, self-review, and subspecialized reports.** *J Am Coll Radiology* 2017;14:1080–86 CrossRef Medline
24. Chang C, Hsieh MK, Chang WY, et al. **Determining the optimal number and location of cutoff points with application to data of cervical cancer.** *PLoS One* 2017;12:e0176231 CrossRef Medline
25. Ivanovic V, Broadhead K, Beck R, et al. **Factors associated with neuroradiology diagnostic errors at a large tertiary-care academic medical center: a case-control study.** *AJR Am J Roentgenol* 2023;221:355–62 CrossRef
26. Chakraborty H, Sen PK. **Resampling method to estimate intra-cluster correlation for clustered binary data.** *Commun Stat Theory Methods* 2016;45:2368–77 CrossRef
27. Wintermark M, Gupta V, Hess C, et al. **Realistic productivity in academic neuroradiology: a national survey of neuroradiology division chiefs.** *AJNR Am J Neuroradiol* 2023;44:759–61 CrossRef Medline
28. Mahesh M, Ansari AJ, Mettle FA. **Patient exposure from radiologic and nuclear medicine procedures in the United States and worldwide: 2009–2018.** *Radiology* 2023;307:e239013 CrossRef Medline
29. Deckey DG, Eltorai AEM, Jindal G, et al. **Analysis of malpractice claims involving diagnostic and interventional neuroradiology.** *J Am Coll Radiology* 2019;16:764–69 CrossRef Medline
30. Khan A, Khunte M, Wu X, et al. **Malpractice litigation related to diagnosis and treatment of intracranial aneurysms.** *AJNR Am J Neuroradiol* 2023;44:460–66 CrossRef Medline
31. Hatzoglou V, Omuro AM, Haque S, et al. **Second-opinion interpretations of neuroimaging studies by oncologic neuroradiologists can help reduce errors in cancer care.** *Cancer* 2016;122:2708–14 CrossRef Medline
32. Lysack JT, Hoy M, Hudon ME, et al. **Impact of neuroradiologist second opinion on staging and management of head and neck cancer.** *J Otolaryngol Head Neck Surg* 2013;42:39 CrossRef Medline
33. Ruutinen AT, Durand DJ, Scanlon MH, et al. **Increased error rates in preliminary reports issued by radiology residents working more than 10 consecutive hours overnight.** *Acad Radiology* 2013;20:305–11 CrossRef Medline
34. Lee MH, Schemmel AJ, Pooler BD, et al. **Workflow dynamics and the imaging value chain: quantifying the effect of designating a nonimage-interpretive task workflow.** *Curr Probl Diagn Radiology* 2017;46:275–81 CrossRef Medline
35. Sokolovskaya E, Shinde T, Ruchman RB, et al. **The effect of faster reporting speed for imaging studies on the number of misses and interpretation errors: a pilot study.** *J Am Coll Radiology* 2015;12:683–88 CrossRef Medline
36. Lu Y, Arenson RL. **The academic radiologist's clinical productivity: an update.** *Acad Radiology* 2005;12:1211–23 CrossRef Medline
37. 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 Radiology* 2015;22:1191–98 CrossRef Medline
38. The Royal College of Radiologists website. Clinical radiology workload: guidance on radiologists' reporting figures. www.rcr.ac.uk. Accessed September 28, 2023
39. Japanese College of Radiology webpage. JCR statement on appropriate workload of radiologists. jcr.or.jp/english-page. Accessed September 28, 2023

# Clinical Evaluation of a 2-Minute Ultrafast Brain MR Protocol for Evaluation of Acute Pathology in the Emergency and Inpatient Settings

Min Lang, Bryan Clifford, Wei-Ching Lo, Brooks P. Applewhite, Azadeh Tabari, Augusto Lio M. Goncalves Filho, Zahra Hosseini, Maria Gabriela Figueiro Longo, Stephen F. Cauley, Kawin Setsompop, Berkin Bilgic, Thorsten Feiweier, Michael H. Lev, Pamela W. Schaefer, Otto Rapalino, Susie Y. Huang, and John Conklin



## ABSTRACT

**BACKGROUND AND PURPOSE:** The use of MR imaging in emergency settings has been limited by availability, long scan times, and sensitivity to motion. This study assessed the diagnostic performance of an ultrafast brain MR imaging protocol for evaluation of acute intracranial pathology in the emergency department and inpatient settings.

**MATERIALS AND METHODS:** Sixty-six adult patients who underwent brain MR imaging in the emergency department and inpatient settings were included in the study. All patients underwent both the reference and the ultrafast brain MR protocols. Both brain MR imaging protocols consisted of T1-weighted, T2/T2\*-weighted, FLAIR, and DWI sequences. The ultrafast MR images were reconstructed by using a machine-learning assisted framework. All images were reviewed by 2 blinded neuroradiologists.

**RESULTS:** The average acquisition time was 2.1 minutes for the ultrafast brain MR protocol and 10 minutes for the reference brain MR protocol. There was 98.5% agreement on the main clinical diagnosis between the 2 protocols. In head-to-head comparison, the reference protocol was preferred in terms of image noise and geometric distortion ( $P < .05$  for both). The ultrafast ms-EPI protocol was preferred over the reference protocol in terms of reduced motion artifacts ( $P < .01$ ). Overall diagnostic quality was not significantly different between the 2 protocols ( $P > .05$ ).

**CONCLUSIONS:** The ultrafast brain MR imaging protocol provides high accuracy for evaluating acute pathology while only requiring a fraction of the scan time. Although there was greater image noise and geometric distortion on the ultrafast brain MR protocol images, there was significant reduction in motion artifacts with similar overall diagnostic quality between the 2 protocols.

**ABBREVIATIONS:** EPI Mix = multicontrast echo-planar imaging; ms-EPI = multishot echo-planar imaging; ss-EPI = single-shot echo-planar imaging

MR imaging offers exquisite sensitivity and accurate characterization of many acute neurologic conditions, including stroke, traumatic brain injury, and intracranial infection. Utilization of MR imaging as the initial imaging technique in the emergency and inpatient settings, however, has been limited

because of long scan time and image quality degradation by patient motion. Lack of timely access to MR imaging for acute ischemic stroke and for hospitalized patients, in general, has been shown to be associated with longer length of hospitalization.<sup>1-4</sup> Delays in diagnosis and management, as well as longer length of hospitalization, contribute to greater institutional cost burden.<sup>5</sup> Therefore, there is a growing need to increase accessibility to emergent MR imaging use to improve the diagnosis and management of a wide variety of acute intracranial pathologies.<sup>6-8</sup>

Long scan times associated with MR imaging result in high sensitivity to motion, a particular challenge in emergency department and inpatient settings. Andre et al<sup>5</sup> found that there was a 4-fold increase in the incidence of moderate to severe motion artifacts for inpatient and emergency department examinations (29.4%) compared with outpatient examinations (7.5%).<sup>5</sup>

Techniques to improve the efficiency of clinical brain MR examinations in acute settings have gained increasing attention. Approaches by using 2D TSE sequences and conventional parallel imaging have been previously validated but suffer from substantial signal-noise-ratio losses when pushed to higher

Received July 17, 2023; accepted after revision December 7.

From the Department of Radiology (M.L., B.P.A., A.T., M.G.F.L., M.H.L., P.W.S., O.R., S.Y.H., J.C.), Massachusetts General Hospital, Boston, Massachusetts; Harvard Medical School (M.L., B.P.A., A.T., M.G.F.L., M.H.L., P.W.S., O.R., S.Y.H., J.C.), Boston, Massachusetts; Siemens Medical Solutions (B.C., W.-C.L., Z.H., S.F.C.), Boston, Massachusetts; Department of Radiology (A.L.M.G.F.), University of Ottawa, Ottawa, Ontario, Canada; Harvard-MIT Health Sciences and Technology (S.F.C., B.B., S.Y.H.), Massachusetts Institute of Technology, Cambridge, Massachusetts; Departments of Radiology and Electrical Engineering (K.S.), Stanford University, Stanford, California; and Siemens Healthcare GmbH (T.F.), Erlangen, Germany.

Supported by U.S. Department of Health and Human Services, National Institutes of Health, P41EB030006, RSNA Research and Education Foundation, Seed Grant, Siemens Healthineers.

Please address correspondence to Min Lang, MD, Department of Radiology, Massachusetts General Hospital, 55 Fruit St, Boston, MA 02141; e-mail: mlang@mgm.harvard.edu; @MinLangMD



Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8143>

acceleration factors, limiting the acceleration that can be achieved.<sup>9,10</sup> Alternatively, single-shot echo-planar imaging (ss-EPI) sequences, which acquire all  $k$ -space data in a single excitation, have been recently proposed to achieve more dramatic acceleration.<sup>11,12</sup> However, ss-EPI suffers from artifacts, including geometric distortion, signal drop-out, and T2\* mediated blurring, limiting diagnostic quality, and clinical utility.

Multishot echo-planar imaging (ms-EPI) is a highly efficient interleaved EPI technique that utilizes multiple excitations to reduce echo-train length.<sup>13</sup> The increased pixel bandwidth along the phase encoding direction, which comes along with segmented  $k$ -space acquisition, reduces off-resonance effects, such as geometric distortion, susceptibility artifact, and signal drop-out.<sup>13-16</sup> In addition, artificial intelligence-assisted reconstruction has achieved reconstruction times that are clinically negligible and are able to further reduce noise and aliasing in accelerated MR imaging techniques. We applied the ms-EPI technique to develop a 2-minute ultrafast brain MR imaging protocol consisting of ms-EPI accelerated T1-weighted, T2/T2\*-weighted, FLAIR, and DWI sequences. The accelerated images are further denoised by using artificial intelligence-assisted reconstruction. This study investigated the clinical feasibility and diagnostic performance of the 2-minute ultrafast brain MR imaging protocol in the emergency department and inpatient settings.

## MATERIALS AND METHODS

This prospective comparative study was performed at a single institution (Massachusetts General Hospital), and was approved by the Institutional Review Board and was compliant with the Health Insurance Portability and Accountability Act. The Institutional Review Board waived the need for signed informed consent.

### Study Population and MR Protocol

Consecutive adult patients >18 years of age who underwent clinical brain MR imaging with the “routine brain without contrast” MR imaging protocol in the emergency and inpatient settings were included in the study. All patients underwent brain MR imaging with the 10-minute reference brain MR imaging protocol first (parameters are listed in Online Supplemental Data), followed by the ultrafast 2-minute brain MR imaging examination (parameters are listed in Online Supplemental Data). The 2-minute ultrafast brain MR imaging protocol was implemented by using an ms-EPI research pulse sequence and reconstruction algorithm. Details of the ms-EPI acquisition parameters and the 2-minute ultrafast brain MR imaging protocol are described by Clifford et al.<sup>17</sup> The 2-minute ultrafast brain MR imaging protocol consisted of 5 sequences: sagittal T1-weighted imaging, axial T2-weighted imaging, axial T2\*-weighted imaging, axial FLAIR, and axial DWI (including ADC reconstruction from the DWIs). Data for the T2 and T2\*-weighted images were obtained from the same spin-echo scan via the incorporation of an additional free induction decay (FID) readout before the refocusing pulse (for more details, see Clifford et al.,<sup>17</sup> their supplemental data). The reference routine brain MR imaging protocol consisted of the same 5 sequences with standard parameters used at our institution. All brain MR examinations were performed on one of two

3T scanners at our institution (Magnetom Prisma and Magnetom Skyra, Siemens Healthineers), by using either a 20-channel head-neck coil or a 32-channel head-only coil. Demographics of the study subjects and clinical indications for MR imaging are listed in Online Supplemental Data.

### Machine-Learning Assisted Reconstruction

The raw  $k$ -space data from the ultrafast brain MR protocol were extracted from the scanners within 48 hours of acquisition and retrospectively reconstructed by using machine-learning assisted reconstruction. The machine-learning reconstruction algorithm mitigated aliasing and g-factor noise amplification to improve the SNR of the highly accelerated ms-EPI. A detailed description of the reconstruction method was previously described by Clifford et al.<sup>17</sup> Briefly, the training data for the neural network were previously collected on a 3T system (Magnetom Prisma) from 16 healthy subjects (8 men, 8 women, ages 19–67 years). The machine-learning-based reconstruction incorporated a tunable parameter for controlling the level of denoising and has been previously validated across various acceleration factors, contrasts, and SNR conditions.

### Clinical Evaluation of the Optimized 2-Minute Brain MR Imaging Protocol

All images were reviewed independently by 2 neuroradiologists (O.R. and B.P.A., with 21 and 6 years of experience, respectively). The reviewing neuroradiologists were blinded to the clinical history and image acquisition method. The main clinical diagnosis was scored as a categorical variable with options of “no acute findings,” “recent infarct,” “intracranial hemorrhage,” “mass,” or “hydrocephalus.” For all cases, the ms-EPI was reviewed first, without seeing the reference sequences, and the neuroradiologists were asked to provide the main clinical diagnosis. The reference sequence images were then subsequently reviewed and the neuroradiologists were asked if there was a change in final diagnosis after reviewing the reference sequences.

Overall image quality was evaluated in a blinded head-to-head comparison between the 2 protocols. The screen position (left versus right) of the ms-EPI sequence images and the counterpart reference images, as well as the order of the cases were all randomized. All cases were graded on a 5-point Likert scale, where positive numbers favored the images on the right side of the screen and negative number favored the images on the left side of the screen. The following image quality variables were evaluated by using this scale: image noise, motion artifact, geometric distortion, other artifacts, and overall diagnostic quality. Disagreements between readers were adjudicated by a third neuroradiologist (J.C.).

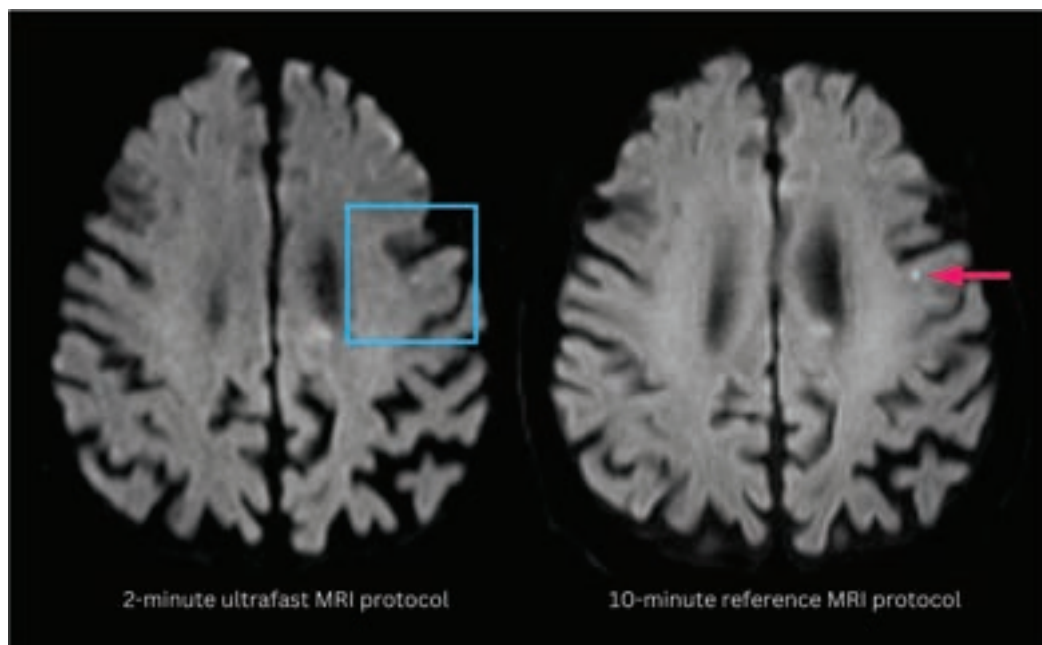
### Statistical Analysis

Comparative image quality analysis between the ultrafast and reference images was performed by using the nonparametric Wilcoxon rank sum test. Statistical analysis was performed by using RStudio. Statistical significance was set at  $P < .05$ .

## RESULTS

A total of 66 patients were included in the study, with 37 women and 29 men (Online Supplemental Data). The mean age was  $58.7 \pm 19$  years. The most common indications for MR imaging





**FIG 1.** A punctate focus of restricted diffusion was less conspicuous on the ultrafast DWI image (*left*) compared with the reference DWI image (*right*). The decreased conspicuity may be due to a combination of differences in image quality between the 2 protocols and section positioning leading to partial volume averaging.

were altered mental status (21.2%), suspected stroke (18.2%), headache (13.6%), vertigo (12.1%), transient ischemic attack (9.1%), suspected tumor (7.6%), vision loss (4.5%), seizure (1.5%), and ataxia (1.5%). The mean acquisition time for the reference brain MR protocol was 10 minutes, compared with 2:06 minutes for the ultrafast brain MR protocol.

There was agreement on the categorically defined main clinical diagnosis in all but 1 case (98.5% agreement between the 2 protocols). The main categorical diagnosis of the 66 examinations after reviewing both the ultrafast and reference protocol images was no acute findings for 71.2% of cases, acute to subacute infarct for 10.6% of cases, intracranial hemorrhage for 6.1% of cases, intracranial mass lesion for 6.1% of cases, old infarct for 3.0% of cases, and Chiari type I malformation for 3.0% of cases. The ultrafast MS-EPI protocol provided excellent image quality for visualization of the various clinical pathologies encountered.

On the single case with discrepant main diagnosis by the 2 reviewing neuroradiologists, there was punctate focus of DWI hyperintensity that may represent acute to subacute infarct that was less conspicuous on the ultrafast DWI (Fig 1). After further reviewing this case with a third adjudicating neuroradiologist, the discrepancy in finding may have been due to a combination of difference in image quality between the 2 protocols and slight differences in section positioning leading to partial volume contamination as the lesion was likely smaller than the section thickness.

In a head-to-head comparison of image quality between the ultrafast and the reference brain MR protocols, the reference protocol was preferred by the evaluating neuroradiologists in terms of image noise and geometric distortion ( $P < .05$  for both; Fig 2). The ultrafast ms-EPI protocol was preferred over the reference protocol in terms of motion artifacts ( $P < .01$ ). There

was no significant difference between the 2 protocols in terms of overall diagnostic quality (Fig 2).

Representative images from 4 clinical cases are shown in Figures 3–6, demonstrating similar diagnostic quality between the ultrafast brain MR protocol images and the reference brain MR protocol images. Figure 3 shows a case of amyloid angiopathy with scattered foci of microhemorrhages. There is blurring of the findings on the reference GRE sequence due to motion artifact, and the foci of susceptibility signal are better appreciated on the ultrafast T2\* images. Figure 4 shows a case of herpes simplex virus encephalitis with T2/FLAIR hyperintensity and restricted diffusion in the left greater than right mesial temporal lobes, visualized equally well on both the reference and the ultrafast brain MR protocols. Figure 5 shows a case of punctate subacute infarct in the left centrum semiovale on a background of chronic small vessel ischemic disease, visualized equally well on both the reference and the ultrafast brain MR protocols demonstrating ill-defined T2/FLAIR hyperintensity centered in the left thalamus.

## DISCUSSION

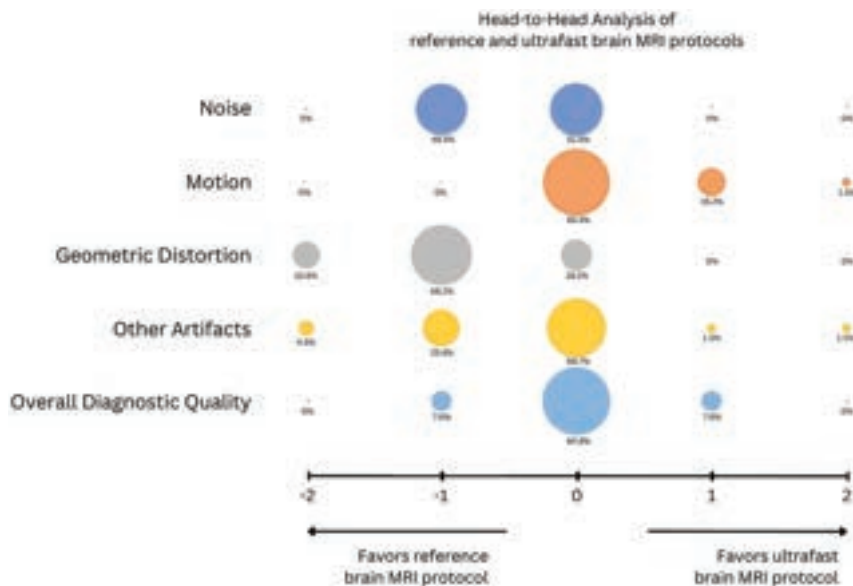
There has been growing interest in accelerated brain MR imaging techniques in recent years, especially for vulnerable patient populations, such as pediatric patients and those presenting with acute symptoms in the emergency and inpatient settings. These patient populations are often motion prone and are in settings where timely diagnosis is critical to guiding treatment, management, and disposition on expedited time scales. In our study, we found that the 2-minute ultrafast ms-EPI brain MR protocol, which was 8 minutes (80%) faster than the reference protocol, provided

similar overall diagnostic image quality and was concordant with the reference brain MR imaging protocol for the main clinical diagnosis in all but 1 case (98.5%).

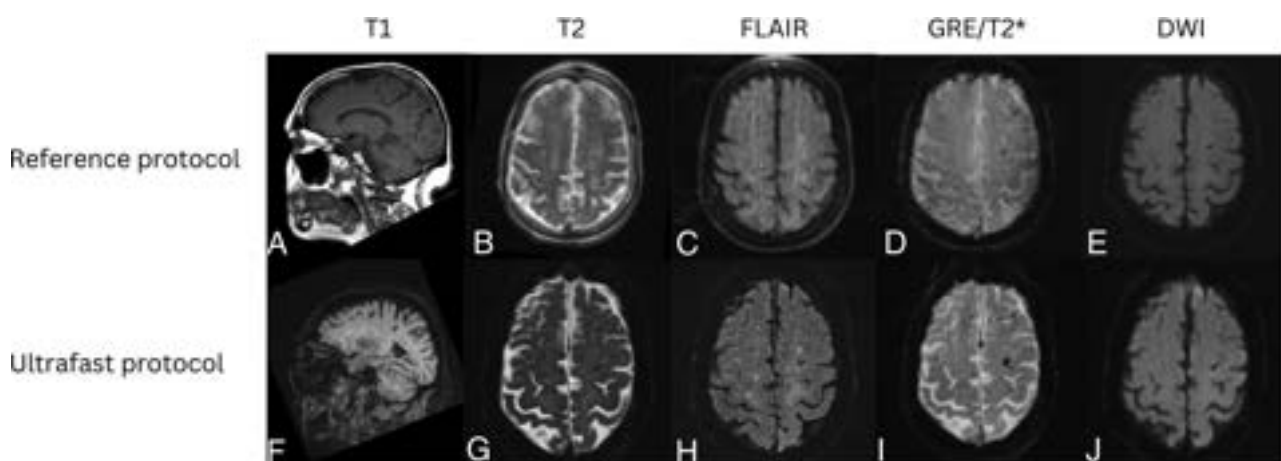
Other brain MR acceleration techniques, including single-shot EPI and previously reported variations, such as multicontrast EPI (EPIMix), have been reported.<sup>9,10</sup> However, at high acceleration factors ( $R > 2$ ), these techniques can suffer from increased noise, geometric distortion, and residual aliasing artifacts.<sup>18,19</sup> A major limitation of single-shot echo-planar FLAIR imaging is the poor tissue contrast between WM and GM.

EPIMix, on the other hand, is an ultrafast brain imaging technique by using single-shot EPI to rapidly acquire multiple contrasts.<sup>11,12,20</sup> ms-EPI uses an interleaved EPI method that utilizes multiple excitations, resulting in significantly reduced geometric distortion and higher SNR than single-shot-EPI.<sup>14-16</sup> The proposed multishot EPI approach employed in this paper requires slightly greater scan time (because of the acquisition of multiple shots) with the expected trade-off of improved image quality in the form of increased SNR and reduced image distortion compared with a single-shot EPI approach, including EPIMix. Single-shot EPI may still have a valuable role in extremely motion-prone patients, or when extreme scan speed is desired.

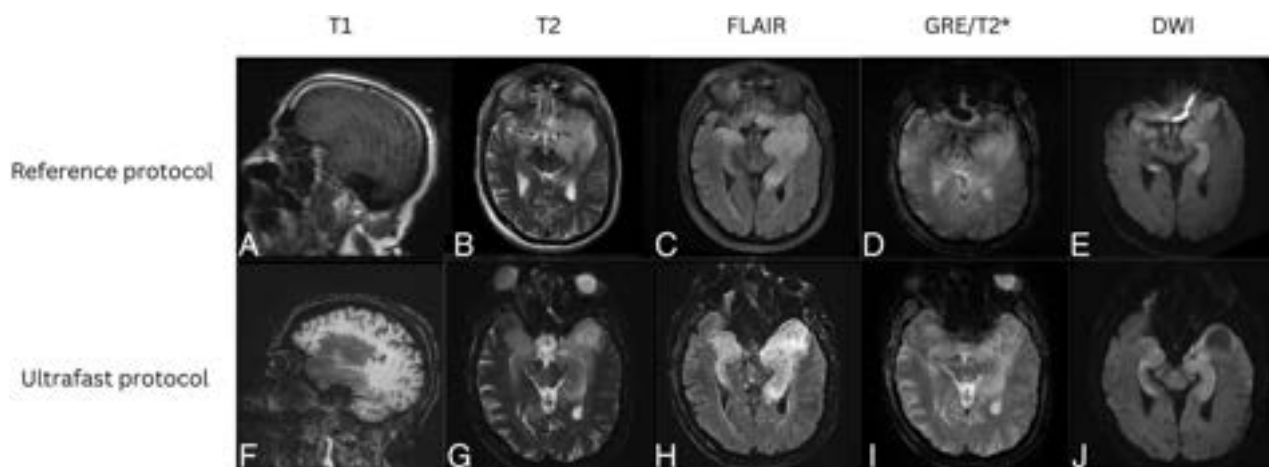
The ultrafast brain protocol used in this study is based on the ms-EPI technique with a machine-learning assisted reconstruction framework that was previously described by Clifford et al.<sup>17</sup> Despite its inherently increased sensitivity to susceptibility-induced variations of the magnetic field, EPI has been previously suggested as an alternative to TSE-based T2-weighted imaging in the clinical setting.<sup>16,21</sup> There are some known limitations to the technique, including signal loss of the extracranial soft tissues and mild geometric distortion in the temporal regions, which are inherent to EPI and are partly mitigated by the multishot acquisition. As expected, the reviewing neuroradiologists rated overall image quality, image noise and geometric distortion to be significantly better in the reference protocol images compared with the ultrafast protocol images.



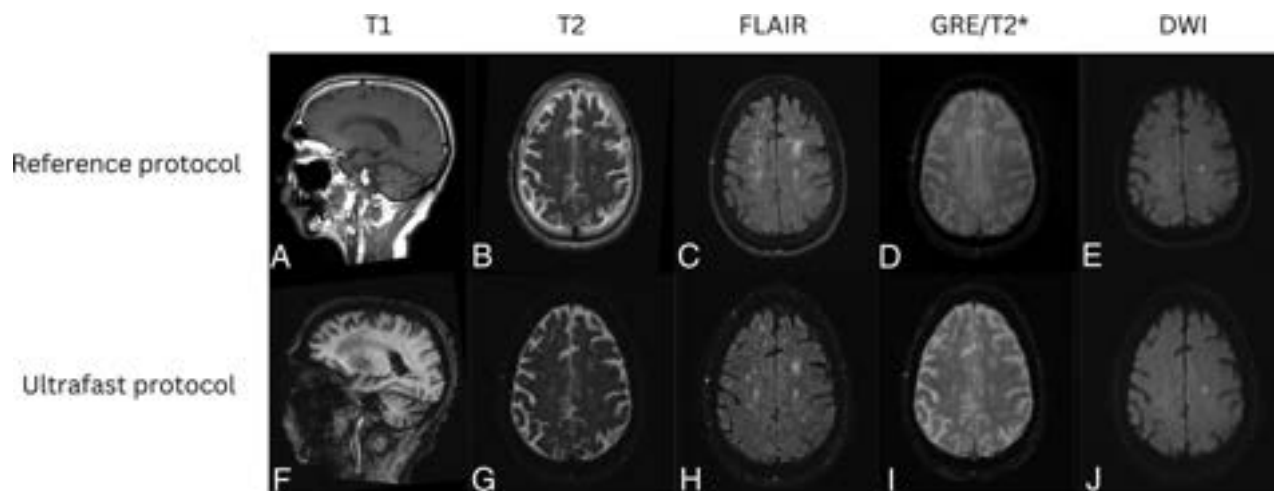
**FIG 2.** Bubble plot shows head-to-head comparison between the reference brain MR protocol and the ultrafast ms-EPI protocol. Negative scores indicate preference of the clinical reference protocol; 0 indicates equivalence between the 2 protocols; and positive scores indicate preference of the ultrafast ms-EPI protocol. The reference protocol was preferred by the evaluating neuroradiologists in terms of image noise and geometric distortion ( $P < .05$  for both). The ultrafast protocol was preferred over the reference protocol in terms of motion artifacts ( $P < .05$ ). There was no significant difference between the 2 protocols in terms of overall diagnostic quality.



**FIG 3.** Patient with a diagnosis of amyloid angiopathy. Reference sagittal T1-weighted (A), axial T2-weighted (B), FLAIR (C), SWI (D), and DWI (E) show scattered foci of susceptibility signal in the left greater than right frontal lobes consistent with chronic microhemorrhages. The findings were more conspicuous on the ultrafast sagittal T1-weighted (F), axial T2-weighted (G), FLAIR (H), SWI (I), and DWI (J) because of motion artifact on the reference MR images.



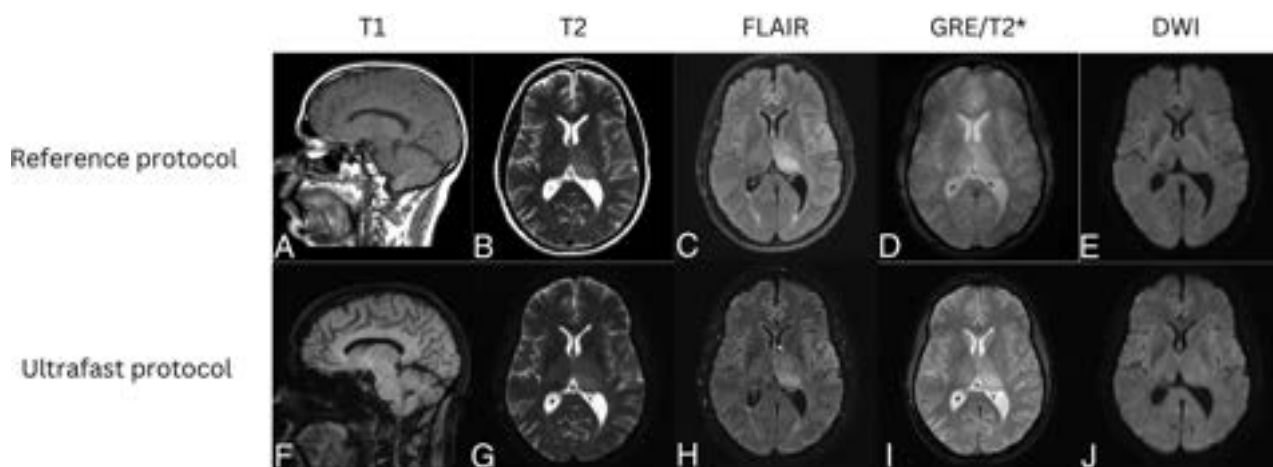
**FIG 4.** Patient with a diagnosis of herpes simplex virus encephalitis. Reference sagittal T1-weighted (A), axial T2-weighted (B), FLAIR (C), SWI (D), and DWI (E) show prominent T1 hypointensity and T2/FLAIR hyperintensity in the left greater than right mesial temporal lobes with scatter foci of susceptibility signal in the left temporal lobe consistent with microhemorrhages. There was associated restricted diffusion in the left greater than right mesial temporal lobes. These signal abnormalities were all seen with similar conspicuity on the ultrafast sagittal T1-weighted (F), and axial T2-weighted (G), FLAIR (H), SWI (I), and DWI (J).



**FIG 5.** Patient with punctate subacute infarct in the left corona radiata. Reference sagittal T1-weighted (A), axial T2-weighted (B), FLAIR (C), SWI (D), and DWI (E) show a punctate focus of restricted diffusion with associated FLAIR hyperintensity in the left centrum semiovale on a background of white matter T2/FLAIR hyperintensities that likely represent chronic small vessel ischemic disease. The same findings were seen on the ultrafast sagittal T1-weighted (F), axial T2-weighted (G), FLAIR (H), SWI (I), and DWI (J).

More importantly, however, the overall diagnostic quality of the ultrafast protocol was not significantly worse and was considered to be equivalent compared with the reference standard. In the acute clinical settings, such as the emergency department and the inpatient wards, time to diagnosis, treatment, and management directly correlate with clinical outcomes, including mortality and morbidities, for multiple neurologic conditions.<sup>22</sup> In addition to patient outcomes, a delay in diagnosis and management also adds to the burden of the hospital systems as it may increase length of hospitalization, the need for increased hospital capacity, and the need for additional health care services.<sup>23</sup> Thus, with equivalent overall clinical diagnostic utility, the ultrafast brain MR imaging may be an appropriate first imaging examination of choice to exclude life-threatening and emergent neurologic pathology.

This study was focused on patients with acute neurologic symptoms and adds to a growing body of literature evaluating ultrafast MR imaging protocols in the brain. A recently published study evaluating patients with acute stroke imaged with a deep-learning enhanced 2-minute multicontrast EPI examination found that characterization of patients with stroke by using this protocol was equivalent to the reference sequences.<sup>24</sup> Another recent study showed that a deep-learning assisted ultrafast multishot EPI examination implemented at 1.5T was effective in detecting acute intracranial pathology.<sup>25</sup> Patients evaluated in the acute setting may have underlying chronic or nonemergent intracranial pathology that would require additional follow-up and consultation. Further evaluation is needed to determine the efficacy of the ultrafast brain MR protocol for detecting nonacute intracranial findings. Furthermore, by



**FIG 6.** Patient with a low-grade left thalamic glioma. Reference sagittal T1-weighted (A), and axial T2-weighted (B), FLAIR (C), SWI (D), and DWI (E) show an ill-defined, mildly expansile, T1 hypointense, and T2/FLAIR hyperintense lesion centered in the left thalamus, compatible with diagnosis of low-grade glioma. Similar findings were appreciated on the ultrafast sagittal T1-weighted (F), axial T2-weighted (G), FLAIR (H), SWI (I), and DWI (J).

greatly accelerating MR acquisition across different contrasts, there is a trade-off of limited evaluation of extracranial structures. This needs to be taken into account when evaluating ms-EPI and clinical suspicion for extracranial processes will require further evaluation with standard MR protocols.

In our study, there was only 1 case (of 66) where there was a finding (punctate focus of cortical restricted diffusion) that was not as well visualized on the ultrafast brain protocol but was seen on the reference protocol. Upon further review, the finding may have been due to differences in section positioning and section thickness (4 mm for the reference DWI and 5 mm for the ultrafast DWI), leading to partial volume contamination. This does raise the question of the sensitivity of the ultrafast MR protocol at picking up small findings and whether the acquisition acceleration to exclude emergent pathology outweighs the slightly lower sensitivity at detecting small findings, which most likely have little impact on clinical care. Future studies with a larger patient population with a wide range of imaging findings is needed to provide a more detailed assessment of the ultrafast brain protocols.

The accelerated acquisition time also benefits from reduced sensitivity to patient motion as the overall acquisition time is shorter. In 11 cases (16.7%), there was less motion on the ultrafast brain protocol images than on the reference protocol images. In patients with altered mental status who do not have the ability to hold still, the ultrafast brain MR protocol may provide superior evaluation compared with the reference protocol, which may be rendered nondiagnostic from motion artifacts. A limitation of the reduced sampling of the ultrafast brain protocol ms-EPI sequences is that, if motion were to occur in between shots, motion artifacts may be inadvertently exacerbated from intershot phase error. Motion-correction techniques, such as navigator-based prospective motion correction or scout accelerated motion estimation and reduction techniques, may be helpful to address intershot phase errors with minimal impact on overall acquisition time.<sup>26,27</sup>

Pediatric patients are an important population for whom motion during MR imaging is a concern, especially infants and very young children who may have difficulty remaining still for

extended periods. General anesthesia may be required for challenging patients, but the use of anesthetics in pediatric patients has known potential negative effects, including nausea, vomiting, and disorientation, or more serious adverse events, such as cardiorespiratory depression.<sup>28</sup> In the acute clinical setting, institutions may implement abbreviated MR protocols to avoid sedation that includes only a single contrast, usually T2-weighted imaging.<sup>29-31</sup> The sensitivity of abbreviated MR protocols, however, is limited with 1 study reporting 14% of cases had undetected findings on a fast-brain MR imaging protocol.<sup>32</sup> An ultrafast 1-minute pediatric brain protocol has been previously proposed, which uses optimized faster versions of commercially available sequences.<sup>33</sup> While the preliminary image quality was good in a previously proposed ultrafast MR protocol, there was no direct evaluation of the final clinical diagnosis. Accelerated protocols, however, have a limited role in sedated patients as the patient is already under anesthesia and standard imaging will provide superior image quality.

There are several limitations to this study. First, although the patient sample size is adequate with 66 patients, most did not exhibit any acute intracranial findings. Future investigations need to include a larger sample size comprising patients with acute findings to further determine sensitivity and specificity of the ultrafast brain MR protocol. It will be important to determine the optimal balance between reducing scan time and potential slight loss in sensitivity in detecting abnormal findings. Second, this study only included adult patients in the emergency and inpatient settings. Pediatric patients, as mentioned, may greatly benefit from accelerated MR examinations; therefore, it will be important to further assess the efficacy of the ultrafast brain MR protocol in this population. Third, this study was conducted exclusively in acute clinical settings. The ultrafast brain MR protocol may also benefit patients who are claustrophobic or unable to tolerate long MR examinations for various reasons. Fourth, given the inherent differences in image characteristics of the reference and ultrafast MR imaging protocols, true blinding was difficult to achieve for side-by-side comparisons. This is unavoidable but does represent real-life situations when new sequences are introduced clinically and

radiologists are asked to evaluate the quality of the new sequence compared with the reference. Last, the neuroradiologist reviewers assessed the ultrafast brain MR protocol and the reference brain MR protocol in succession rather than by randomized evaluation. This method is more stringent on the ultrafast brain MR protocol as the reference brain MR protocol is reviewed after the ultrafast brain MR protocol to identify any missed findings.

## CONCLUSIONS

The 2-minute ultrafast brain MR imaging protocol provides high accuracy for evaluation of acute pathology in emergency department and inpatient settings. While there was greater image noise and geometric distortion on the ultrafast brain MR protocol images compared with the reference protocol, there was reduced motion artifact with similar overall diagnostic quality.

**Disclosure forms** provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

- Robinson NB, Gao M, Patel PA, et al. **Secondary review reduced inpatient MRI orders and avoidable hospital days.** *Clin Imaging* 2022;82:156–60 CrossRef Medline
- Drose JA, Pritchard NL, Honce JM, et al. **Utilizing process improvement methodology to improve inpatient access to MRI.** *Radiographics* 2019;39:2103–10 CrossRef Medline
- Tokur S, Lederle K, Terris DD, et al. **Process analysis to reduce MRI access time at a German University Hospital.** *Int J Qual Health Care* 2012;24:95–99 CrossRef Medline
- Manwani B, Rath S, Lee NS, et al. **Early magnetic resonance imaging decreases hospital length of stay in patients with ischemic stroke.** *J Stroke Cerebrovasc Dis* 2019;28:425–29 CrossRef Medline
- Andre JB, Bresnahan BW, Mossa-Basha M, et al. **Toward quantifying the prevalence, severity, and cost associated with patient motion during clinical MR examinations.** *J Am Coll Radiol* 2015;12:689–95 CrossRef Medline
- Buller M, Karis JP. **Introduction of a dedicated emergency department MR imaging scanner at the Barrow Neurological Institute.** *AJNR Am J Neuroradiol* 2017;38:1480–85 CrossRef Medline
- Pakpoor J, Saylor D, Izbudak I, et al. **Emergency department MRI scanning of patients with multiple sclerosis: worthwhile or wasteful?** *AJNR Am J Neuroradiol* 2017;38:12–17 CrossRef Medline
- Algethamy HM, Alzawahmah M, Young GB, et al. **Added value of MRI over CT of the brain in intensive care unit patients.** *Can J Neurol Sci* 2015;42:324–32 CrossRef Medline
- Prakkamakul S, Witzel T, Huang S, et al. **Ultrafast brain MRI: clinical deployment and comparison to conventional brain MRI at 3T.** *J Neuroimaging* 2016;26:503–10 CrossRef Medline
- Fagundes J, Longo MG, Huang SY, et al. **Diagnostic performance of a 10-minute gadolinium-enhanced brain MRI protocol compared with the standard clinical protocol for detection of intracranial enhancing lesions.** *AJNR Am J Neuroradiol* 2017;38:1689–94 CrossRef Medline
- Skare S, Sprenger T, Norbeck O, et al. **A 1-minute full brain MR exam using a multicontrast EPI sequence.** *Magn Reson Med* 2018;79:3045–54 CrossRef Medline
- Delgado AF, Kits A, Bystam J, et al. **Diagnostic performance of a new multicontrast one-minute full brain exam (EPIMix) in neuro-radiology: a prospective study.** *J Magn Reson Imaging* 2019;50:1824–33 CrossRef Medline
- Liu W, Zhao X, Ma Y, et al. **DWI using navigated interleaved multi-shot EPI with realigned GRAPPA reconstruction.** *Magn Reson Med* 2016;75:280–86 CrossRef Medline
- van Pul C, Roos FG, Derksen OS, et al. **A comparison study of multi-shot vs. single-shot DWI-EPI in the neonatal brain: reduced effects of ghosting compared to adults.** *Magn Reson Imaging* 2004;22:1169–80 CrossRef Medline
- Bilgic B, Chatnuntawech I, Manhard MK, et al. **Highly accelerated multishot echo planar imaging through synergistic machine learning and joint reconstruction.** *Magn Reson Med* 2019;82:1343–58 CrossRef Medline
- Poustchi-Amin M, Mirowitz SA, Brown JJ, et al. **Principles and applications of echo-planar imaging: a review for the general radiologist.** *Radiographics* 2001;21:767–79 CrossRef Medline
- Clifford B, Conklin J, Huang SY, et al. **An artificial intelligence-accelerated 2-minute multi-shot echo planar imaging protocol for comprehensive high-quality clinical brain imaging.** *Magn Reson Med* 2022;87:2453–63 CrossRef Medline
- Nael K, Khan R, Choudhary G, et al. **Six-minute magnetic resonance imaging protocol for evaluation of acute ischemic stroke: pushing the boundaries.** *Stroke* 2014;45:1985–91 CrossRef Medline
- Yanasak NE, Kelly MJ. **MR imaging artifacts and parallel imaging techniques with calibration scanning: a new twist on old problems.** *Radiographics* 2014;34:532–48 CrossRef Medline
- Kits A, De Luca F, Kolloch J, et al. **One-minute multi-contrast echo planar brain MRI in ischemic stroke: a retrospective observational study of diagnostic performance.** *J Magn Reson Imaging* 2021;54:1088–95 CrossRef Medline
- Patel MR, Siewert B, Klufas R, et al. **Echoplanar MR imaging for ultra-fast detection of brain lesions.** *AJR Am J Roentgenol* 1999;173:479–85 CrossRef Medline
- Siddique SM, Tipton K, Leas B, et al. **Interventions to reduce hospital length of stay in high-risk populations: a systematic review.** *JAMA Netw Open* 2021;4:e2125846 CrossRef Medline
- Doctoroff L, Herzig SJ. **Predicting patients at risk for prolonged hospital stays.** *Med Care* 2020;58:778–84 CrossRef Medline
- Verclytte S, Gnanh R, Verdun S, et al. **Ultrafast MRI using deep learning echoplanar imaging for a comprehensive assessment of acute ischemic stroke.** *Eur Radiol* 2023;33:3715–25 CrossRef Medline
- Altmann S, Abello Mercado MA, Brockstedt L, et al. **Ultrafast brain MRI protocol at 1.5 T using deep learning and multi-shot EPI.** *Acad Radiol* 2023;30:2988–98 CrossRef Medline
- Sarlls JE, Lalonde F, Rettmann D, et al. **Effectiveness of navigator-based prospective motion correction in MPRAGE data acquired at 3T.** *PLoS One* 2018;13:e0199372 CrossRef Medline
- Lang M, Tabari A, Polak D, et al. **Clinical evaluation of scout accelerated motion estimation and reduction technique for 3D MR imaging in the inpatient and emergency department settings.** *AJNR Am J Neuroradiol* 2023;44:125–33 CrossRef Medline
- Arlachov Y, Ganatra RH. **Sedation/anaesthesia in paediatric radiology.** *Br J Radiol* 2012;85:e1018–31 CrossRef Medline
- Engelbrecht V, Malms J, Kahn T, et al. **Fast spin-echo MR imaging of the pediatric brain.** *Pediatr Radiol* 1996;26:259–64 CrossRef Medline
- Penzkofer AK, Pfluger T, Pochmann Y, et al. **MR imaging of the brain in pediatric patients: diagnostic value of HASTE sequences.** *AJR Am J Roentgenol* 2002;179:509–14 CrossRef Medline
- Ashley WW Jr, McKinstry RC, Leonard JR, et al. **Use of rapid-sequence magnetic resonance imaging for evaluation of hydrocephalus in children.** *J Neurosurg* 2005;103:124–30 CrossRef Medline
- Rozovsky K, Ventureyra EC, Miller E. **Fast-brain MRI in children is quick, without sedation, and radiation-free, but beware of limitations.** *J Clin Neurosci* 2013;20:400–05 CrossRef Medline
- Ha JY, Baek HJ, Ryu KH, et al. **One-minute ultrafast brain MRI with full basic sequences: can it be a promising way forward for pediatric neuroimaging?** *AJR Am J Roentgenol* 2020;215:198–205 CrossRef Medline



# Increased Intracranial Arterial Pulsatility and Microvascular Brain Damage in Pseudoxanthoma Elasticum

J.W. Bartstra, T. van den Beukel, G. Kranenburg, L.J. Geurts, A.M. den Harder, T. Witkamp, J.M. Wolterink, J.J.M. Zwanenburg, E. van Valen, H.L. Koek, W.P.T.M. Mali, P.A. de Jong, J. Hendrikse, and W. Spiering



## ABSTRACT

**BACKGROUND AND PURPOSE:** Carotid siphon calcification might contribute to the high prevalence of cerebrovascular disease in pseudoxanthoma elasticum through increased arterial flow pulsatility. This study aimed to compare intracranial artery flow pulsatility, brain volumes, and small-vessel disease markers between patients with pseudoxanthoma elasticum and controls and the association between arterial calcification and pulsatility in pseudoxanthoma elasticum.

**MATERIALS AND METHODS:** Fifty patients with pseudoxanthoma elasticum and 40 age- and sex-matched controls underwent 3T MR imaging, including 2D phase-contrast acquisitions for flow pulsatility in the assessment of ICA and MCA and FLAIR acquisitions for brain volumes, white matter lesions, and infarctions. All patients with pseudoxanthoma elasticum underwent CT scanning to measure siphon calcification. Flow pulsatility (2D phase-contrast), brain volumes, white matter lesions, and infarctions (3D T1 and 3D T2 FLAIR) were compared between patients and controls. The association between siphon calcification and pulsatility in pseudoxanthoma elasticum was tested with linear regression models.

**RESULTS:** Patients with pseudoxanthoma elasticum (mean age, 57 [SD, 12] years; 24 men) had significantly higher pulsatility indexes (1.05; range, 0.94–1.21 versus 0.94; range, 0.82–1.04;  $P = .02$ ), lower mean GM volumes (597 [SD, 53] mL versus 632 [SD, 53] mL;  $P < .01$ ), more white matter lesions (2.6; range, 0.5–7.5 versus 1.1; range, 0.5–2.4) mL;  $P = .05$ ), and more lacunar infarctions (64 versus 8,  $P = .04$ ) than controls (mean age, 58 [SD, 11] years; 20 men). Carotid siphon calcification was associated with higher pulsatility indexes in patients with pseudoxanthoma elasticum ( $\beta = 0.10$ ; 95% CI, 0.01–0.18).

**CONCLUSIONS:** Patients with pseudoxanthoma elasticum have increased intracranial artery flow pulsatility and measures of small-vessel disease. Carotid siphon calcification might underlie the high prevalence of cerebrovascular disease in pseudoxanthoma elasticum.

**ABBREVIATIONS:** CSVD = cerebral small-vessel disease; IEL = internal elastic lamina; LDL = low-density lipoprotein; PI = pulsatility index; PPI = inorganic pyrophosphate; PXE = pseudoxanthoma elasticum; RR = rate ratio; WML = white matter lesions

Pseudoxanthoma elasticum (PXE, Online Mendelian Inheritance in Man [OMIM] 264800) is a rare monogenetic disorder caused by mutations in the *ABCC6* gene, which result in degradation and calcification of the elastic fibers of the skin, the Bruch membrane in the eyes, and the peripheral arteries.<sup>1,2</sup> In the skin, this disorder results in yellow papules known as pseudoxanthomas. Calcification of the Bruch membrane results in peau d'orange, angioid streaks,

choroidal neovascularizations, and macular atrophy.<sup>3</sup> Arterial calcification of the peripheral arteries of the legs results in peripheral artery disease.<sup>4</sup> Patients with PXE are at increased risk of TIAs and strokes, and several case reports have confirmed increased white matter lesions (WML).<sup>5–7</sup> Recently, mutations in the *ABCC6* gene were implicated as one of the monogenetic causes of cerebral small-vessel disease (CSVD).<sup>8</sup> However, the underlying mechanisms that cause brain damage in PXE remain uncertain. Carotid hypoplasia,<sup>9</sup> vascular inflammation,<sup>10</sup> increased elastin degradation,<sup>11</sup> and calcification might all play a role. In addition to calcification of the legs, the intracranial carotid siphon is more frequently and severely calcified compared with the general population in PXE.<sup>2</sup> This intracranial calcification might underlie the increased prevalence of cerebrovascular disease in PXE. Recently, it was shown that the carotid siphon attenuates arterial pulsatility in the carotid artery in healthy volunteers<sup>12</sup> and patients with PXE.<sup>13</sup> Furthermore, it was found that calcification of the carotid siphon is related to increased

Received February 1, 2023; accepted after revision December 2.

From the Departments of Radiology (J.W.B., T.v.d.B., L.J.G., A.M.d.H., T.W., J.J.M.Z., W.P.T.M.M., P.A.d.J., J.H.), Vascular Medicine (G.K., W.S.), and Geriatrics (E.v.V., H.L.K.), University Medical Center Utrecht, Utrecht University, the Netherlands; and Department of Applied Mathematics (J.M.W., E.v.V., H.L.K.), Technical Medical Centre, University of Twente, Enschede, the Netherlands.

Please address correspondence to J.W. Bartstra, MD, Department of Radiology, University Medical Center Utrecht, Postbox 85500, 3508 GA Utrecht, the Netherlands; e-mail: jbartstra@hotmail.com



Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8212>

pulsatility in the perforating arteries in patients with CSVD.<sup>14</sup> The burden of structural brain disease in PXE and the relation with structural and functional abnormalities in the brain vessels are unknown.

Arterial calcifications develop in the intimal and medial layers of the arterial wall. In the intima, they occur in atherosclerotic plaques and are a measure of the atherosclerotic involvement of the arterial bed.<sup>15</sup> In the medial layer, calcification of the internal elastic lamina (IEL) is shown to contribute to arterial stiffness.<sup>16,17</sup> Increased arterial stiffness reduces the physiologic Windkessel function of the arterial tree and allows increased pulsatility in end-organs like the brain.<sup>18</sup> Moreover in atherosclerotic plaques, arterial calcifications typically develop along the IEL and in the medial arterial wall in PXE.<sup>19</sup> We hypothesized that these carotid siphon calcifications might contribute to the high prevalence of cerebrovascular disease in PXE through increased arterial pulsatility. The aim of this study was to compare intracranial arterial flow pulsatility, brain volumes, infarctions, and WML between patients with PXE and controls. In addition, we investigated the association between carotid siphon calcification and arterial pulsatility, brain volumes, infarctions, and WML in patients with PXE.

## MATERIALS AND METHODS

### Study Design and Population

Patients were recruited from the Dutch National Expertise Center for PXE in the University Medical Center Utrecht. We conducted a patient-control study in 50 patients with PXE from the Dutch National Expertise Center for Pseudoxanthoma Elasticum (DNECP) who volunteered to participate and 40 healthy controls. All patients with PXE had a confirmed diagnosis of PXE. PXE was diagnosed if 2 of the following criteria were present: skin involvement (eg, yellow papules or plaques), eye involvement (eg, peau d'orange, angioid streaks), and genetic confirmation (biallelic *ABCC6* mutations).<sup>1</sup> Age- and sex-matched controls were recruited from the surrounding areas of patients with PXE. All patients with PXE and controls underwent MR imaging to assess brain structure and blood flow. Patients with PXE also underwent CT of the brain. Patients or controls were excluded if they were younger than 18 years of age, had an estimated glomerular filtration rate of  $<30$  mL/min/1.73m<sup>2</sup>, had a pacemaker or implantable cardioverter-defibrillator, had a metallic foreign body in the eye, or had claustrophobia. Controls were also excluded if they were first- or second-degree relatives of patients with PXE. The study was approved by the institutional review board of the University Medical Center Utrecht (No. 16–622/M-X), and written informed consent was obtained from all participants.

### Demographics

From all 50 patients and 40 controls, a medical history on medication and smoking status and a physical examination including height, weight, and blood pressure were obtained. Laboratory measurements included calcium, phosphate, estimated glomerular filtration rate, low-density lipoprotein (LDL) cholesterol, and 25-hydroxy vitamin D levels. Hypertension was defined as office systolic blood pressure of  $>140$  mm Hg and/or office diastolic blood pressure of  $>90$  mm Hg and/or the use of antihypertensive drugs. Hypercholesterolemia was defined as statin use and/or an LDL cholesterol level of  $>2.5$  mmol/L.

### CT Acquisition and Processing

Calcification was visualized on unenhanced, thin-section CT scans (120 kV[peak], mAs depending on body weight) on a Biograph 40 (Siemens) or a Brilliance CT 64 (Philips Healthcare) scanner. The carotid siphon calcification relative mass score was quantified using an in-house-developed software tool (iX Viewer; Image Sciences Institute) on the basis of the area and density of arterial wall lesions with an attenuation of  $>130$  HU (J.W.B.).<sup>20</sup> Twenty-five randomly selected scans were scored by 2 investigators. The intraclass correlation coefficient was assessed and was excellent (0.915; range, 0.751–0.966).

### MR Imaging Acquisition and Processing

Fifty patients with PXE and 42 healthy controls underwent a standardized MR imaging protocol (Ingenia CX 3T; Philips Healthcare) including time-resolved 2D phase-contrast, 3D T1-weighted, and 3D T2-weighted FLAIR acquisitions of the brain. In 2 controls, MR imaging was ended prematurely due to claustrophobia.

The 2D phase-contrast acquisitions were planned separately for both sides proximal and distal to the carotid siphon at the cavernous ICA segment (C4) and at the origin of the MCA M1 segment, respectively (Online Supplemental Data).<sup>21,22</sup> We used the following imaging parameters: FOV =  $250 \times 250$  mm<sup>2</sup>, acquired resolution =  $0.5 \times 0.5 \times 3$  mm<sup>3</sup> (reconstructed spatial resolution =  $0.25 \times 0.25 \times 3$  mm<sup>3</sup>), acquired temporal resolution = 64 ms, and unidirectional through-plane velocity-encoding sensitivities = 100 cm/s for C4 and MCA. The flow acquisitions provided time-resolved measurements of the blood flow velocity and volumetric flow rates over the cardiac cycle.

An in-house-developed Matlab script (MathWorks) was used to process the 2D phase-contrast acquisitions. The script created ROIs for each phase of the cardiac cycle based on image intensity and corrected potential phase wraps in the velocity maps. All ROIs were manually checked for ROI variation (L.J.G.). The acquisitions were included if they met the following previously described quality criteria:<sup>12</sup> The image section was planned perpendicular to the vessel (circular vessel appearance), the contour of the ROI followed the vessel contour adequately (circular ROI), and the contour was stable over the cardiac cycle with only minor ROI variations. From the included acquisitions, blood flow was calculated by integrating blood flow velocity over the vessel lumen. Next, the pulsatility index (PI) was calculated as peak systolic flow minus the lowest diastolic flow divided by the mean flow during 1 cardiac cycle.<sup>21</sup> The mean PI was calculated from the left and right carotid siphon and MCA. Of 360 pulsatility measurements, 35 were missing due to scan artifacts or suboptimal planning, and data were imputed using a single imputation. These measurements were previously shown to be reproducible among observers.<sup>21</sup>

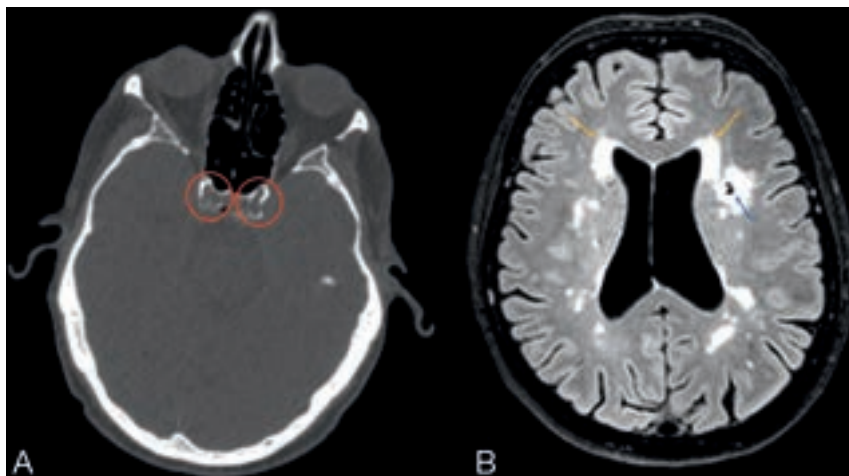
Brain volumes (GM, WM, CSF, and cortical thickness) and WML were computed automatically from the 3D T1 and 3D T2 FLAIR sequences using the Computational Anatomy Toolbox 12 (CAT12, <http://www.neuro.uni-jena.de/cat/>) and the Lesion Segmentation Tool (LST, <https://www.applied-statistics.de/lst.html>) within SPM12 software (<https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>). WML were quantified on the T2 FLAIR scans using LST, and the T1 scans were used to quantify the other brain volumes using CAT12. Before segmentation with CAT12, the WML on T1 scans

**Table 1: Clinical characteristics of 50 patients with PXE and 40 controls<sup>a</sup>**

Characteristics	PXE (n = 50)	Controls (n = 40)	P Value
Clinical			
Age (yr)	57 (SD, 12)	58 (SD, 11)	.49
Male sex (No.)	24 (48)	20 (50)	.85
BMI (kg/m <sup>2</sup> )	26 (SD, 4)	26 (SD, 4)	.74
Current smokers (No.)	6 (12)	5 (13)	.86
Systolic blood pressure (mm Hg)	136 (SD, 21)	133 (SD, 15)	.49
Diastolic blood pressure (mm Hg)	77 (SD, 12)	79 (SD, 10)	.41
Hypertension (No.)	22 (44)	18 (45)	.92
Hypercholesterolemia (No.)	43 (86)	37 (93)	.33
Medication			
Glucose-lowering medication (No.)	1 (2)	0 (0)	.38
Antihypertensive medication (No.)	13 (26)	7 (18)	.40
Cholesterol-lowering medication (No.)	26 (52)	5 (13)	<.01
Laboratory			
Calcium (mmol/L)	2.38 (SD, 0.09)	2.39 (SD, 0.09)	.61
Phosphate (mmol/L)	1.02 (SD, 0.16)	0.98 (SD, 0.16)	.25
Estimated glomerular filtration rate (mL/min/1.73m <sup>2</sup> )	90 (82–90)	88 (75–90)	.17
LDL-cholesterol (mmol/L)	2.8 (SD, 0.9)	3.6 (SD, 0.8)	<.01
25-OH vitamin D (nmol/L)	74 (SD, 36)	61 (SD, 20)	.04

**Note:**—BMI indicates body mass index; 25-OH vitamin D = 25-hydroxy vitamin D.

<sup>a</sup>Data are presented as mean (SD), median (interquartile range), or No. (%) when appropriate. Data were analyzed using the Student *t* test, Mann Whitney *U* test, or  $\chi^2$  test when appropriate. A *P* value < .05 was statistically significant.



**FIG 1.** Clinical example of vascular and brain disease in PXE. Carotid siphon calcification on CT (A, red circles) and small-vessel disease on 3T MR imaging (B). The blue arrow shows a lacunar infarction; the orange arrows show WML.

were filled on the basis of the results of the LST segmentation. All segmentations were visually checked (J.W.B.). Lacunar, subcortical, and cortical infarctions were scored on the basis of the Standards for Reporting Vascular changes on nEuroimaging recommendations<sup>23</sup> on the 3D T1 and 3D T2-FLAIR scans by an experienced, independent neuroradiologist blinded to the disease state, demographics, calcification mass, and pulsatility index (T.W.).

### Statistical Analysis

Descriptive data are presented as categoric (number and percentage), normally distributed continuous (mean [SD]), or non-normally distributed continuous variables (median [interquartile range]) as appropriate.

The difference in the PI, brain volumes, and small-vessel disease burden between patients with PXE and healthy controls was evaluated with the unpaired Student *t* test with 2-tailed probabilities when normally distributed or the Mann-Whitney *U* test when non-normally distributed. Separate analyses were performed for lacunar infarctions and all infarctions, in which lacunar, subcortical, and cortical infarctions were summed.

<sup>10</sup>Log transformation was performed on the carotid siphon calcification mass. Linear regression models were built to test the association between calcification mass and the PI and brain volumes and small-vessel disease. Negative binomial regression models were built to test the association between carotid siphon calcification and lacunar and all infarctions. The different models were adjusted for age and sex and for age, sex, and hypertension. A *P* value < .05 was statistically significant. Data analyses were performed in SPSS statistics, Version 25 (IBM); figures were made in R statistical and computing software, Version 3.3.2 (<http://www.r-project.org/>).

## RESULTS

### Baseline Characteristics

The baseline characteristics of the 50 patients with PXE and 40 healthy controls are presented in Table 1. The mean age of patients with PXE was 57 versus 58 years in controls. Forty-eight percent of patients with PXE versus 50% of controls were men. The PI measurement of 1 patient with PXE was excluded because of an ophthalmic artery aneurysm, which might affect blood flow. The brain volume measurements from 2 patients with PXE were

excluded from analysis due to demyelinating disease and a communicating hydrocephalus.

### Imaging Findings in Those with PXE versus Controls

We provide an illustration of the disease on CT and MR imaging in 1 of the patients with PXE in Fig 1. The median PI was significantly higher in patients with PXE than in controls (1.05; range, 0.94–1.21 versus 0.94; range, 0.82–1.04; *P* = .01) (Table 2). Patients with PXE had significantly lower mean GM volumes (597 [SD, 53] mL versus 632 [SD, 53] mL, *P* < .01) and more WML (2.6; range, 0.5–7.5 mL versus 1.1; range, 0.5–2.4; *P* = .05). Patients with PXE had more lacunar infarctions (total 64 versus 8, *P* = .04) and all types of infarctions (75 versus 14, *P* = .02).

## Arterial Calcification, Flow Pulsatility, and Small-Vessel Disease in PXE

Forty-eight of 50 (96%) patients with PXE had calcifications in the carotid siphon, and the median calcification mass score was 17 (range, 7–51). Calcification was associated with a higher PI ( $\beta = 0.10$ ; 95% CI, 0.01–0.18; Fig 2), higher CSF volume ( $\beta = 67.1$ ; 95% CI, 18.7–115.5), more WML ( $\beta = 6.2$ ; 95% CI, 1.6–10.9), and all infarctions (rate ratio [RR] = 2.8; 95% CI, 1.1–7.0) when adjusted for age, sex, and hypertension (Table 3). The PI was associated with more WML (Table 4).

## DISCUSSION

In this patient-control study, we show that patients with PXE have higher intracranial artery flow pulsatility, lower GM volumes, and

more lacunar infarctions and WML as a measure of CSVD, compared with an age- and sex-matched control population. In addition, we show that carotid siphon calcification in patients with PXE is associated with higher flow pulsatility of the intracranial arteries and that calcification is associated with reduced brain volumes and a higher small-vessel disease burden. These findings support the pulsatility hypothesis and show that arterial calcification is associated with brain damage in PXE (Fig 3). This pulsatility hypothesis states that increased pulsatility of the blood flow to the brain can damage brain tissue.<sup>18</sup> Large elastic arteries such as the aorta and carotid artery expand during systole and thereby dampen the pulsatile blood flow to the brain. This dampening results in laminar blood flow in the small vessels of the brain. Stiffening of the large arteries results in less dampening of the pulsatile blood flow and, therefore, increased pulsatility in the microvasculature, which might damage the brain parenchyma.

Arterial calcification is thought to contribute to arterial stiffening. These calcifications develop in the intimal and medial layers of the arterial wall.<sup>16</sup>

Intimal calcifications occur in atherosclerotic plaques and are suggested to contribute to plaque stabilization. IEL calcification is mainly thought to contribute to arterial stiffness and hypertension and might cause increased pulsatility in high-flow organs such as the kidney and the brain.<sup>18</sup> Carotid siphon calcification is very prevalent in the general population in which it co-occurs, with atherosclerotic intimal calcification.<sup>24</sup> It is, therefore, difficult to investigate the effect of IEL calcifications in cardiovascular disease in general populations. Arterial calcifications in PXE typically develop along the IEL in the medial arterial wall.<sup>19</sup> Although patients with PXE do develop atherosclerotic calcifications, these are not considered to be more than in the general population. PXE can, therefore, be seen as a model disease for the role of relatively isolated IEL calcifications in cardiovascular disease.

The findings of this study are in line with those in previous research in more specific and general populations. A study on patients with ischemic stroke showed that intracranial artery calcification is correlated with increased arterial pulsatility in the MCA as measured with ultrasound.<sup>25,26</sup> Intracranial artery calcifications are associated with small-vessel disease,<sup>27</sup> infarctions,<sup>28</sup> decreased brain volumes,<sup>27</sup> dementia, and cognitive decline in the general and aging

**Table 2: Imaging findings in patients with PXE compared with controls<sup>a</sup>**

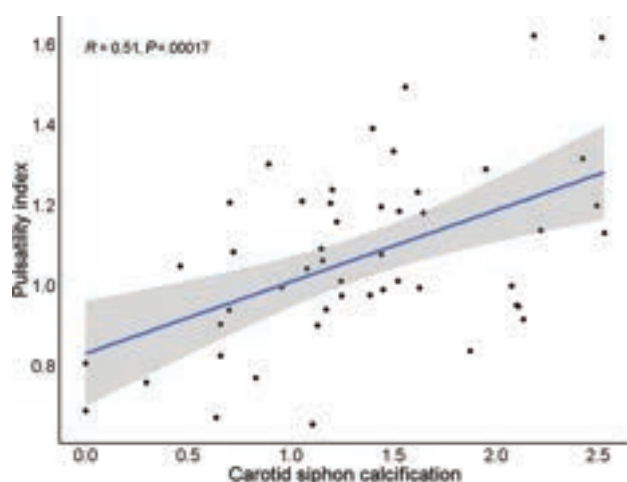
	Patients with PXE (n = 50)	Controls (n = 40)	P Value
Carotid siphon			
Calcification mass score	17 (7–51)	—	—
PI	1.05 (0.94–1.21)	0.94 (0.82–1.04)	.02 <sup>b</sup>
Brain MR imaging markers			
GM (mL)	597 (SD, 53)	632 (SD, 53)	<.01 <sup>c</sup>
WM (mL)	492 (SD, 52)	509 (SD, 56)	.15
CSF (mL)	337 (SD, 104)	311 (SD, 82)	.18
Cortical thickness (mm)	2.6 (SD, 0.1)	2.6 (SD, 0.1)	.16
WML (mL)	2.6 (0.5–7.5)	1.1 (0.5–2.4)	.05 <sup>b</sup>
Infarctions			
Lacunar infarctions			
No. of patients	16 (32)	6 (15)	.06
No. of infarctions	64	8	.04 <sup>b</sup>
All infarctions			
No. of patients	19 (38)	7 (18)	.03 <sup>b</sup>
No. of infarctions	75	14	.02 <sup>b</sup>

**Note:**—The en dash indicates that data on calcification mass are not available for controls.

<sup>a</sup> Data are presented as mean (SD), median (interquartile range), or No. (%) when appropriate. Data were analyzed using the Student *t* test, Mann-Whitney *U* test, or  $\chi^2$  test when appropriate. All infarctions include lacunar infarctions, 3 subcortical and 8 cortical infarctions in patients with PXE, and 6 cortical infarctions in controls.  $P < .05$  was statistically significant.

<sup>b</sup>  $P < .05$ .

<sup>c</sup>  $P < .01$ .



**FIG 2.** Correlation between carotid siphon calcification and the PI in patients with PXE. Carotid siphon calcification was correlated with the intracranial PI ( $R = 0.51$ ,  $P < .01$ , regression line; 95% CI, gray band). This association remained significant after multivariate adjustment for age and sex ( $\beta = 0.10$ ; 95% CI, 0.02–0.18\*) and age, sex, and hypertension ( $\beta = 0.10$ ; 95% CI, 0.01–0.18\*). Carotid siphon calcification was analyzed as  $^{10}\log(1 + \text{carotid siphon calcification})$ .  $P < .05$  was statistically significant. A single asterisk indicates  $P < .05$ .



**Table 3: Associations of carotid siphon calcification with brain MR imaging findings<sup>a</sup>**

	Carotid Siphon Calcification in PXE	
	Age- and Sex-Adjusted (n = 50)	Multivariable-Adjusted (n = 50)
MR imaging volumes ( $\beta$ [95% CI])		
GM (mL)	-17.3 [-39.9–5.2]	-16.8 [-39.9–6.4]
WM (mL)	-3.6 [-27.3–20.2]	-4.2 [-28.4–20.0]
CSF (mL)	74.7 [25.5–123.9] <sup>c</sup>	67.1 [18.7–115.5] <sup>c</sup>
Cortical thickness ( $\mu$ m)	-39.8 [-89.9–10.1]	-40.5 [-91.6–10.7]
WML (mL)	6.4 [1.9–11.0] <sup>c</sup>	6.2 [-1.6–10.9] <sup>b</sup>
Brain infarctions (RR [95%CI])		
Lacunar infarctions	2.5 [0.98–6.4]	2.3 [0.9–6.1]
All infarctions	3.1 [1.3–7.5] <sup>b</sup>	2.8 [1.1–7.0] <sup>b</sup>

<sup>a</sup> Carotid siphon calcifications are analyzed as  $10 \log 1 +$  carotid siphon calcification. For carotid siphon calcification, MR imaging volumes and infarctions are adjusted for age and sex (model 1) and age, sex, and hypertension (model 2).  $P < .05$  was statistically significant.

<sup>b</sup>  $P < .05$ .

<sup>c</sup>  $P < .01$ .

**Table 4: Associations of pulsatility with brain MR imaging findings<sup>a</sup>**

	Pulsatility in PXE	
	Age and Sex-Adjusted (n = 50)	Multivariable-Adjusted (n = 50)
MR imaging volumes ( $\beta$ [95% CI])		
GM (mL)	3.8 [-73.9–81.6]	7.4 [-71.4–86.1]
WM (mL)	22.7 [-55.8–101.3]	23.1 [-57.0–103.2]
CSF (mL)	68.6 [-115.5–252.7]	45.6 [-131.8–223.0]
Cortical thickness ( $\mu$ m)	29.5 [-145.0–204.1]	31.9 [-146.0–209.7]
WML (mL)	26.5 [11.6–41.3] <sup>c</sup>	25.9 [10.8–40.9] <sup>c</sup>
Brain infarctions (RR [95%CI])		
Lacunar infarctions	1.7 [0.1–23.6]	1.8 [0.1–25.0]
All infarctions	2.3 [0.2–1.2]	2.5 [0.2–33.6]

<sup>a</sup> For pulsatility, MR imaging volumes and infarctions are adjusted for age and sex (model 1) and age, sex, and hypertension (model 2).  $P < .05$  was statistically significant.

<sup>b</sup>  $P < .05$ .

<sup>c</sup>  $P < .01$ .

populations.<sup>27,29</sup> Increased pulsatility is associated with brain atrophy in the elderly and in patients along the dementia spectrum<sup>30,31</sup> and with small-vessel disease in patients with an acute stroke or TIA.<sup>32</sup> These findings suggest that arterial calcification and increased pulsatility might also play a role in brain damage in more general populations. Recently, we developed a histologic validated CT calcium score to differentiate between a dominant pattern of atherosclerotic intimal calcification and IEL calcification in the CS for use in the general population.<sup>33</sup> This score was used in 2 studies from the general population of the Rotterdam study. It showed that IEL calcification increases from 23% below 65 years of age to 79% above 80 years of age.<sup>34</sup> Furthermore, it was shown that in the general population, IEL calcification of the carotid siphon (CS) is the leading mechanism explaining the link between blood pressure and CSVD.<sup>35</sup> Second, IEL calcification was more strongly associated with stroke than intimal calcification, certainly in the elderly population.<sup>34</sup> Longitudinal studies that modify calcification or pulsatility are needed to provide further causal evidence.<sup>36</sup> We performed a patient-control study in a calcific disease, PXE, and our findings support the hypothesis, but we cannot claim causality on the basis of this cross-sectional study.

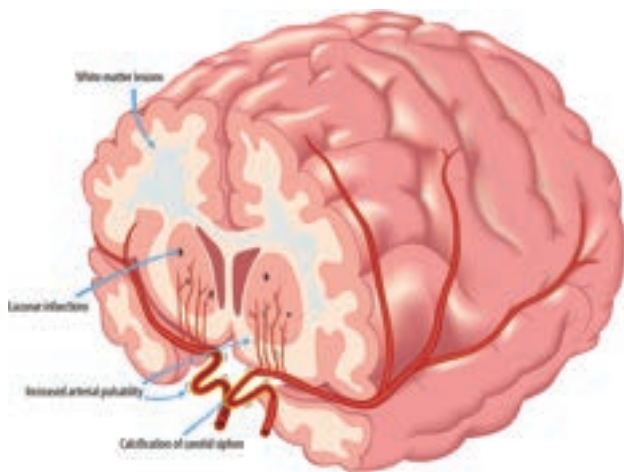
In PXE, low levels of inorganic pyrophosphate (PPi) are thought to underlie the increased propensity for arterial calcification.<sup>37</sup> PPi is a potent calcification inhibitor that also plays an important role in other genetic disorders like generalized arterial calcification of infancy (OMIM 208000) and arterial calcification due to a deficiency in CD73 (OMIM #211800).<sup>16</sup> In generalized arterial calcification of infancy, the complete lack of PPi results in extensive arterial calcification at birth, which is often lethal in the first half year of life.<sup>16</sup> Arterial calcification due to a deficiency in CD73 results in extensive calcification of the arteries and the joints of the hands and feet.<sup>16</sup> The low levels of PPi in PXE result in ectopic calcification of the skin, eyes, and arteries.<sup>1</sup> Recent studies for the treatment of PXE focus on increasing PPi with oral supplementation,<sup>38</sup> preventing PPi degradation by inhibiting tissue-nonspecific alkaline phosphatase,<sup>39</sup> or supplying the stable PPi analog etidronate.<sup>40,41</sup> Although all are promising studies, long-term longitudinal data on the effect of these therapies on arterial calcification progression, pulsatility, or clinical outcomes are currently lacking.

### Strengths and Limitations

A strength of this study is the extensive imaging. We used CT and functional and structural MR imaging to test every step from our hypothesis from large-artery carotid siphon calcification to microvascular brain damage.

Because we did not obtain CT scans of the healthy controls, the difference in carotid siphon calcifications could not be assessed. However, it was previously shown that the prevalence and severity of carotid siphon calcification are higher in patients with PXE than in hospital controls.<sup>2</sup> In this cohort, patients with PXE were more often on statin treatment than the healthy controls. Because of their increased cardiovascular risk, patients with PXE in our clinical practice are started on statin treatment early to control additional cardiovascular risk factors. The healthy controls had significantly higher LDL cholesterol levels. To compare the prevalence of hypercholesterolemia between the groups, we, therefore, made a composite hypercholesterolemia score based on statin use and an LDL cholesterol level of  $>2.5$  mmol/L; the score showed no differences between the groups. In addition, patients with PXE had significantly lower vitamin D levels. Adjustment of the regression models for LDL cholesterol, statin treatment, and vitamin D levels did not significantly change the outcomes of the study. Finally, we performed a cross-sectional study that supports the pulsatility hypothesis; however, causal





**FIG 3.** Arterial calcification, pulsatility, and microvascular brain damage in PXE. Carotid siphon calcification as a measure of intracranial arterial calcification is associated with increased flow pulsatility in PXE. The subsequent pulse pressure-induced microvascular damage, including lacunar infarctions and WML, might contribute to cognitive decline.

claims can only be proved in longitudinal studies, preferably with an intervention that modifies the pulsatility. Also, the calcification-induced change in arterial pulsatility occurs locally and is transmitted more by downward flow toward the cerebral tissue. Thus, in this later trajectory, the pulsatility can be modulated again. Thus, another way to more certainly assess whether pulsatility reaches the cerebral tissue is by measuring the pulsatility in perforating arteries.

In our controls, we observed lacunar infarctions in 6 (15%), which appears somewhat higher than those observed in the Rotterdam study (7.6%). This finding could be explained by the sensitivity of the method (ie, we used 3T instead of 1.5T) or by the small sample size of our control group.<sup>42</sup> Due to the cross-sectional design, caution is needed with causal assumptions. Larger and longitudinal studies in other populations must show whether our results are generalizable and whether progression of carotid siphon calcification results in progression of microvascular brain damage in patients with PXE and other populations.

## CONCLUSIONS

We showed that patients with PXE have more structural brain disease and increased intracranial arterial flow pulsatility compared with healthy controls. Arterial calcification of the carotid siphon is associated with increased flow pulsatility and microvascular brain damage, including infarctions and WML in patients with PXE. Larger, longitudinal studies are needed to further elucidate the causal role of arterial calcification and pulsatility in PXE-related brain disease.

**Disclosure forms** provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

- Plomp AS, Toonstra J, Bergen AA, et al. **Proposal for updating the pseudoxanthoma elasticum classification system and a review of the clinical findings.** *Am J Med Genet A* 2010;152A:1049–58 CrossRef Medline
- Kranenburg G, de Jong PA, Mali WP, et al. **Prevalence and severity of arterial calcifications in pseudoxanthoma elasticum (PXE) compared to hospital controls: novel insights into the vascular phenotype of PXE.** *Atherosclerosis* 2017;256:7–14 CrossRef Medline
- Risseuw S, Ossewaarde-van Norel J, van Buchem C, et al. **The extent of angioid streaks correlates with macular degeneration in pseudoxanthoma elasticum.** *Am J Ophthalmol* 2020;220:82–90 CrossRef Medline
- Leftheriotis G, Kauffenstein G, Hamel JF, et al. **The contribution of arterial calcification to peripheral arterial disease in pseudoxanthoma elasticum.** *PLoS One* 2014;9:e96003 CrossRef Medline
- Kauw F, Kranenburg G, Kappelle LJ, et al. **Cerebral disease in a nationwide Dutch pseudoxanthoma elasticum cohort with a systematic review of the literature.** *J Neurol Sci* 2017;373:167–72 CrossRef Medline
- Cerrato P, Giraudo M, Baima C, et al. **Asymptomatic white matter ischemic lesions in a patient with pseudoxanthoma elasticum.** *J Neurol* 2005;252:848–49 CrossRef Medline
- Pavlovic AM, Zidverc-Trajkovic J, Milovic MM, et al. **Cerebral small vessel disease in pseudoxanthoma elasticum: three cases.** *Can J Neurol Sci* 2005;32:115–18 CrossRef Medline
- Manini A, Pantoni L, Pantoni L, et al. **Genetic causes of cerebral small vessel diseases: a practical guide for neurologists.** *Neurology* 2022;100:766–83 CrossRef Medline
- Omarjee L, Fortrat JO, Larralde A, et al. **Internal carotid artery hypoplasia: a new clinical feature in pseudoxanthoma elasticum.** *J Stroke* 2019;21:108–11 CrossRef Medline
- Mention PJ, Lacoëuille F, Leftheriotis G, et al. **18F-fluorodeoxyglucose and 18F-sodium fluoride positron emission tomography/computed tomography imaging of arterial and cutaneous alterations in pseudoxanthoma elasticum.** *Circ Cardiovasc Imaging* 2018;11:e007060 CrossRef Medline
- Bartstra JW, Spiering W, van den Ouweland JM, et al. **Increased elastin degradation in pseudoxanthoma elasticum is associated with peripheral arterial disease independent of calcification.** *J Clin Med* 2020;9:2771–11 CrossRef Medline
- van Tuijl RJ, Ruigrok YM, Velthuis BK, et al. **Velocity pulsatility and arterial distensibility along the internal carotid artery.** *J Am Heart Assoc* 2020;9:e016883 CrossRef Medline
- Bartstra JW, van Tuijl RJ, de Jong PA, et al. **Pulsatility attenuation along the carotid siphon in pseudoxanthoma elasticum.** *AJNR Am J Neuroradiol* 2021;42:2030–33 CrossRef Medline
- van Tuijl RJ, Ruigrok YM, Geurts LJ, et al. **Does the internal carotid artery attenuate blood-flow pulsatility in small vessel disease? A 7 T 4D-flow MRI study.** *J Magn Reson Imaging* 2022;56:527–35 CrossRef Medline
- Panh L, Lairez O, Ruidavets JB, et al. **Coronary artery calcification: from crystal to plaque rupture.** *Arch Cardiovasc Dis* 2017;110:550–61 CrossRef Medline
- Lanzer P, Boehm M, Sorribas V, et al. **Medial vascular calcification revisited: review and perspectives.** *Eur Heart J* 2014;35:1515–25 CrossRef Medline
- Chirinos JA, Segers P, Hughes T, et al. **Large-artery stiffness in health and disease: JACC State-of-the-Art Review.** *J Am Coll Cardiol* 2019;74:1237–63 CrossRef Medline
- Mitchell GF. **Aortic stiffness, pressure and flow pulsatility, and target organ damage.** *J Appl Physiol* (1985) 2018;125:1871–80 CrossRef Medline
- Vos A, Kranenburg G, de Jong PA, et al. **The amount of calcifications in pseudoxanthoma elasticum patients is underestimated in computed tomographic imaging; a post-mortem correlation of histological and computed tomographic findings in two cases.** *Insights Imaging* 2018;9:493–98 CrossRef Medline
- Hoffmann U, Kwait DC, Handwerker J, et al. **Vascular calcification in ex vivo carotid specimens: precision and accuracy of measurements**

- with multi-detector row CT. *Radiology* 2003;229:375–81 CrossRef Medline
21. Correia de Verdier M, Wikstrom J. **Normal ranges and test-retest reproducibility of flow and velocity parameters in intracranial arteries measured with phase-contrast magnetic resonance imaging.** *Neuroradiology* 2016;58:521–31 CrossRef Medline
  22. Buijs PC, Krabbe-Hartkamp MJ, Bakker CJ, et al. **Effect of age on cerebral blood flow: measurement with ungated two-dimensional phase-contrast MR angiography in 250 adults.** *Radiology* 1998;209:667–74 CrossRef Medline
  23. Wardlaw JM, Smith EE, Biessels GJ, et al; STandards for ReportIng Vascular changes on nEuroimaging (STRIVE v1). **Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration.** *Lancet Neurol* 2013;12:822–38 CrossRef Medline
  24. Bartstra JW, van den Beukel TC, Van Hecke W, et al. **Intracranial arterial calcification: prevalence, risk factors, and consequences: JACC Review Topic of the Week.** *J Am Coll Cardiol* 2020;76:1595–604 CrossRef Medline
  25. Wu X, Wang L, Zhong J, et al. **Impact of intracranial artery calcification on cerebral hemodynamic changes.** *Neuroradiology* 2018;60:357–63 CrossRef Medline
  26. Park KY, Chung PW, Kim YB, et al. **Increased pulsatility index is associated with intracranial arterial calcification.** *Eur Neurol* 2013;69:83–88 CrossRef Medline
  27. Bos D, Vernooij MW, Elias-Smale SE, et al. **Atherosclerotic calcification relates to cognitive function and to brain changes on magnetic resonance imaging.** *Alzheimers Dement* 2012;8:S104–111 CrossRef Medline
  28. Hong NR, Seo HS, Lee YH, et al. **The correlation between carotid siphon calcification and lacunar infarction.** *Neuroradiology* 2011; 53:643–49 CrossRef Medline
  29. Bos D, Vernooij MW, de Bruijn RF, et al. **Atherosclerotic calcification is related to a higher risk of dementia and cognitive decline.** *Alzheimers Dement* 2015;11:639–47.e1 CrossRef Medline
  30. Berman SE, Rivera-Rivera LA, Clark LR, et al. **Intracranial arterial 4D-flow is associated with metrics of brain health and Alzheimer's disease.** *Alzheimers Dement (Amst)* 2015;1:420–28 CrossRef Medline
  31. Wahlin A, Ambarki K, Birgander R, et al. **Intracranial pulsatility is associated with regional brain volume in elderly individuals.** *Neurobiol Aging* 2014;35:365–72 CrossRef Medline
  32. Birnefeld J, Wahlin A, Eklund A, et al. **Cerebral arterial pulsatility is associated with features of small vessel disease in patients with acute stroke and TIA: a 4D flow MRI study.** *J Neurol* 2020;267:721–30 CrossRef Medline
  33. Kockelkoren R, Vos A, Hecke WV, et al. **Computed tomographic distinction of intimal and medial calcification in the intracranial internal carotid artery.** *PLoS One* 2017;12:e0168360–11 CrossRef
  34. Van Den Beukel TC, Van Der Toorn JE, Vernooij MW, et al. **Morphological subtypes of intracranial internal carotid artery arteriosclerosis and the risk of stroke.** *Stroke* 2022;53:1339–47 CrossRef Medline
  35. Melgarejo JD, Vernooij MW, Ikram MA, et al. **Intracranial carotid arteriosclerosis mediates the association between blood pressure and cerebral small vessel disease.** *Hypertension* 2023;80:618–28 CrossRef Medline
  36. Tsao CW, Pencina KM, Massaro JM, et al. **Cross-sectional relations of arterial stiffness, pressure pulsatility, wave reflection, and arterial calcification.** *Arterioscler Thromb Vasc Biol* 2014;34:2495–500 CrossRef Medline
  37. Jansen RS, Duijst S, Mahakena S, et al. **ABCC6-mediated ATP secretion by the liver is the main source of the mineralization inhibitor inorganic pyrophosphate in the systemic circulation: brief report.** *Arterioscler Thromb Vasc Biol* 2014;34:1985–89 CrossRef Medline
  38. Dedinszki D, Szeri F, Kozak E, et al. **Oral administration of pyrophosphate inhibits connective tissue calcification.** *EMBO Mol Med* 2017;9:1463–70 CrossRef Medline
  39. Ziegler SG, Ferreira CR, MacFarlane EG, et al. **Ectopic calcification in pseudoxanthoma elasticum responds to inhibition of tissue-nonspecific alkaline phosphatase.** *Sci Transl Med* 2017;9:eaal1669 CrossRef Medline
  40. Bartstra JW, de Jong PA, Kranenburg G, et al. **Etidronate halts systemic arterial calcification in pseudoxanthoma elasticum.** *Atherosclerosis* 2020;292:37–41 CrossRef Medline
  41. Kranenburg G, de Jong PA, Bartstra JW, et al. **Etidronate for prevention of ectopic mineralization in patients with pseudoxanthoma elasticum.** *J Am Coll Cardiol* 2018;71:1117–26 CrossRef Medline
  42. Vernooij MW, Ikram MA, Tanghe HL, et al. **Incidental findings on brain MRI in the general population.** *N Engl J Med* 2007;357:1821–28 CrossRef Medline

# Hemorrhage Volume Drives Early Brain Injury and Outcome in Poor-Grade Aneurysmal SAH

<sup>15</sup>Pietro Panni, Franco Simionato, Roberta Cao, <sup>15</sup>Alessandro Pedicelli, <sup>15</sup>Enrico Marchese, Anselmo Caricato, <sup>15</sup>Andrea Alexandre, <sup>15</sup>Alberto Feletti, Mattia Testa, Paolo Zanatta, Nicola Gitti, <sup>15</sup>Simone Piva, Dikran Mardighian, <sup>15</sup>Vittorio Semeraro, Giordano Nardin, <sup>15</sup>Emilio Lozupone, Giafranco Paiano, <sup>15</sup>Edoardo Picetti, Vito Montanaro, Massimo Petranca, Carlo Bortolotti, <sup>15</sup>Antonino Scibilia, <sup>15</sup>Luigi Cirillo, <sup>15</sup>Raffaele Aspide, Andrea Luigi Lanterna, <sup>15</sup>Alessandro Ambrosi, Pietro Mortini, Maria Luisa Azzolini, Maria Rosa Calvi, and Andrea Falini, on behalf of the POGASH Investigators



## ABSTRACT

**BACKGROUND AND PURPOSE:** Early brain injury is a major determinant of clinical outcome in poor-grade (World Federation of Neurosurgical Societies [WFNS] IV–V) aneurysmal SAH and is radiologically defined by global cerebral edema. Little is known, though, about the effect of global intracranial hemorrhage volume on early brain injury development and clinical outcome.

**MATERIALS AND METHODS:** Data from the multicentric prospective Poor-Grade Aneurysmal Subarachnoid Hemorrhage (POGASH) Registry of consecutive patients with poor-grade aneurysmal SAH admitted from January 1, 2015, to August 31, 2022, was retrospectively evaluated. Poor grade was defined according to the worst-pretreatment WFNS grade. Global intracranial hemorrhage volume as well as the volumes of intracerebral hemorrhage, intraventricular hemorrhage, and SAH were calculated by means of analytic software in a semiautomated setting. Outcomes included severe global cerebral edema (defined by Subarachnoid Hemorrhage Early Brain Edema Score grades 3–4), in-hospital mortality (mRS 6), and functional independence (mRS 0–2) at follow-up.

**RESULTS:** Among 400 patients (median global intracranial hemorrhage volume of 91 mL; interquartile range, 59–128), severe global cerebral edema was detected in 218/400 (54.5%) patients. One hundred twenty-three (30.8%) patients died during the acute phase of hospitalization. One hundred fifty-five (38.8%) patients achieved mRS 0–2 at a median of 13 (interquartile range, 3–26) months of follow-up. Multivariable analyses showed global intracranial hemorrhage volume as independently associated with severe global cerebral edema (adjusted OR, 1.009; 95% CI, 1.004–1.014;  $P < .001$ ), mortality (adjusted OR, 1.006; 95% CI, 1.001–1.01;  $P = .018$ ) and worse clinical outcome (adjusted OR, 0.992; 95% CI, 0.98–0.996;  $P < .010$ ). The effect of global intracranial hemorrhage volume on clinical-radiologic outcomes changed significantly according to different age groups (younger than 50, 50–70, older than 70 year of age). Volumes of intracerebral hemorrhage, intraventricular hemorrhage, and SAH affected the 3 predefined outcomes differently. Intracerebral hemorrhage volume independently predicted global cerebral edema and long-term outcome, intraventricular hemorrhage volume predicted mortality and long-term outcome, and SAH volume predicted long-term clinical outcome.

**CONCLUSIONS:** Global intracranial hemorrhage volume plays a pivotal role in global cerebral edema development and emerged as an independent predictor of both mortality and long-term clinical outcome. Aging emerged as a reducing predictor in the relationship between global intracranial hemorrhage volume and global cerebral edema.

**ABBREVIATIONS:** aSAH = aneurysmal SAH; aOR = adjusted OR; DCI = delayed cerebral ischemia; EBI = early brain injury; EVD = external ventricular drain; GCE = global cerebral edema; GHV = global intracranial hemorrhage volume; ICH = intracerebral hemorrhage; ICP = intracranial pressure; IQR = interquartile range; IVH = intraventricular hemorrhage; LOC = loss of consciousness; POGASH = Poor-Grade Aneurysmal Subarachnoid Hemorrhage; SEBES = Subarachnoid Hemorrhage Early Brain Edema Score; WFNS = World Federation of Neurosurgical Societies; V = volume

**A**neurysmal rupture, microcirculation compromise, and elevated intracranial pressure (ICP) contribute to the development of


early brain injury (EBI), which plays a pivotal role in determining disability and mortality in patients with aneurysmal SAH (aSAH).<sup>1,2</sup>

Received July 17, 2023; accepted after revision December 6.

From the Department of Neuroradiology (P.P., F.S., R.C., A. Falini), Division of Interventional Neuroradiology, and Department of Neurosurgery (P.P., P.M.), IRCCS San Raffaele Scientific Institute, Vita-Salute San Raffaele University, Milan, Italy; Institute of Radiological Sciences (A.P., A. Alexandre), Department of Neurosurgery (E.M.), and Department of Anesthesia and Critical Care Medicine (A.C.), Fondazione Policlinico Universitario A. Gemelli IRCCS Catholic University of Rome, Rome, Italy; Institute of Neurosurgery (A. Feletti, M.T.), Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Verona, Italy; Anesthesia and Intensive Care A (P.Z.), Integrated University Hospital, Verona, Italy; Departments of Anesthesia, Critical Care and Emergency (N.G., S.P.), and Neuroradiology (D.M.), Spedali Civili University Hospital, Brescia, Italy; Departments of Radiology (V.S.), and Critical Care (G.N.), SS Annunziata Hospital, Taranto, Italy; Departments of Neuroradiology (E.L.), and Anaesthesia and Critical Care (G.P.), Vito-Fazzi Hospital, Lecce, Italy; Department of

Anesthesia and Intensive Care (E.P., V.M., M.P.), Parma University Hospital, Parma, Italy; Departments of Neurosurgery (C.B., A.S.), Neuroradiology (L.C.), and Neurointensive Care (R.A.), IRCCS Institute of Neurological Sciences “Bellaria,” Bologna, Italy; Department of Neurosurgery (A.L.L.), Papa Giovanni XXIII Hospital, Bergamo, Italy; and Biostatistics, School of Medicine (A. Ambrosi) and Department of Neurocritical Care (M.L.A., M.R.C.), San Raffaele University Hospital, Milan, Italy.

Please address correspondence to Pietro Panni, MD, Department of Neuroradiology, Interventional Neuroradiology Division, and Department of Neurosurgery, IRCCS San Raffaele University Hospital, Via Olgettina 60, 20132 Milan, Italy; e-mail: panni.pietro@hsr.it; @PietroPanni

 Indicates article with online supplemental data.  
<http://dx.doi.org/10.3174/ajnr.A8135>

EBI, which translates clinically in the severe neurologic deficits (World Federation of Neurosurgical Societies [WFNS] IV–V) on admission as well as in loss of consciousness (LOC) for the effect of transient cerebral ischemia,<sup>3</sup> has emerged as a leading cause of mortality in the past decades, outnumbering the proportion of patients dying from either rebleeding or delayed cerebral ischemia (DCI) related brain swelling.<sup>4,5</sup> Poor aSAH grades and ictal LOC have been shown to be clinical predictors of global cerebral edema (GCE) which is, in turn, the only radiologic marker of EBI, evaluated on a 4-grade scale, Subarachnoid Hemorrhage Early Brain Edema Score (SEBES).<sup>6,7</sup> The modified Fisher grade, a commonly used and accepted scale grading the amount of intracranial bleeding, has not been shown to be predictive of GCE.<sup>6</sup> In this study we sought to evaluate the effects of the primary hemorrhage volume, calculated on the first admission CT examination after aneurysmal rupture, on EBI and clinical outcome in a national prospective multicentric registry on patients with poor-grade aSAH.

## MATERIALS AND METHODS

### Data Availability

All data and materials can be accessed on reasonable request addressed to the Principal Investigator of the Registry, Pietro Panni, MD.

### Standard Protocol Approvals, Registrations, and Patient Consents

The local ethics committees approved both the inclusion and use of patient data in the Poor-Grade Aneurysmal Subarachnoid Hemorrhage (POGASH) Registry (NCT04945603). Patients' data were managed according to the Declaration of Helsinki. Informed consent for the scientific use of clinical-radiologic anonymized data was signed by patients' relatives or proxies.

### Study Population and Variable Description

POGASH (NCT04945603) is an ongoing prospective multicentric registry pooling anonymized data from prospectively collected and maintained institutional databases at 9 academic institutions/tertiary referral centers on a national level. All consecutive poor grade aSAH patients admitted to the emergency department of each participating center from June 2016 to August 2022 to provide at least 6 months of follow-up for patients surviving the acute phase of hospitalization were included for the present study.

Inclusion criteria for the present study were as follows:

- Patients consecutively admitted to the emergency department because of WFNS IV–V aSAH, regardless from treatment indication/abstention
- Available clinical follow-up
- Available anonymized DICOM files of the admission CT scan.

Exclusion criteria were as follows:

- Patients younger than 18 years of age
- aSAH due to trauma or vascular malformations other than cerebral aneurysms.

A detailed description of the entire variable list is provided as Online Supplemental Data.

Variables were grouped according to baseline and demographic features as well as the 4 different phases of the clinical course of the disease: EBI, intensive care unit, hospitalization, follow-up. Considering the EBI phase, poor grade was defined according to the worst pretreatment clinical grade, ie, at nadir.<sup>8,9</sup> LOC was recorded only in cases of at least >60-minute length, for this has been shown to be peculiar to early brain injury in the most severe forms of aSAH, regardless of subsequent recovery or direct intubation.<sup>3</sup> Time-metrics variables (time to external ventricular drain [EVD] placement and time to treatment of the culprit aneurysm) were derived by directly examining the documents of the ambulance service. Time of the ambulance call was considered as a proxy for aSAH onset. DCI was defined according to a previously reported definition.<sup>10</sup> Clinical outcome was scored according to the mRS assessed by qualified personnel in each center according to a predefined follow-up schedule.

### Neuroradiologic Evaluation of the Admission CT Scan

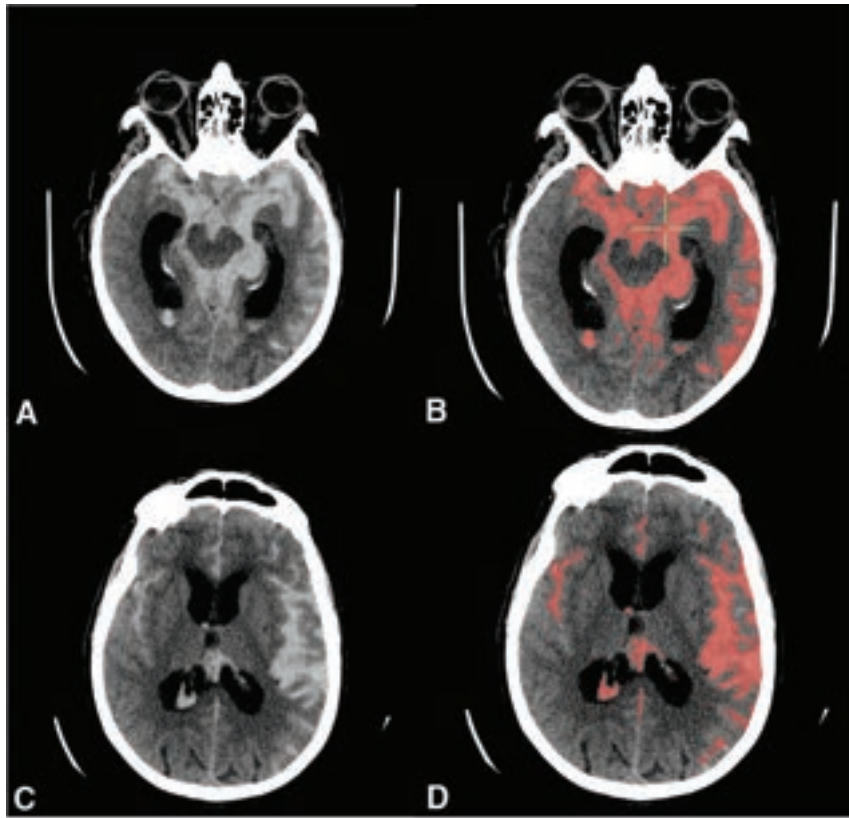
CT axial slices obtained for the selected patients on admission were retrieved from each institution's digital archive system, stored as DICOM files, and sent, anonymized, to the promoting center core imaging lab (Neuroradiology Department, San Raffaele University Hospital in Milan) for a re-evaluation blinded to the results of treatment or clinical follow-up. Each CT scan was evaluated for the presence of acute hydrocephalus, modified Fisher grade,<sup>11</sup> the presence of acute subdural hematoma, intraventricular hemorrhage (IVH), intracerebral hemorrhage (ICH), and GCE.

The presence of GCE was scored according to the presence, at the first admission CT scan, of complete or near-complete effacement of the hemispheric sulci and basal cisterns and/or bilateral and extensive disruption of the hemispheric gray-white matter junction.<sup>6</sup> This evaluation has been performed on 2 admission CT slices, the first one at the level of basal ganglia/subinsular region and the second, more cranial, at the level of the centrum semiovale.<sup>7</sup> GCE was dichotomized and considered to be present for SEBES grades 3–4 (clear and extensive bihemispheric involvement).

### Volumetric Analysis of Global Intracranial Hemorrhage Volume

The volumetric analysis method has been previously described.<sup>12</sup> Software-based semiautomatic segmentation and volumetric quantification were performed using OsiriX MD (<https://www.osirix-viewer.com/>). The determination of blood volume according to the ROI methodology was based on a tissue-specific threshold interval between 45 and 90 HU, to avoid bone-related artifacts in anatomically challenging regions like the skull base and the parenchyma-skull interface at the convexity and the manual outline of areas of blood on each CT section. The pixels with a similar attenuation in the neighboring areas were automatically connected by the software. Hemorrhage volumes were calculated in milliliters obtained by multiplying section thickness by the hemorrhage area in cases of volumetric scans. For nonvolumetric scans (eg, posterior fossa imagined with 2.5-mm slices and supratentorial space with 5-mm slices), 2 separate volumes were derived, one for posterior fossa and one for supratentorial space, and their sum was calculated. An example of global hemorrhage





**FIG 1.** Example of GHV semiautomatic segmentation with the neighboring method. The threshold was set between 45 and 90 HU. A and C, Basal hyperdensities. B and D, The results of the semiautomatic segmentation of hemorrhage with a threshold set between 45 and 90 HU.

volume (GHV) segmentation is provided in Fig 1. The same process was applied to derive volumes for GHV distributions, namely the volume of ICH (ICH-V), volume of IVH (IVH-V) and volume of SAH (SAH-V).

### Statistical Analysis

Quantitative variables are expressed as means (SD), and categorical variables are expressed as numbers (percentage). To evaluate interobserver agreement for GHV, 2 independent neuroradiologists, with >10 years of experience, blinded to treatment choice, clinical evolution, and outcome, evaluated a random sample made up of the first 80 consecutive CT scans included in the Registry (20% of the studied population), chosen regardless of center provenience, and calculated the GHV. A Bland-Altman graph was constructed and inspected to assess value distribution, and it is reported as Online Supplemental Data. Age stratification was based on its median value and interquartile range (IQR) distribution in the studied population. The aim of the present study was to ascertain the role of GHV on the development of EBI (SEBES grades 3–4 GCE as a radiologic proxy for EBI) and clinical outcome: in-hospital mortality (mRS 6 vs 0–5) and long-term clinical outcome (mRS 0–2 vs 3–6 at follow-up). Bivariate comparison and multivariable binary logistic regression analyses were performed for the 3 predefined outcomes, determining unadjusted and adjusted ORs (aORs) along with corresponding 95% CIs and *P* values for significance.

All variables emerging from bivariate comparison with *P* < .2 were included in the multivariable models. The results of

multivariable modeling were expressed after adjustment for covariates and admission center, to eliminate potential biases due to center-related treatment variations. Bivariate comparison of baseline variables between the studied population and excluded patients was performed as a sensitivity analysis and is provided as Online Supplemental Data. To rule out the potentially confounding effect of hydrocephalus, it was forced into the 3 multivariable models. Bivariate and multivariable analyses for independent predictors of GCE were repeated in the subpopulation younger than 70 years of age, to rule out the potentially confounding effect of brain atrophy (Online Supplemental Data.). Scatterplots were constructed on the basis of multivariable-derived predicted probabilities of predefined outcomes. The Hosmer-Lemeshow test was performed to verify the appropriateness of each analysis. Collinearity among variables entered in the multivariable models was assessed by the variance inflation factor, with multicollinearity considered present for variance inflation factor values of any variable of at least 10. All statistical analyses were

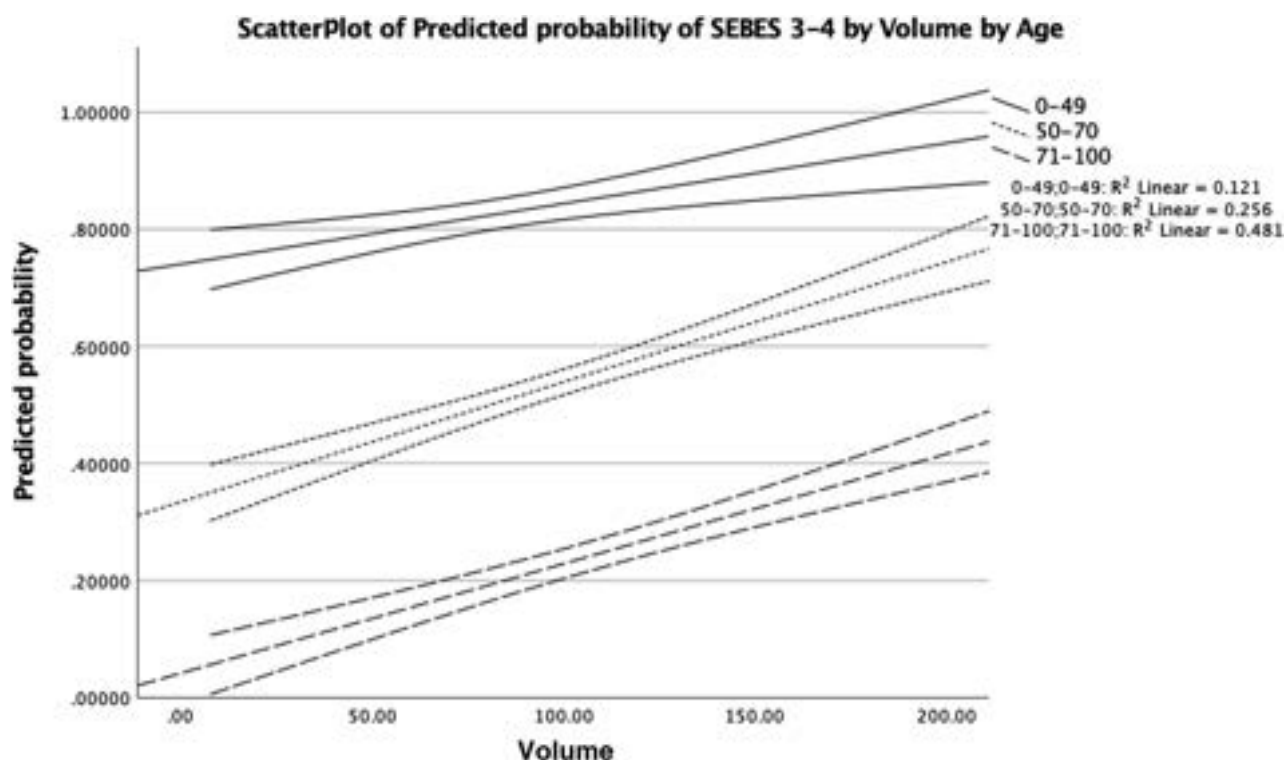
performed using SPSS, Version 20.0 (IBM) and in R environment (Version 4.1.3; <http://www.r-project.org/>).

### RESULTS

Of 466 consecutive patients with poor-grade aSAH with available follow-up included in the Registry up to August 2022, data concerning 400 patients with available admission CT for volumetric quantification of GHV were included (flow chart included as Online Supplemental Data). Baseline characteristics and treatment-related hospitalization and follow-up of the population are reported in the Online Supplemental Data. LOC occurred more frequently in the excluded patients (*P* = .017).

Considering clinical-radiologic features of EBI, 257/400 (64.3%) patients were admitted in coma (WFNS V) and experienced prolonged LOC (74.7%). SEBES grades 3 and 4 were detected in 218/400 (54.5%) patients. The median GHV was 91 mL (IQR, 59–128 mL), median ICH-V was 14 (IQR, 4.5–32.5), median IVH-V was 12.5 (IQR, 4–26), and the median SAH-V was 73.6 (44–29). One hundred ten patients were treated surgically (27.5%), while 258 (64.5%) were treated endovascularly. Treatment abstinence was chosen in 32 (8%) patients. Aneurysms were secured at a median of 5 hours (IQR 4–9 hours) since onset. One hundred twenty-three (30.8%) patients died during hospitalization. At a median follow-up of 24 months (IQR, 9–43 months), survivors had severe disabilities (median mRS, 4; IQR, 1–6), even though a non-negligible proportion achieved functional independency (mRS 0–2; 155/400, 38.8%). The Online Supplemental Data report the





**FIG 2.** Scatterplot showing multivariable analysis–derived predicted probabilities of severe SEBES (SEBES grades 3–4) on the admission CT scan according to age (3 groups: younger than 50, 50–70, older than 70 years of age) and baseline GHV.

results of univariable analyses for factors associated with EBI development, in-hospital mortality, and long-term clinical outcome, respectively. The Online Supplemental Data show the results of multivariable analyses for independent predictors of severe GCE (SEBES 3–4) presence at the admission CT, in-hospital mortality, and long-term clinical outcome.

Worst pretreatment clinical grading (WFNS V at nadir), (aOR, 1.9; 95% CI, 1.2–3.2;  $P = .002$ ) and GHV (aOR for every milliliter, 1.007; 95% CI, 1.002–1.012;  $P = .007$ ) independently increased the chance of severe SEBES on the admission CT scan, while older age (aOR for every year, 0.90; 95% CI, 0.88–0.93;  $P < .001$ ) independently reduced its chances. Considering in-hospital mortality, older age (aOR, 1.02; 95% CI, 1.00–1.04;  $P = .014$ ), WFNS V (aOR, 3.7; 95% CI, 2–6.7;  $P < .001$ ), rebleeding (aOR, 2.5; 95% CI, 1.3–4.5;  $P = .003$ ), and GHV (aOR for every milliliter, 1.006; 95% CI, 1.001–1.01;  $P = .018$ ) independently predicted in-hospital mortality, while EVD use (aOR, 0.48; 95% CI, 0.27–0.85;  $P = .012$ ) was independently associated with reduced mortality.

The independent predictors were retained by the multivariable analysis performed in the population younger than 70 years of age (Online Supplemental Data). Age (aOR for every year, 0.93; 95% CI, 0.91–0.96;  $P < .001$ ), GHV (aOR for every milliliter, 0.992; 95% CI, 0.98–0.996;  $P < .010$ ), cardiac arrest on admission (aOR, 0.12; 95% CI, 0.037–0.41;  $P < .001$ ), WFNS V (aOR, 0.51; 95% CI, 0.28–0.9;  $P = .029$ ), rebleeding (aOR, 0.42; 95% CI, 0.20–0.85;  $P = .017$ ), the need for decompressive craniectomy (aOR, 0.18; 95% CI, 0.08–0.40;  $P < .001$ ), and DCI (aOR, 0.24; 95% CI, 0.11–0.51;  $P < .001$ ) independently predicted increased disability, while EVD use (aOR, 3.4; 95% CI, 1.6–7.4;  $P = .002$ ) was

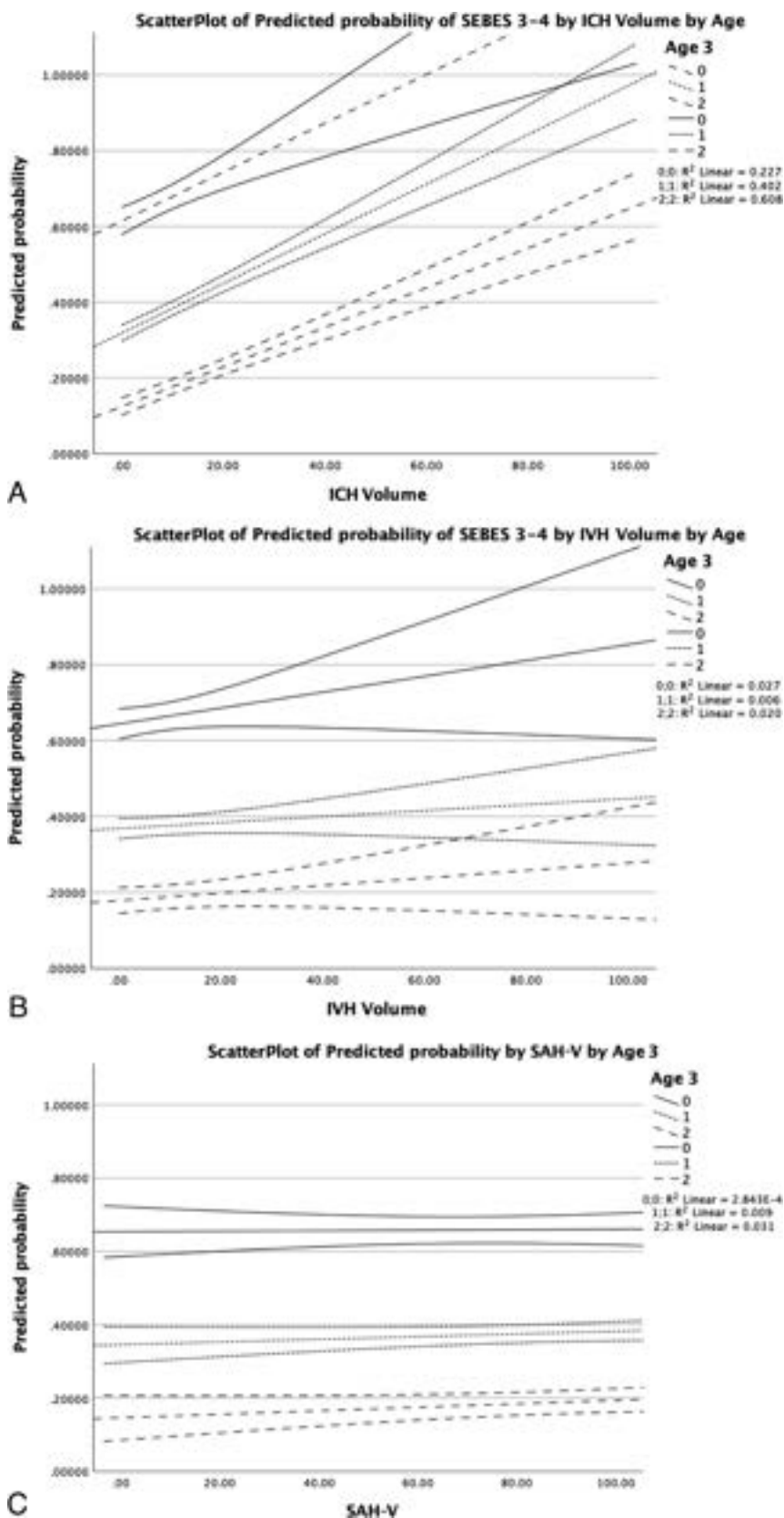
independently associated with improved outcome. Only ICH-V, among bleeding distributions, was independently associated with SEBES 3–4, (aOR, 1.02; 95% CI, 1.01–1.04;  $P < .001$ ). IVH-V and SAH-V emerged as independently associated with in-hospital mortality (aOR for IVH-V, 1.03; 95% CI, 1.01–1.05;  $P < .001$ ; aOR for SAH-V, 1.015; 95% CI, 1.01–1.02;  $P = .014$ ).

All 3 bleeding distributions emerged as independently associated with increased disability at long-term follow-up (Online Supplemental Data). Scatterplots showing multivariable analysis–derived predicted probabilities of severe SEBES on admission CT according to age and baseline GHV is reported in Fig 2, and according to age and different bleeding distributions in Fig 3. No collinearity was detected among independent predictors resulting from the multivariable models (variance inflation factor,  $<10$ ). The Online Supplemental Data report number of patients along with their clinical outcome stratified per SEBES grades, age, and corresponding GHV volumes.

## DISCUSSION

The volume of the primary bleed emerged as an independent predictor of both GCE and clinical outcome, and its clinical impact changed significantly throughout different age groups.

These results are notable because the population with poor-grade aSAH is the ideal one to study EBI, due to the high proportion of patients with severe neurologic deficit (WFNS IV–V), prolonged LOC, and the presence of GCE.<sup>6,7</sup> Blood extravasation occurring during aneurysmal rupture is known to cause an increase in ICP, a decrease in cerebral perfusion pressure, and transient ischemia (leading to LOC), part of EBI that,



**FIG 3.** Scatterplot showing multivariable analysis–derived predicted probabilities of severe SEBES (SEBES grades 3–4) on the admission CT scan according to age (3 groups: younger than 50, 50–70, older than 70 years of age and baseline ICH-V (A), IVH-V (B), and SAH-V (C).

radiologically, is shown by the effacement of sulci and progressive disruption of the gray-white matter junction,<sup>13</sup> the so-called GCE.

The association between volumetric clot determination and outcome is not novel per se. Several articles have studied it, with most aiming at DCI prediction.<sup>14–17</sup> Nonetheless, the present work reports novel data concerning the largest multicentric poor-grade population available to date, which is the ideal one to study GCE, the radiologic proxy of early brain injury. In addition, the relatively long follow-up (median, 24 months) is ideal to better capture the true proportion of long-term independent survivors, strengthening the results concerning independent outcome predictors.<sup>18</sup>

Previous studies failed to detect an association between the modified Fisher scale and GCE.<sup>6</sup> This failure may likely be due to the greater granularity provided by the volumetric assessment of intracranial bleeding, which emerged as independently associated with GCE and predictive of clinical outcome. Of particular interest is the relationship between the volumetrically assessed primary bleed, age and the studied outcomes. Few data are available concerning the impact of age on EBI development;<sup>6,13</sup> the reported results add further insight into the pathophysiologic relationship between age and GHV (Fig 2), showing a significant variation of the effect of GHV on severe GCE development throughout different age groups. In the studied population, a given GHV can either be tolerated or associated with GCE, in-hospital death, and worse long-term clinical outcome according to the age of the patient (Fig 2, and Online Supplemental Data). This finding may suggest the presence of age-related thresholds of intracranial compliance, reminiscent of the Monro-Kellie doctrine and its subsequent modifications.<sup>19,20</sup> Even though we cannot completely exclude the effect of brain atrophy due to aging, the results were confirmed by sensitivity analysis in the population younger than 70 years of

age, reinforcing the hypothesis that the effect is likely due to surpassing the intracranial compliance impairment threshold.<sup>21</sup> It has been shown in animal models of pathologic ICP after severe ischemic or hemorrhagic stroke that besides the reduction of blood and CSF within the cranium, the brain parenchyma itself plays an active role in intracranial compliance, with evidence of neuronal somata and extracellular space shrinkage far from the lesion site.<sup>20</sup> GCE has recently been shown to be independently associated with ICP-related secondary events like the need for decompressive craniectomy for the treatment of refractory high ICP,<sup>22</sup> making it a radiologic marker of intracranial hypertension.

The different effects of the GHV call for a better understanding of age-related pathologic responses. For example, cutoff values to start ICP-reducing therapies have varied historically according to different diseases (eg, diffuse traumatic brain injury versus focal mass effect due to ICH),<sup>23</sup> but there is no age-related cutoff of what can or should be considered pathologic. If confirmed by future studies, the critical role of GHV, a reproducible and measurable outcome predictor, could be of interest for practitioners directly involved in the emergency management of poor-grade aSAH and could help develop tailored, patient-specific treatments and timing in the EBI phase. Most interesting, even though GHV emerged as consistently and independently associated with the 3 predefined outcomes (GCE, in-hospital mortality, and long-term clinical outcome), its different distributions, namely ICH, IVH, and SAH, affected the predefined outcomes differently (Online Supplemental Data). ICH-V, a novel finding, emerged as independently associated with GCE development, likely due to the direct mass effect and swelling associated with large ICH and clinical outcome, but not with mortality. The impact on clinical outcome of the different bleeding distributions may be, at least partly, explained by the challenges inherent in their removal: Large space-occupying ICHs are usually removed during aneurysm clipping and associated with the use of decompressive craniectomy. This outcome may explain a reduced effect on mortality, while retaining a significant effect on disability. The removal of large IVH and SAH volumes still poses significant challenges and is less straightforward in clinical practice.<sup>24,25</sup>

Exemplar is the effect of EVD placement on clinical outcome, which is again modified by volume and varies (Online Supplemental Data) throughout different age groups. Besides, the choice to perform a primary decompressive craniectomy, currently investigated in one ad hoc randomized controlled trial,<sup>26</sup> could be influenced by the relationship between GHV and age. GHV could also be used for preclinical research, because the primary bleed precedes BBB disruption, GCE, neuroinflammation and oxidative cascades, which ultimately result in neuronal death.<sup>27</sup> In line with recent research,<sup>4</sup> all mortality predictors found by the reported multivariable model refer to the hyperacute phase of aSAH, with WFNS grade, rebleeding, and LOC confirmed as outcome predictors in aSAH.<sup>3,9,28</sup> The beneficial role of EVD has been previously reported.<sup>29,30</sup> Unlike prior results, we did not find treatment complications, IVH, or DCI as independent mortality predictors in the studied population, even though all were significantly associated with it in bivariate comparison.<sup>31</sup> The reason may be the disproportionate importance of EBI as a mortality predictor compared with the

delayed effect of DCI, which has been shown, along with decompressive craniectomy and the presence of ICH, as a predictor of increased disability.<sup>31</sup>

The strength of this study lies in the largest sample to date of poor-grade aSAH prospectively included in the Registry, allowing a robust analysis of outcome predictors. Moreover, it shows the feasibility of GHV calculation with DICOM sources coming from different settings and acquired with different scanners.

We acknowledge, nonetheless, several limitations, particularly those inherent to its Registry status. The absence of guidelines for the management of poor-grade aSAH adds complexity to the search for outcome predictors, and the reported results, even though each multivariable analysis was corrected for participating centers, could have unavoidable intrinsic biases. We could not provide reliable and easy-to-be-implemented cutoff volumes to predict functional independence in the studied population because they are necessarily influenced by treatment strategies. Separate studies, addressing each GHV distribution, are going to be better informative.

The reported results, though, picture the real-life poor-grade aSAH management, which deserves further research to standardize the management of such a fragile patient population. We quantified GCE, which is only a radiologic surrogate of EBI, a much more complex pathophysiologic mechanism during the early stage of aSAH. The reported results are far from fully delineating all the causative mechanisms of it, and we hope that future studies are going to address this topic. Moreover, the present study does not report cerebral edema resolution and its predictors. We hope that the results are going to be considered by future research addressing this topic.

## CONCLUSIONS

GHV after aneurysmal rupture emerged as both a radiologic marker of EBI and an outcome predictor. Its effects seem to be significantly modified by age. ICH-V, IVH-V, and SAH-V have different effects on clinical outcome. This finding may be of relevance for the design of future therapeutic studies. Moreover, the reported results could be of help for practitioners involved in the management of poor-grade aSAH, particularly in the EBI phase, as well as for caregivers and practitioners involved in the subacute management of these patients.

**Disclosure forms** provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

1. Hoh BL, Ko NU, Amin-Hanjani S, et al. **2023 Guideline for the Management of Patients With Aneurysmal Subarachnoid Hemorrhage: A Guideline From the American Heart Association/American Stroke Association.** *Stroke* 2023;54:e314–70 CrossRef Medline
2. Kusaka G, Ishikawa M, Nanda A, et al. **Signaling pathways for early brain injury after subarachnoid hemorrhage.** *J Cereb Blood Flow Metab* 2004;24:916–25 CrossRef Medline
3. Suwatcharakoon S, Meyers E, Falo C, et al. **Loss of consciousness at onset of subarachnoid as an important marker of early brain injury.** *JAMA Neurol* 2016;73:28–35 CrossRef Medline
4. Komotar RJ, Schmidt JM, Starke RM, et al. **Resuscitation and critical care of poor-grade subarachnoid hemorrhage.** *Neurosurgery* 2009;64:397–410 CrossRef Medline

5. Broderick JP, Brott TG, Duldner JE, et al. **Initial and recurrent bleeding are the major causes of death following subarachnoid hemorrhage.** *Stroke* 1994;25:1342–47 CrossRef Medline
6. Claassen J, Carhuapoma JR, Kreiter KT, et al. **Global cerebral edema after subarachnoid hemorrhage: frequency, predictors, and impact on outcome.** *Stroke* 2002;33:1225–32 CrossRef Medline
7. Ahn SH, Savarraj JP, Pervez M, et al. **The subarachnoid hemorrhage early brain edema score predicts delayed cerebral ischemia and clinical outcomes.** *Neurosurgery* 2018;83:137–45 CrossRef Medline
8. **Report of World Federation of Neurological Surgeons Committee on a Universal Subarachnoid Hemorrhage Grading Scale.** *J Neurosurg* 1988;68:985–86 CrossRef Medline
9. Chiang VL, Claus EB, Awad IA. **Toward more rational prediction of outcome in patients with high-grade subarachnoid hemorrhage.** *Neurosurgery* 2000;46:28–35; discussion 35–36 CrossRef Medline
10. Vergouwen MD, Vermeulen M, van Gijn J, et al. **Definition of delayed cerebral ischemia after aneurysmal subarachnoid hemorrhage as an outcome event in clinical trials and observational studies: proposal of a multidisciplinary research group.** *Stroke* 2010;41:2391–95 CrossRef Medline
11. Claassen J, Bernardini GL, Kreiter K, et al. **Effect of cisternal and ventricular blood on risk of delayed cerebral ischemia after subarachnoid hemorrhage: the Fisher scale revisited.** *Stroke* 2001;32:2012–20 CrossRef Medline
12. Panni P, Colombo E, Donofrio CA, et al. **Hemorrhagic burden in poor-grade aneurysmal subarachnoid hemorrhage: a volumetric analysis of different bleeding distributions.** *Acta Neurochir (Wien)* 2019;161:791–97 CrossRef Medline
13. Rass V, Helbok R. **Early brain injury after poor-grade subarachnoid hemorrhage.** *Curr Neurol Neurosci Rep* 2019;19:78 CrossRef Medline
14. van der Steen WE, Zijlstra IA, Verbaan D, et al. **Association of quantified location-specific blood volumes with delayed cerebral ischemia after aneurysmal subarachnoid hemorrhage.** *AJNR Am J Neuroradiol* 2018;39:1059–64 CrossRef Medline
15. van der Steen WE, Marquering HA, Ramos LA, et al. **Prediction of outcome using quantified blood volume in aneurysmal SAH.** *AJNR Am J Neuroradiol* 2020;41:1015–21 CrossRef Medline
16. Ko SB, Choi HA, Carpenter AM, et al. **Quantitative analysis of hemorrhage volume for predicting delayed cerebral ischemia after subarachnoid hemorrhage.** *Stroke* 2011;42:669–74 CrossRef Medline
17. Lagares A, Jiménez-Roldán L, Gomez PA, et al. **Prognostic value of the amount of bleeding after aneurysmal subarachnoid hemorrhage: a quantitative volumetric study.** *Neurosurgery* 2015;77:898–907; discussion 907 CrossRef Medline
18. Wilson DA, Nakaji P, Albuquerque FC, et al. **Time course of recovery following poor-grade SAH: the incidence of delayed improvement and implications for SAH outcome study design.** *J Neurosurg* 2013;119:606–12 CrossRef Medline
19. Benson JC, Madhavan AA, Cutsforth-Gregory JK, et al. **The Monro-Kellie doctrine: a review and call for revision.** *AJNR Am J Neuroradiol* 2022;44:2–6 CrossRef Medline
20. Kalisvaart AC, Wilkinson CM, Gu S, et al. **An update to the Monro-Kellie doctrine to reflect tissue compliance after severe ischemic and hemorrhagic stroke.** *Sci Rep* 2020;10:22013 CrossRef Medline
21. Godoy DA, Brasil S, Iaccarino C, et al. **The intracranial compartmental syndrome: a proposed model for acute brain injury monitoring and management.** *Crit Care* 2023;27:137 CrossRef Medline
22. Said M, Gümüş M, Herten A, et al. **Subarachnoid Hemorrhage Early Brain Edema Score (SEBES) as a radiographic marker of clinically relevant intracranial hypertension and unfavorable outcome after subarachnoid hemorrhage.** *Eur J Neurol* 2021;28:4051–59 CrossRef Medline
23. Wijdicks EF. **10 or 15 or 20 or 40 mmHg? What is increased intracranial pressure and who said so?** *Neurocrit Care* 2022;36:1022–26 CrossRef Medline
24. Hanley DF, Lane K, McBee N, et al; CLEAR III Investigators. **Thrombolytic removal of intraventricular haemorrhage in treatment of severe stroke: results of the randomised, multicentre, multicenter, placebo-controlled CLEAR III trial.** *Lancet* 2017;389:603–11 CrossRef Medline
25. Wolf S, Mielke D, Barner C, et al. **EARLYDRAIN Study Group. Effectiveness of lumbar cerebrospinal fluid drain among patients with aneurysmal subarachnoid hemorrhage: a randomized clinical trial.** *JAMA Neurol* 2023;80:833–42 CrossRef Medline
26. Güresir E, Lampmann T, Brandecker S, et al. **Primary Decompressive Craniectomy in Aneurysmal Subarachnoid hemorrhage (PICASSO) trial: study protocol for a randomized controlled trial.** *Trials* 2022;23:1027 CrossRef Medline
27. Wu Y, Pang J, Peng J, et al. **Apolipoprotein E deficiency aggravates neuronal injury by enhancing neuroinflammation via the JNK/c-Jun pathway in the early phase of experimental subarachnoid hemorrhage in mice.** *Oxid Med Cell Longev* 2019;2019:3832648 CrossRef Medline
28. Goldberg J, Schoeni D, Mordasini P, et al. **Survival and outcome after poor-grade aneurysmal subarachnoid hemorrhage in elderly patients.** *Stroke* 2018;49:2883–89 CrossRef Medline
29. Baggiani M, Graziano F, Rebora P, et al. **Intracranial pressure monitoring practice, treatment, and effect on outcome in aneurysmal subarachnoid hemorrhage.** *Neurocrit Care* 2022;38:741–51 CrossRef Medline
30. Bailes JE, Spetzler RF, Hadley MN, et al. **Management morbidity and mortality of poor-grade aneurysm patients.** *J Neurosurg* 1990;72:559–66 CrossRef Medline
31. Stienen MN, Germans M, Burkhardt JK, et al; Swiss SOS Study Group. **Predictors of in-hospital death after aneurysmal subarachnoid hemorrhage: analysis of a nationwide database (Swiss SOS [Swiss Study on Aneurysmal Subarachnoid Hemorrhage]).** *Stroke* 2018;49:333–40 CrossRef Medline

# The Management of Persistent Distal Occlusions after Mechanical Thrombectomy and Thrombolysis: An Inter- and Intrarater Agreement Study

W. Boisseau, A. Benomar, C. Ducroux, R. Fahed, S. Smajda, J. D. B. Diestro, G. Charbonnier, J. Ognard, J. Burel, A. Ter Schiphorst, M. Boulanger, A. Nehme, J. Boucherit, G. Marnat, D. Volders, Q. Holay, G. Forestier, M. Bretzner, D. Roy, S. Vingadassalom, M. Elhorany, L. Nico, G. Jacquin, M. Abdalkader, A. Guedon, P. Seners, K. Janot, V. Dumas, R. Olatunji, S. Gazzola, G. Milot, J. Zehr, T.E. Darsaut, D. Iancu, and J. Raymond



## ABSTRACT

**BACKGROUND AND PURPOSE:** The best management of patients with persistent distal occlusion after mechanical thrombectomy with or without IV thrombolysis remains unknown. We sought to evaluate the variability and agreement in decision-making for persistent distal occlusions.

**MATERIALS AND METHODS:** A portfolio of 60 cases was sent to clinicians with varying backgrounds and experience. Responders were asked whether they considered conservative management or rescue therapy (stent retriever, aspiration, or intra-arterial thrombolysis) a treatment option as well as their willingness to enroll patients in a randomized trial. Agreement was assessed using  $\kappa$  statistics.

**RESULTS:** The electronic survey was answered by 31 physicians (8 vascular neurologists and 23 interventional neuroradiologists). Decisions for rescue therapies were more frequent ( $n = 1116/1860$ , 60%) than for conservative management ( $n = 744/1860$ , 40%;  $P < .001$ ). Interrater agreement regarding the final management decision was “slight” ( $\kappa = 0.12$ ; 95% CI, 0.09–0.14) and did not improve when subgroups of clinicians were studied according to background, experience, and specialty or when cases were grouped according to the level of occlusion. On delayed re-questioning, 23 of 29 respondents (79.3%) disagreed with themselves on at least 20% of cases. Respondents were willing to offer trial participation in 1295 of 1860 (69.6%) cases.

**CONCLUSIONS:** Individuals did not agree regarding the best management of patients with persistent distal occlusion after mechanical thrombectomy and IV thrombolysis. There is sufficient uncertainty to justify a dedicated randomized trial.

**ABBREVIATIONS:** CA = contact aspiration; IAT = Intra-arterial thrombolysis; IVT = IV thrombolysis; MT = mechanical thrombectomy; RCT = randomized controlled trial; SR = stent retriever

Incomplete recanalization after mechanical thrombectomy (MT) and/or IV thrombolysis (IVT) can result in persistent

distal occlusion of the middle, anterior, or posterior cerebral arteries or branches.<sup>1</sup> Despite increased expertise and a multitude

Received August 22, 2023; accepted after revision October 24.

From the Department of Interventional Neuroradiology (W.B., S.S.), Fondation Rothschild Hospital, Paris, France; Department of Radiology (W.B., A.B., D.R., D.L., J.R.), Centre Hospitalier de l'Université de Montréal Montréal, Canada; Department of Neurology (C.D., R.F.), Ottawa Hospital Research Institute & University of Ottawa, Ottawa, Canada; Division of Diagnostic and Therapeutic Neuroradiology (J.D.B.D., R.O.), Department of Medical Imaging, St. Michael's Hospital, University of Toronto, Toronto, Ontario, Canada; Interventional Neuroradiology Department (G.C.), Besançon University Hospital, Besançon, France; Interventional Neuroradiology Department (J.O.), Hôpital de la Cavale Blanche, Brest, Bretagne, France; Department of Radiology (J. Burel), Rouen University Hospital, Rouen, France; Neurology Department (A.T.S.), CHRU Gui de Chauliac, Montpellier, France; Department of Neurology (M. Boulanger, A.N.), Caen University Hospital, Caen, France; Department of Radiology (J. Boucherit), Rennes University Hospital, Rennes, France; Department of Neuroradiology (G. Marnat), University Hospital of Bordeaux, Bordeaux, France; Department of Radiology (D.V.), Dalhousie University, Halifax, Nova Scotia, Canada; Radiology Department (Q.H.) and Neurology Department (S.G.), Hôpital d'Instruction des armées Saint-Anne, Toulon, France; Neuroradiology Department (G.F.), University Hospital of Limoges, Limoges, France; Neuroradiology Department (M. Bretzner), CHU Lille, University Lille, Inserm, U1172 Lille Neuroscience & Cognition, F-59000, Lille, France; Interventional Neuroradiology Department (S.V.), CHRU Marseille, La Timone, France; Department of Neuroradiology (M.E.), Groupe Hospitalier de Pitié Salpêtrière, Paris, France; Department of Neurology (M.E.), Tanta University,

Tanta, Egypt; Department of Radiology (L.N.), University Hospital of Padova, Padova, Italy; Neurovascular Health Program (G.J.), Centre Hospitalier de l'Université de Montréal (CHUM), Montreal, Quebec, Canada; Department of Radiology (M.A.), Boston Medical Center, Boston University Chobanian & Avedisian School of Medicine, Boston, Massachusetts; Department of Neuroradiology (A.G.), Lariboisière Hospital, Paris, France; Neurology Department (P.S.), Hôpital Fondation A. de Rothschild, Fondation Rothschild Hospital, Paris, France; Institut de Psychiatrie et Neurosciences de Paris (P.S.), UMR\_S1266, INSERM, Université de Paris, Paris, France; Interventional Neuroradiology (K.J.), University Hospital of Tours, Tours, France; Radiology Department (V.D.), University Hospital of Poitiers, Poitiers, France; Department of Neurosurgery (G. Milot), CHU de Québec, Québec, Canada; Department of Mathematics and Statistics (J.Z.), Pavillon André-Aisenstadt, Montreal, Quebec, Canada; and Department of Surgery (T.E.D.), Division of Neurosurgery, University of Alberta Hospital, Mackenzie Health Sciences Centre, Edmonton, Alberta, Canada.

Please address correspondence to Jean Raymond, MD, Centre Hospitalier de l'Université de Montréal (CHUM), Department of Radiology, Room D03.5462b, Montreal, PQ, Canada H2X 0C1; e-mail: jean.raymond@umontreal.ca; @INR\_CHUM; @tdarsaut; @dana\_iancu00; @RaymondJeanMD1; @AnassBenomarMD

Indicates article with online supplemental data.  
<http://dx.doi.org/10.3174/ajnr.A8149>



of new thrombectomy devices, incomplete recanalization remains a frequent clinical scenario, observed in almost 50% of patients in recent randomized trials.<sup>1,2</sup>

Better recanalization grades after thrombectomy are associated with improved clinical outcomes, but this post hoc observation does not mean that persistent distal occlusions should be treated.<sup>3-7</sup> Various interventions have been proposed for treating such occlusions, including small adjustable stent retrievers,<sup>8-12</sup> distal aspiration catheters,<sup>8,12</sup> and intra-arterial thrombolytics, but there is no consensus on which management is best.<sup>12,13</sup> Although these treatments can improve the reperfusion status of patients with reportedly little added risk,<sup>8-12,14,15</sup> there is currently no randomized evidence that these complementary treatments lead to better outcomes.

There have been several reliability studies on thrombolysis or thrombectomy decisions, but none have specifically addressed distal occlusions.<sup>16</sup> Reliability and agreement studies on management decisions can measure the degree of clinical uncertainty and inform the design of randomized trials.<sup>17,18</sup> With this aim in mind, we investigated the agreement of various experts in making decisions regarding the management of patients with persistent distal occlusion after MT or IVT.

## MATERIALS AND METHODS

This study was prepared in accordance with the Guidelines for Reporting Reliability and Agreement Studies (GRRAS).<sup>19</sup>

### Case Selection

We selected and assembled an electronic portfolio of catheter angiograms of 60 patients who underwent endovascular thrombectomy for acute ischemic stroke with large-vessel occlusion (internal carotid, proximal MCA, or basilar trunk occlusion) in a single comprehensive stroke center (Centre hospitalier de l'Université de Montréal [CHUM]) between December 2020 and July 2021. To minimize  $\kappa$  paradoxes,<sup>20</sup> 1 author (W.B.) selected the cases so that approximately one-third were cases for whom conservative management was expected to be a frequent choice, one-third of cases were those for whom rescue therapy was expected to be a frequent choice, and one-third of cases were "gray zone" cases for whom both options would likely be considered.

A wide variety of distal occlusions were included (middle, anterior, and posterior cerebral arteries) at various levels (such as M2-to-M4 segments of the MCA, A1-to-A5 segments of the anterior, and P1-to-P5 segments of the posterior cerebral arteries, as in previous reports).<sup>21-23</sup> In cases with multiple tandem occlusions, the most proximal was defined as the level of occlusion.

The number of cases in the present study ( $n = 60$ ) was selected to cover a wide spectrum of distal occlusions and to ensure relatively small confidence intervals according to tables provided by Donner and Michael.<sup>24</sup>

### Raters

Two hundred eleven clinicians were invited to participate, including 148 interventional neuroradiologists and 63 vascular neurologists from 6 different countries.

### Agreement Study

Catheter angiograms were de-identified and uploaded onto a local secure server. Raters had no access to other imaging studies

or clinical information other than sex, age, symptoms (ie, left or right motor deficit, aphasia), the initial NIHSS score at presentation, the ASPECTS or its variant for the posterior circulation on pretreatment imaging,<sup>25</sup> pretreatment with IVT, the time from symptom onset to proximal recanalization (when available; or in case of unknown onset, the time from last seen well and from stroke discovery to proximal recanalization were provided), and the time from CT to proximal recanalization. Study data were collected and managed using the REDCap electronic data capture tools (<https://www.project-redcap.org/>) hosted at our institution.<sup>26,27</sup>

For each case, raters were asked to independently answer the 5 following questions:

1) Is rescue therapy (stent retriever [SR], contact aspiration [CA], or intra-arterial thrombolysis) an option for this patient (Yes/No)? 2) Is conservative management an option for this patient (Yes/No)? 3) What is your final best treatment choice? Choose one: rescue therapy; conservative management. If you choose rescue therapy, please choose one: SR, CA, combined use of SR and CA, and intra-arterial thrombolytic (ie, recombinant tPA or urokinase). 4) How confident are you regarding your final treatment choice? (answers in 10% increments, ranging from 0% to 100%) 5) Would you be willing to recruit this patient for a randomized controlled trial (RCT) that would give a 50% chance of rescue therapy and a 50% chance of conservative management? (Yes/No).

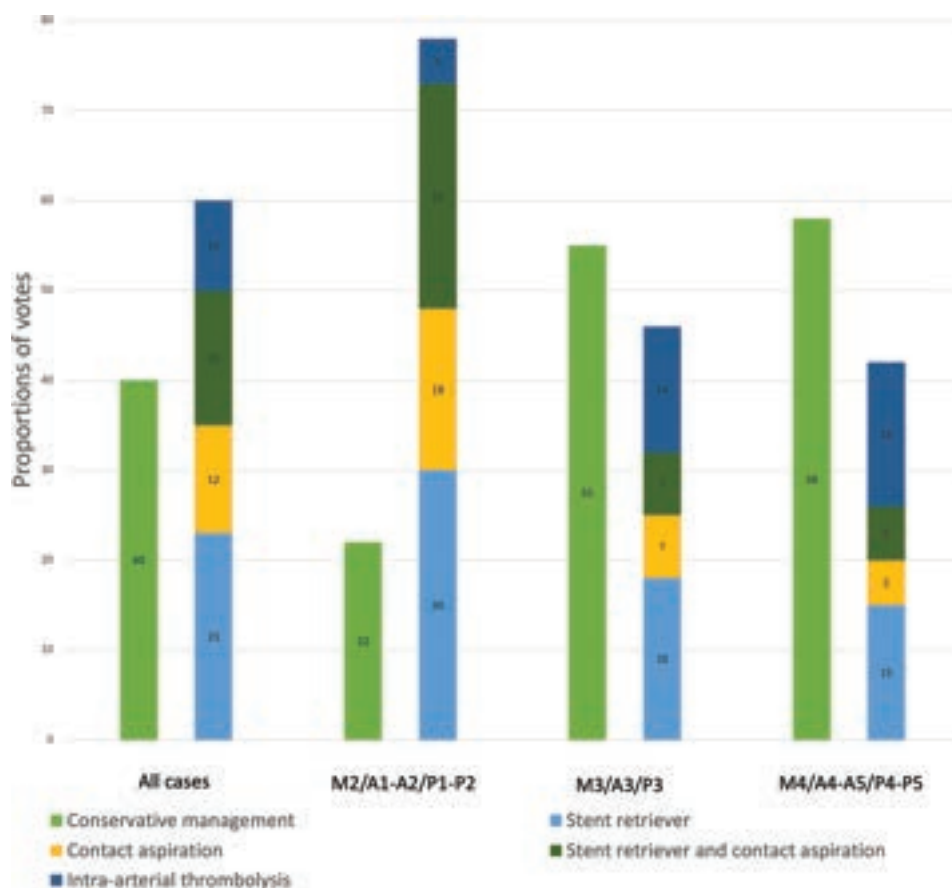
All raters were asked to perform a second reading session (with cases presented in a permuted order), at least 2 months later, to independently evaluate intrarater agreement.

### Statistical Analysis

Descriptive statistics are presented using percentage for categorical variables and mean (SD) for continuous variables. The proportions of answers for each question among different groups of raters (background, specialty, years of experience) of the level of persistent occlusion (M2/A1-A2/P1-P2; M3/A3/P3; M4/A4-A5/P4-P5) were compared using a 1-way ANOVA. When applicable, we used pair-wise comparisons using a Bonferroni adjustment. Correlations between treatment choice and continuous variables (age, NIHSS score, ASPECTS, time between onset and proximal recanalization and between CT and recanalization, number of passes) were analyzed using the Pearson correlation coefficient. Confidence in decision-making (scale of 0–100) was analyzed using ANOVA.  $\chi^2$  tests were used to compare the willingness to recruit patients in a RCT. Agreement between and within raters was measured using  $\kappa$  statistics and 95% bootstrap confidence intervals and was interpreted according to Landis and Koch.<sup>28</sup>  $\kappa$  values ranged from  $-1$  (perfect disagreement) to  $+1$  (perfect agreement), with zero indicating no agreement among the raters other than that expected by chance alone. Analyses were performed using STATA (Version 16.1; StataCorp) and SPSS software (Version 26; IBM) with significance set at 5%.

## RESULTS

Patients, treatment characteristics, and main outcomes of cases included in the portfolio are detailed in the Online Supplemental Data. Details regarding raters are provided in the Online Supplemental Data. Among the 211 clinicians who were invited



**FIG 1.** Proportions of final treatment choices for all cases and according to several levels of persistent distal occlusion. Note the strong correlation between the level of persistent distal occlusion and the choice of treatment.

to participate, 31 (14.7%) responded, including 8 vascular neurologists and 23 interventional neuroradiologists from 20 comprehensive stroke centers; 29 of 31 raters (93.6%) agreed to review the same cases in a permuted order at least 2 months later.

Rescue therapy was judged to be a treatment option by a mean of 20 (SD, 7) raters (minimum 6; maximum, 31). Interrater agreement regarding whether rescue therapy was an option was fair ( $\kappa = 0.22$ ; 95% CI, 0.17–0.28) for all raters and subgroups (Online Supplemental Data).

Conservative management was judged to be an option by a mean of 25 (SD, 6) raters (minimum, 6; maximum, 31). Interrater agreement regarding conservative management was fair ( $\kappa = 0.21$ ; 95% CI, 0.13–0.30) for all raters and subgroups (Online Supplemental Data).

Overall, rescue therapy was more frequently selected as the final best management: 1116/1860 responses (60%; 95% CI, 59%–61%), with conservative management garnering 744/1860 votes (40%; 95% CI, 39%–41%) ( $P < .001$ ) (Online Supplemental Data). Final treatment choices did not vary according to years of experience ( $P = .145$ ), background ( $P = .153$ ), or interventional practice ( $P = .897$ ). The level of the occlusion influenced the final treatment choice: Rescue therapy was more frequently selected as the final management for M2/A1-A2/P1-P2 than for more distal occlusions ( $P < .001$ ). The proportions of votes for each rescue therapy for all raters and subgroups and according to the level of occlusion are illustrated in Fig 1 and detailed in the Online Supplemental Data.

The number of rescue therapy and conservative choices for each of the 60 cases is illustrated in Fig 2.

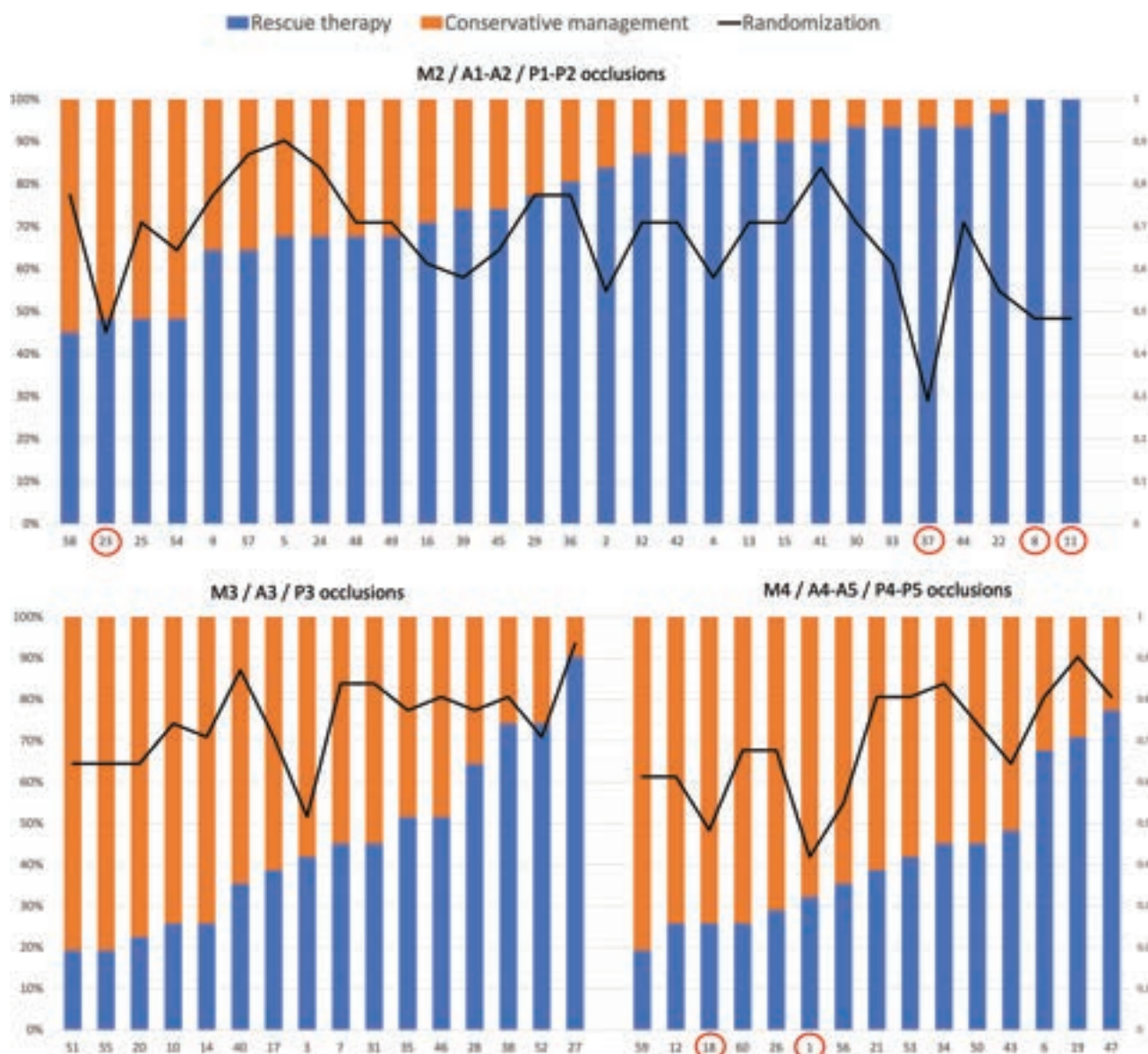
The initial NIHSS score (0.279,  $P = .031$ ) correlated with more rescue therapy choices. The previous use of IVT was associated with more conservative choices ( $P = .037$ ).

Interrater agreement regarding the final management decision was “slight” ( $\kappa = 0.12$ ; 95% CI, 0.09–0.14) and did not improve when answers were dichotomized (conservative management versus any rescue therapy); when subgroups of clinicians were studied according to background, experience, and specialty; or when cases were grouped according to the level of occlusion (Fig 3 and Online Supplemental Data).

The proportion of final decisions for rescue therapy, intrarater agreement (ie, proportion of cases with the same judgment between both rating sessions), and the proportion of patients recruited in an RCT for each rater are illustrated in Fig 4.

Raters changed their final decision (rescue therapy versus conservative treatment) between readings in a mean 25.5% of cases; 23/29 (79.3%) raters changed their final decision between readings in at least 20% of cases (Online Supplemental Data). Intrarater  $\kappa$  values regarding the best final management (conservative management versus any rescue therapy) reached a substantial level (ie,  $\kappa > 0.6$ ) for 4/29 (13.8%) raters.

Clinicians were not very confident in their final decisions (mean confidence, 65%; minimum-maximum, 24%–68%). Recruitment of patients in an RCT comparing rescue therapy and conservative



**FIG 2.** Proportions of votes for rescue therapy and conservative management for several levels of occlusion and the proportion of clinicians willing to include the patient in an RCT (line) for each case ( $n = 60$ ). Red circles indicate cases with  $<50\%$  of raters willing to include the patient in an RCT (4/29 for patients with M2/A1-A2/P1-P2 occlusion, none for M3/A3/P3 occlusions, and 2/15 for patients with M4/A4-A5/P4-P5).

management is graphically displayed in Fig 2 (per patient) and Fig 4 (per rater). Trial participation was offered 69.6% of the time (1295/1860 Yes responses to question 5). In 54/60 (90%) cases, a majority ( $\geq 51\%$ ) of responders were willing to include the patient in a randomized trial. Willingness to be included in a trial did not vary according to years of experience ( $P = .624$ ), background ( $P = .328$ ), or specialty ( $P = .445$ ) (Online Supplemental Data). Interrater agreement regarding recruitment in an RCT was slight ( $\kappa = 0.05$ ; 95% CI, 0.03–0.09) (Online Supplemental Data).

## DISCUSSION

The current study highlights the clinical uncertainty regarding the management of patients with persistent distal occlusions after MT or IVT. Overall, agreement on the management of persistent

occlusions was barely above that expected by chance alone for all cases and all clinicians. The lack of agreement among clinicians has a more fundamental nature than a divergence in the opinions of experts from different specialties, for there was substantial variability at the level of individual clinicians themselves when they were asked on 2 different occasions to judge the management of the same patients. Only 4 raters reached substantial agreement with themselves, and 75% recommended different management for  $\geq 20\%$  of the cases.

A recent trend to actively intervene is perhaps reflected in this study by the 60%/40% proportions of answers in favor of intervention. A more proximal persistent occlusion was found to be the most influential factor in clinical decision-making; the presenting NIHSS score and the use of IVT were other factors that seemed to influence treatment decisions.

The uncertainty brought to light by this study should encourage our community to pursue the clinical research necessary to better define the optimal management of patients with persistent distal occlusions after MT and/or IVT. We believe that in the presence of such uncertainty, patients are best managed within the context of a care trial,<sup>17,29</sup> in which they are given a 50% chance of receiving a promising treatment of yet-unknown benefit but an equal 50% chance of receiving more conservative alternative management.<sup>30</sup>

This survey also provides information regarding the feasibility of such a trial. Most clinicians (27 of 31 or 87%) claimed that they would recommend RCT participation to at least 50% of patients. In order to be eligible for the trial, both rescue therapy and conservative management should be considered reasonable treatment options, which was the case for nearly all patients. The survey also

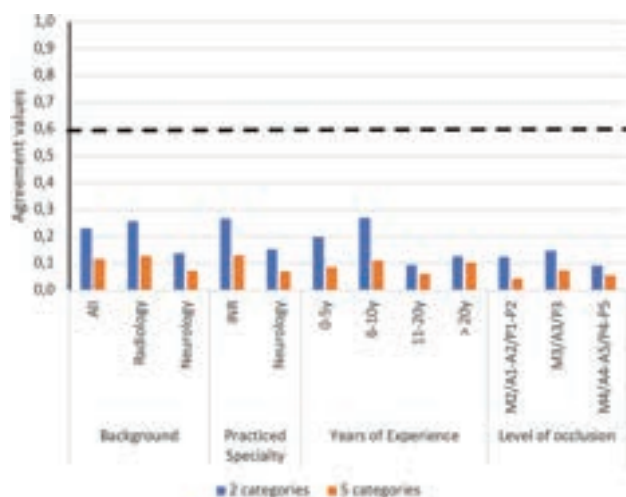
provides some indications regarding trial design. For example, patients with all sorts of persistent occlusions should be included because the uncertainty concerned all sites. Moreover, given the lack of agreement regarding the wide variety of interventional management strategies that were chosen by raters, we believe that the trial should not be limited to a specific device. The main question that needs to be addressed is whether rescue therapy (any selection or combination of SR, CA, and IAT) truly offers a safe and more effective alternative than conservative management.

There are several limitations to this study. The number of cases was relatively small ( $n=60$ ), to encourage participation of multiple raters. Creating a portfolio of artificially balanced cases was important to minimize  $\kappa$  paradoxes,<sup>15</sup> but a different selection might have produced different results. Raters were not a random sample representative of a population of clinicians, a requirement that is not necessary when reliability cannot be shown within or between observers in a sufficiently wide range of clinicians.<sup>18</sup> The survey did not include patients with a primary distal occlusion (without proximal occlusion). Thus, our study does not provide any insight into the management of such patients. All raters were informed that the cases included in this survey had at least 1 persistent distal occlusion. Therefore, our study did not evaluate the raters' ability to detect the presence or absence of one or multiple distal occlusions. Finally, completing an electronic survey is different from caring for real patients. The degree to which responders imagined they were dealing with serious clinical decisions can only be surmised.

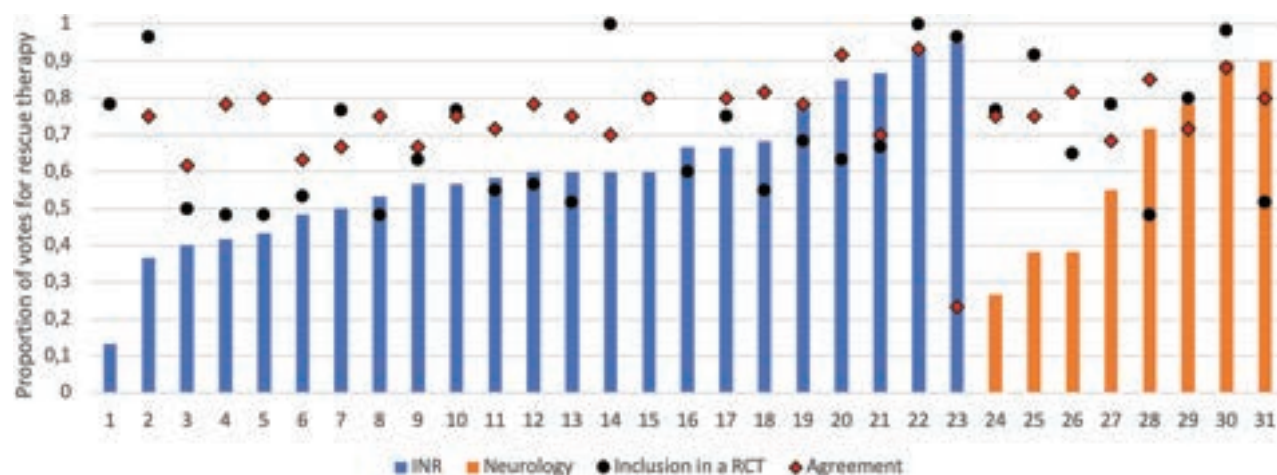
## CONCLUSIONS

There is a lack of consensus among clinicians regarding the optimal management approach for patients with persistent distal occlusion following MT with or without IVT. This level of uncertainty justifies the need for a randomized trial to compare rescue therapy and conservative management in this context.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).



**FIG 3.** Interrater agreement for final treatment choices (5 categories: conservative management, SR, CA, combined use of SR and CA, intra-arterial thrombolysis) was dichotomized as conservative management versus all others (2 categories) for all raters, according to background, practice specialty, years of experience, and level of occlusion. All  $\kappa$  values are well below the dashed "substantial" line (0.600). Y indicates years.



**FIG 4.** Proportion of final decisions for rescue therapy (blue and orange bars, ordered by numbers and classified by specialty), intrarater agreement (ie, proportion of cases with same judgment between readings [red diamonds]), and proportion of patients recruited in an RCT (black circles) for each rater ( $n=31$ ). INR indicates interventional neuroradiology.



## REFERENCES

- Hill MD, Goyal M, Menon BK, et al; ESCAPE-NA1 Investigators. Efficacy and safety of tenecteplase in the treatment of acute ischaemic stroke (ESCAPE-NA1): a multicentre, double-blind, randomised controlled trial. *Lancet* 2020;395:878–87 CrossRef Medline
- van der Steen W, van de Graaf RA, Chalos V, et al; MR CLEAN-MED investigators. Safety and efficacy of aspirin, unfractionated heparin, both, or neither during endovascular stroke treatment (MR CLEAN-MED): an open-label, multicentre, randomised controlled trial. *Lancet* 2022;399:1059–69 CrossRef Medline
- Almekhlafi MA, Mishra S, Desai JA, et al. Not all “successful” angiographic reperfusion patients are an equal validation of a modified TICI scoring system. *Interv Neuroradiol* 2014;20:21–27 CrossRef Medline
- Tung EL, McTaggart RA, Baird GL, et al. Rethinking thrombolysis in cerebral infarction 2b: which thrombolysis in cerebral infarction scales best define near complete recanalization in the modern thrombectomy era? *Stroke* 2017;48:2488–93 CrossRef Medline
- Dargazanli C, Fahed R, Blanc R, et al; ASTER Trial Investigators. Modified thrombolysis in cerebral infarction 2C/thrombolysis in cerebral infarction 3 reperfusion should be the aim of mechanical thrombectomy: insights from the ASTER Trial (Contact Aspiration Versus Stent Retriever for Successful Revascularization). *Stroke* 2018;49:1189–96 CrossRef Medline
- Liebeskind DS, Bracard S, Guillemin F, et al; HERMES Collaborators. eTICI reperfusion: defining success in endovascular stroke therapy. *J Neurointerv Surg* 2019;11:433–38 CrossRef Medline
- Kaesmacher J, Dobrocky T, Heldner MR, et al. Systematic review and meta-analysis on outcome differences among patients with TICI2b versus TICI3 reperfusion: success revisited. *J Neurol Neurosurg Psychiatry* 2018;89:910–17 CrossRef Medline
- Grossberg JA, Rebello LC, Haussen DC, et al. Beyond large vessel occlusion strokes: distal occlusion thrombectomy. *Stroke* 2018;49:1662–68 CrossRef Medline
- Hofmeister J, Kulcsar Z, Bernava G, et al. The Catch Mini stent retriever for mechanical thrombectomy in distal intracranial occlusions. *J Neuroradiol* 2018;45:305–09 CrossRef Medline
- Rikhtegar R, Mosimann PJ, Weber R, et al. Effectiveness of very low profile thrombectomy device in primary distal medium vessel occlusion, as rescue therapy after incomplete proximal recanalization or following iatrogenic thromboembolic events. *J Neurointerv Surg* 2021;13:1067–72 CrossRef Medline
- Guenego A, Mine B, Bonnet T, et al. Thrombectomy for distal medium vessel occlusion with a new generation of Stent retriever (Tigertriever 13). *Interv Neuroradiol* 2022;28:444–54 CrossRef Medline
- Meyer L, Stracke P, Wallocha M, et al; from the TOPMOST Study Group. Aspiration versus stent retriever thrombectomy for distal, medium vessel occlusion stroke in the posterior circulation: a subanalysis of the TOPMOST Study. *Stroke* 2022;53:2449–57 CrossRef Medline
- Kashani N, Cimflova P, Ospel JM, et al. Endovascular device choice and tools for recanalization of medium vessel occlusions: insights from the MeVO FRONTIERS International Survey. *Front Neurol* 2021;12:735899 CrossRef Medline
- Kaesmacher J, Bellwald S, Dobrocky T, et al. Safety and efficacy of intra-arterial urokinase after failed, unsuccessful, or incomplete mechanical thrombectomy in anterior circulation large-vessel occlusion stroke. *JAMA Neurol* 2020;77:318–26 CrossRef Medline
- Kaesmacher J, Meinel TR, Kurmann C, et al. Safety and efficacy of intra-arterial fibrinolytics as adjunct to mechanical thrombectomy: a systematic review and meta-analysis of observational data. *J Neurointerv Surg* 2021;13:1073–80 CrossRef Medline
- Ducroux C, Fahed R, Khoury NN, et al; FAMOUS Collaborative Group. Intravenous thrombolysis and thrombectomy decisions in acute ischemic stroke: an interrater and intrarater agreement study. *Rev Neurol (Paris)* 2019;175:380–89 CrossRef Medline
- Fahed R, Darsaut TE, Farzin B, et al. Measuring clinical uncertainty as a preliminary step to randomized controlled trials. *J Clin Epidemiol* 2019;112:96–98 CrossRef Medline
- Fahed R, Darsaut TE, Farzin B, et al. Measuring clinical uncertainty and equipoise by applying the agreement study methodology to patient management decisions. *BMC Med Res Methodol* 2020;20:214 CrossRef Medline
- Kottner J, Audigé L, Brorson S, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol* 2011;64:96–106 CrossRef Medline
- Feinstein AR, Cicchetti DV. High agreement but low kappa: I, the problems of two paradoxes. *J Clin Epidemiol* 1990;43:543–49 CrossRef Medline
- Gibo H, Carver CC, Rhoton AL Jr, et al. Microsurgical anatomy of the middle cerebral artery. *J Neurosurg* 1981;54:151–69 CrossRef Medline
- Párraga RG, Ribas GC, Andrade SE, et al. Microsurgical anatomy of the posterior cerebral artery in three-dimensional images. *World Neurosurg* 2011;75:233–57 CrossRef Medline
- Perlmutter D, Rhoton AL Jr. Microsurgical anatomy of the distal anterior cerebral artery. *J Neurosurg* 1978;49:204–28 CrossRef Medline
- Donner A, Michael R. A. Sample size requirements for interval estimation of the kappa statistic for interobserver agreement studies with a binary outcome and multiple raters. *Int J Biostat* 2010;6: Article 31 CrossRef Medline
- Barber PA, Hill MD, Eliasziw M, et al; Imaging ASPECTS Study Group. Imaging of the brain in acute ischaemic stroke: comparison of computed tomography and magnetic resonance diffusion-weighted imaging. *J Neurol Neurosurg Psychiatry* 2005;76:1528–33 CrossRef Medline
- Harris PA, Taylor R, Minor BL, et al; REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform* 2019;95:103208 CrossRef Medline
- Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap): a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81 CrossRef Medline
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74 CrossRef Medline
- Raymond J, Darsaut TE, Altman DG. Pragmatic trials can be designed as optimal medical care: principles and methods of care trials. *J Clin Epidemiol* 2014;67:1150–56 CrossRef Medline
- Raymond J, Fahed R, Darsaut TE. Randomize the first patient. *J Neuroradiol* 2017;44:291–94 CrossRef Medline



# A Clinical and Imaging Fused Deep Learning Model Matches Expert Clinician Prediction of 90-Day Stroke Outcomes

Yongkai Liu, Preya Shah, Yannan Yu, Jai Horsey, Jiahong Ouyang, Bin Jiang, Guang Yang, Jeremy J. Heit, Margy E. McCullough-Hicks, Stephen M. Hugdal, Max Wintermark, Patrik Michel, David S. Liebeskind, Maarten G. Lansberg, Gregory W. Albers, and Greg Zaharchuk



## ABSTRACT

**BACKGROUND AND PURPOSE:** Predicting long-term clinical outcome in acute ischemic stroke is beneficial for prognosis, clinical trial design, resource management, and patient expectations. This study used a deep learning–based predictive model (DLPD) to predict 90-day mRS outcomes and compared its predictions with those made by physicians.

**MATERIALS AND METHODS:** A previously developed DLPD that incorporated DWI and clinical data from the acute period was used to predict 90-day mRS outcomes in 80 consecutive patients with acute ischemic stroke from a single-center registry. We assessed the predictions of the model alongside those of 5 physicians (2 stroke neurologists and 3 neuroradiologists provided with the same imaging and clinical information). The primary analysis was the agreement between the ordinal mRS predictions of the model or physician and the ground truth using the Gwet Agreement Coefficient. We also evaluated the ability to identify unfavorable outcomes (mRS >2) using the area under the curve, sensitivity, and specificity. Noninferiority analyses were undertaken using limits of 0.1 for the Gwet Agreement Coefficient and 0.05 for the area under the curve analysis. The accuracy of prediction was also assessed using the mean absolute error for prediction, percentage of predictions  $\pm 1$  categories away from the ground truth ( $\pm 1$  accuracy [ACC]), and percentage of exact predictions (ACC).

**RESULTS:** To predict the specific mRS score, the DLPD yielded a Gwet Agreement Coefficient score of 0.79 (95% CI, 0.71–0.86), surpassing the physicians' score of 0.76 (95% CI, 0.67–0.84), and was noninferior to the readers ( $P < .001$ ). For identifying unfavorable outcome, the model achieved an area under the curve of 0.81 (95% CI, 0.72–0.89), again noninferior to the readers' area under the curve of 0.79 (95% CI, 0.69–0.87) ( $P < .005$ ). The mean absolute error,  $\pm 1$ ACC, and ACC were 0.89, 81%, and 36% for the DLPD.

**CONCLUSIONS:** A deep learning method using acute clinical and imaging data for long-term functional outcome prediction in patients with acute ischemic stroke, the DLPD, was noninferior to that of clinical readers.

**ABBREVIATIONS:** AC = Agreement Coefficient; ACC = accuracy;  $\pm 1$ ACC = mRS accuracy within  $\pm 1$  score; AIS = acute ischemic stroke; AUC = area under the curve; DL = deep learning; DLPD = deep learning–based predictive model; IQR = interquartile range; MAE = mean absolute error; ROC = receiver operating characteristic

Stroke affects nearly 800,000 people annually in the United States and is a major global cause of disability and mortality.<sup>1</sup> Survivors often face significant functional impairment that impacts their quality of life.<sup>2</sup> Predicting long-term clinical impairment from early-stage information in acute ischemic strokes (AIS) is crucial for enhancing rehabilitation strategies

and informing clinical trial designs, resource allocation, and patient expectations.<sup>1–3</sup> However, prediction is complex due to the many factors influencing a patient's eventual disability level and the known weak correlation between initial infarct size and outcome.<sup>4–7</sup>

Although some studies<sup>1,4,8</sup> have attempted to predict long-term functional outcomes, these traditional methodologies have suboptimal performance due to 2 primary factors: their reliance on manually crafted imaging features, which may not be optimal predictors, and the subjective inclusion of clinical measurements,

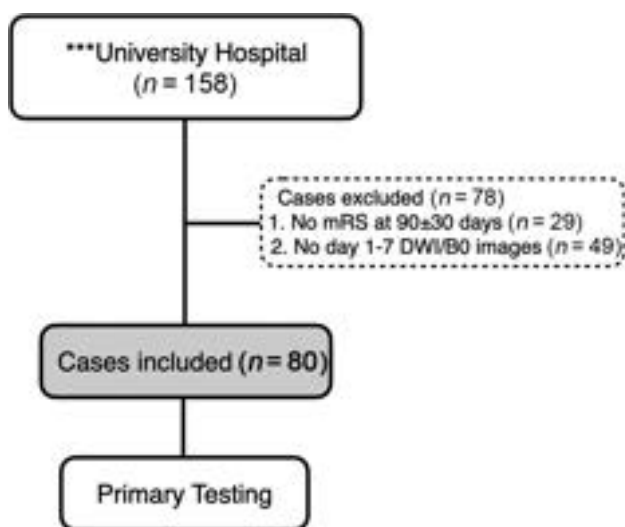
Received October 27, 2023; accepted after revision December 7.

From the Departments of Radiology (Y.L., P.S., Y.Y., J.O., B.J., J.J.H., S.M.H., G.Z.), Electrical Engineering (J.O.), and Neurology (M.G.L., G.W.A.), Stanford University, Stanford, California; Meharry Medical College (J.H.), Nashville, Tennessee; National Heart and Lung Institute (G.Y.), Imperial College London, London, UK; Department of Neurology (M.E.M.-H.), University of Minnesota Medical School, Minneapolis, Minnesota; Department of Neuroradiology (M.W.), University of Texas MD Anderson Center, Houston, Texas; Neurology Service (P.M.), Department of Clinical Neurosciences, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland; and Department of Neurology (D.S.L.), University of California, Los Angeles, Los Angeles, California.

Y. Liu and P. Shah contributed equally to this work.

This work was supported by the US Department of Health and Human Services, National Institutes of Health, National Institute of Neurological Disorders and Stroke, R01-NS130172, R01-NS066506.

Please address correspondence to Yongkai Liu, PhD, Department of Radiology, 1201 Welch Rd., Stanford, CA 94305-5488; e-mail: yongkai@stanford.edu; @Focus\_on\_aca  
<http://dx.doi.org/10.3174/ajnr.A8140>



**FIG 1.** Flow chart for patients in the current study.

some of which may not be available in nonspecialist centers. Selecting and extracting imaging features further compounds these issues, adding yet more subjectivity and often requiring resource-intensive manual postprocessing. In recent years, deep learning (DL), particularly convolutional neural networks, has shown promise in enhancing medical imaging diagnostics and prognostics by adaptively learning from raw images.<sup>9,10</sup> Few existing studies have used DL to discern optimal features from medical imaging for the prediction of long-term disability, especially for the task of predicting the patient's exact score on the 90-day mRS.

The alternative to automated systems might be the predictions of expert physicians, who have significant real-world experience correlating imaging and clinical results with eventual outcomes and may be required to make judgments that can impact individual patients. However, there are few systematic evaluations of prediction of stroke outcome by humans.<sup>11</sup> Benchmarking against clinical readers provides a meaningful context for measuring improvement and may serve as a reference for future research and clinical evaluations. This study aimed to evaluate a previously developed deep learning-based predictive model (DLPD), which uses DWI and clinical variables to predict 90-day clinical outcomes in patients with AIS, using a prospective registry from a comprehensive stroke center. We compared its performance with that of clinical readers, including neurologists and neuroradiologists.

## MATERIALS AND METHODS

### **Patients and MR Imaging Data Sets**

This study adhered to the guidelines set forth by the US Health Insurance Portability and Accountability Act of 1996 and received approval from the institutional review board (Stanford University). Under the institutional review board guidelines, we either obtained written informed consent from all subjects or the requirement for consent was waived. The initial study population included 158 patients with AIS from a single registry who were randomly selected from a database of large-vessel occlusion

candidates undergoing triage for possible thrombectomy between 2010 and 2019. To assess long-term clinical outcomes, we used the mRS,<sup>12,13</sup> a grading system for measuring disability levels ranging from 0 (no disability) to 6 (death). Inclusion criteria were the acquisition of MR images with DWI acquired between days 1 and 7 after the index event and after all acute therapies were complete and the 90-day mRS evaluation. Routinely collected clinical parameters included the following: age, sex, premorbid mRS, presenting and 24-hour NIHSS, and a history of hypertension or diabetes. Figure 1 provides a detailed flow chart outlining the study subjects.

### **DLPD**

The DLPD model uses DWI and the previously mentioned clinical variables as input to predict 90-day mRS outcome. It was based on a previously developed and validated model that was trained using data from 861 patients across multiple institutions.<sup>14</sup> In brief, it is a fused model that takes DWI and B0 images as input to define deep features relevant to mRS prediction and then fuses these with a separate predictive support-vector machine model using the clinical variables. Part of the nonsensitive code is available at outcome prediction.

### **Physician mRS Prediction**

Outcome prediction using the mRS scale was independently performed by 5 physicians who were given the same information as in the DLPD model. These included 2 neuroradiologists (one being a neurointerventional radiologist) with 13 and 22 years of experience, respectively; 2 stroke neurologists with 8 and 38 years of experience, respectively; and a neuroradiology fellow with 5 years of experience. To compare against the DLPD, we used the consensus mRS prediction using the median score from all the physicians, given the high agreement between them. The consensus mRS score for the physicians was created using a majority score when present (ie, the same score in  $\geq 3$  readers); otherwise, the median score was used.

### **Statistical Analysis**

The performance in predicting ordinal mRS outcomes was evaluated using several metrics: the Gwet Agreement Coefficient (AC),<sup>15</sup> mean absolute error (MAE), mRS accuracy within  $\pm 1$  score ( $\pm 1$ ACC), and accuracy (ACC). The Gwet AC, applied with ordinal weighting, quantifies the level of agreement of predictions of both the DLPD and clinicians with the ground truth. The MAE assesses the average absolute difference between the predicted scores and the actual 90-day mRS scores, with a smaller MAE indicating superior performance.  $\pm 1$ ACC evaluates the proportion of predictions that fall within 1 mRS category of the actual score. ACC measures the proportion of predictions that precisely match the actual score. For each of these metrics, the noninferiority of the DLPD compared with the consensus of clinical readers was determined using a predefined margin of 0.1 (MAE) or 10% ( $\pm 1$ ACC, ACC). Additionally, the area under the curve (AUC), sensitivity, and specificity were measured to evaluate the predictive accuracy of the model for unfavorable outcome (mRS  $> 2$ ), with a predefined noninferiority margin of 0.05. Analyses for ordinal outcome prediction were performed using

Stata 17.0 (StataCorp),<sup>16</sup> while analyses for unfavorable outcome prediction were performed using Python 3.9.12.

## RESULTS

### Patient Characteristics

From an initial pool of 158 patients sourced from the stroke registry, a total of 80 patients, median age of 62 years (interquartile range [IQR]: 51–75 years), including 44 men (55%), met the inclusion criteria of the study and were subsequently included in its testing set. Details about the included cohort can be found in Tables 1 and 2.

### Performance of Ordinal mRS Prediction

Table 3 compares the DLPD model and the consensus of clinical readers across multiple metrics. The DLPD model had improved

**Table 1: Summary of the characteristics of patients with AIS included in the Stanford University Hospital cohort (n = 80)<sup>a</sup>**

Summary	
Characteristics	
Male	44 (55.0)
Age (yr) (median) (IQR)	62 (51–75)
History of hypertension	53 (66.3)
History of diabetes	20 (25.0)
Baseline NIHSS (median) (IQR)	12 (7–17)
24-Hour NIHSS	9 (4–17)
	3.8% <sup>b</sup>
Days after stroke for MR imaging (median) (IQR)	1 (1–3)
90-Day outcome	
Favorable outcome (90-Day mRS ≤ 2)	36 (45.0)
Unfavorable outcome (90-Day mRS > 2)	44 (55.0)

<sup>a</sup> Unless otherwise mentioned, data are expressed as number (percentage) of patients.

<sup>b</sup> Percentage of variables missing. If no data are missing, then there will be no percentage reported.

**Table 2: MRS score<sup>a</sup>**

Scale	Pre morbid mRS	90-Day mRS
0	67 (83.8)	6 (7.5)
1	6 (7.5)	17 (21.3)
2	3 (3.8)	13 (16.3)
3	4 (5.0)	19 (23.8)
4	0 (0.0)	12 (15.0)
5	0 (0.0)	10 (12.5)
6	0 (0.0)	3 (3.8)

<sup>a</sup> Data are expressed as number (percentage) of patients.

**Table 3: Performance comparisons for ordinal mRS prediction between the DLPD and the clinical readers<sup>a</sup>**

	Gwet AC	MAE	±1ACC (%)	ACC (%)
Model/readers				
Neuroradiologist I	0.70 (0.60–0.80)	1.15 (0.94–1.38)	71 (60–81)	26 (18–36)
Neuroradiologist II	0.69 (0.59–0.79)	1.14 (0.93–1.38)	70 (60–80)	29 (19–39)
Neuroradiology fellow	0.73 (0.65–0.81)	1.04 (0.85–1.24)	74 (6–84)	31 (21–41)
Clinical readers				
Neurologist I	0.75 (0.66–0.84)	1.03 (0.83–1.25)	75 (65–84)	32 (22–44)
Neurologist II	0.77 (0.67–0.86)	0.91 (0.70–1.15)	79 (69–88)	41 (30–51)
Consensus read	0.76 (0.67–0.84)	0.95 (0.75–1.17)	79 (70–88)	36 (25–46)
DLPD	0.79 (0.71–0.86)	0.89 (0.70–1.11)	81 (73–90)	36 (26–46)
P value	P < .001	P = .02	P < .001	P = .07

<sup>a</sup> The data in the parentheses represent the 95% confidence interval. The P value is for the noninferiority test between the consensus clinical reads and the DLPD with the predefined margin of 0.1 (MAE)/10% (±1ACC, ACC). The Gwet AC for agreement among 5 clinical readers is 0.83 (95% CI, 0.80–0.86), justifying the comparison with a consensus.

values in all evaluated metrics, achieving a Gwet AC of 0.79 (95% CI, 0.71–0.86), an MAE of 0.89 (95% CI, 0.70–1.11), ±1ACC of 81% (95% CI, 73%–90%), and an ACC of 36% (95% CI, 26%–46%). The level of agreement among the 5 clinical readers, as gauged by a strong Gwet AC of 0.83 (95% CI, 0.80–0.86), affirms the consistency in their judgments and justifies the use of a clinical consensus score to compare with the DLPD. The clinical consensus score achieved a Gwet AC of 0.76 (95% CI, 0.67–0.84), MAE of 0.95 (95% CI, 0.75–1.17), ±1ACC of 79% (95% CI, 70%–88%), and ACC of 36% (95% CI, 25%–46%). Noninferiority tests confirmed that the performance of the DLPD model was noninferior to that of the clinicians across all evaluated metrics, except for ACC. The significant P values were P < .001 for the Gwet AC, P = .02 for MAE, and P < .001 for ±1ACC, while the P value for ACC was not significant (P = .07). Figure 2 presents 3 illustrative examples of outcome predictions made by the DLPD and physicians.

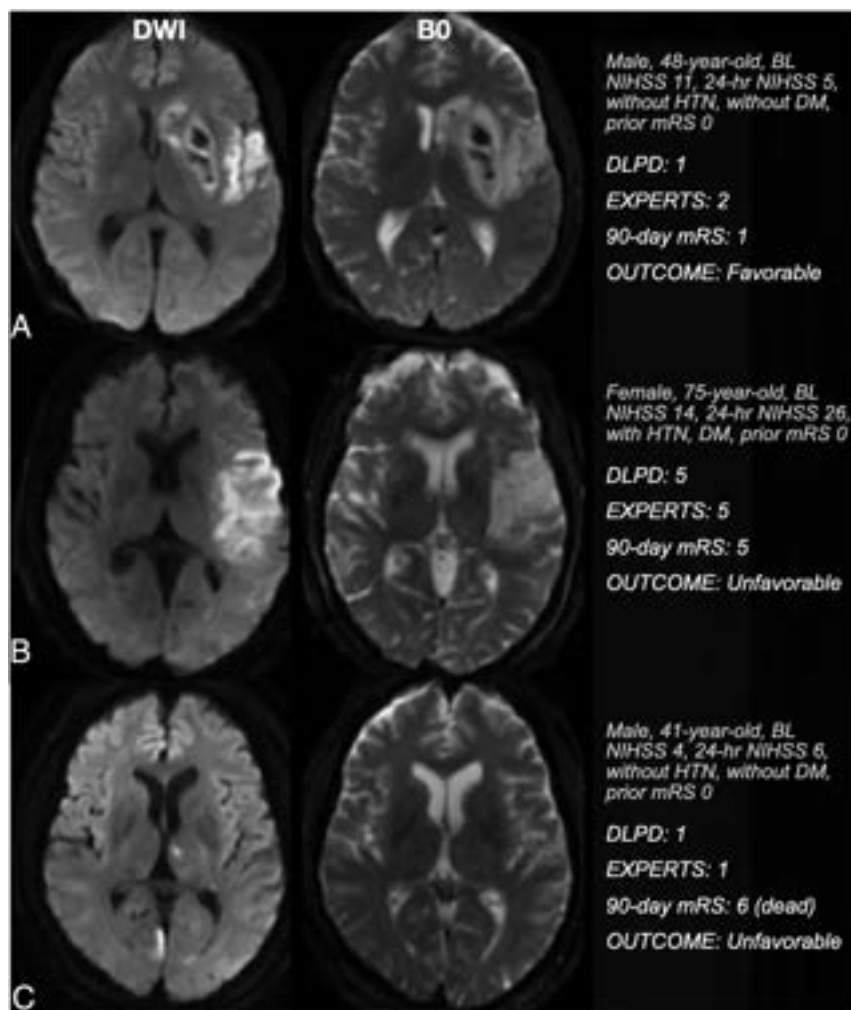
### Predicting Unfavorable Outcome

Table 4 compares the performance of the DLPD and physicians to predict unfavorable outcome (mRS > 2). The DLPD model surpassed the readers by achieving an AUC of 0.81 (95% CI, 0.72–0.89), compared with the physicians' AUC of 0.79 (95% CI, 0.69–0.87). The model was again noninferior to the physicians for this task (P = .005). The DLPD model had a higher specificity of 0.81 (95% CI, 0.67–0.92), which was noninferior to the clinical consensus specificity of 0.75 (95% CI, 0.60–0.88) (P = .03). The sensitivity of the DLPD model (0.68 [95% CI, 0.54–0.81]) was lower and did not satisfy the noninferiority margin compared with the clinical consensus (0.70 [95% CI, 0.56–0.83]). Figure 3 shows the receiver operating characteristic (ROC) of the DLPD together with the data points representing the individual and consensus physicians. The physicians' operating points are located just beneath the ROC curve of the DLPD, suggesting that for the same level of specificity or sensitivity, the DL model generally achieves slightly better performance.

## DISCUSSION

This study demonstrates that a clinical and imaging fused DL model is noninferior to expert physicians in predicting specific mRS outcomes and unfavorable prognoses. Building on our prior

work—which established a robust methodology across multiple institutions and demonstrated consistent performance in 2 distinct cohorts—this work not only further enhances the generalizability of the model but also provides critical benchmarks against human expert performance for the task at hand. By evaluating our methodology alongside clinical expert judgments within a unique patient cohort, we investigated the practical implications of our approach in a real-world clinical setting, an element not addressed in our prior work. Slightly better performance was observed in this cohort



**FIG 2.** MR images (the first and second columns represent DWI and B0 images, respectively) for 3 patients with diverse clinical histories and 90-day mRS scores. Patient A is a 48-year-old man with a baseline NIHSS of 11, 24-hour NIHSS of 5, and a 90-day mRS of 1. He has no medical history of either diabetes or hypertension. The DL model accurately predicted his score. However, the readers overestimated his score by 1 point. Patient B, a 75-year-old woman, has a medical history that includes diabetes and hypertension and a 90-day mRS of 5. Both the DL model and the readers accurately predicted her 90-day mRS score of 5. Patient C, a 41-year-old man with no history of diabetes or hypertension, has a 90-day mRS score of 6. However, both the DL model and the readers incorrectly predicted his 90-day mRS score as 1. HTN indicates hypertension; DM, diabetes mellitus.

**Table 4: Performance comparisons for unfavorable-outcome prediction (mRS > 2) between the DLPD and the clinical readers<sup>a</sup>**

	AUC	Sensitivity	Specificity
Model/readers			
Neuroradiologist I	0.76 (0.64–0.85)	0.68 (0.53–0.81)	0.69 (0.54–0.83)
Neuroradiologist II	0.81 (0.71–0.89)	0.61 (0.47–0.77)	0.86 (0.73–0.97)
Neuroradiology fellow	0.82 (0.72–0.89)	0.75 (0.62–0.87)	0.69 (0.53–0.83)
Clinical readers			
Neurologist I	0.79 (0.69–0.88)	0.82 (0.7–0.93)	0.53 (0.36–0.68)
Neurologist II	0.77 (0.67–0.86)	0.64 (0.49–0.78)	0.83 (0.7–0.95)
Consensus read	0.79 (0.68–0.87)	0.70 (0.56–0.83)	0.75 (0.60–0.88)
DLPD	0.81 (0.72–0.89)	0.68 (0.54–0.81)	0.81 (0.67–0.92)
P value	P = .005	P = .25	P = .03

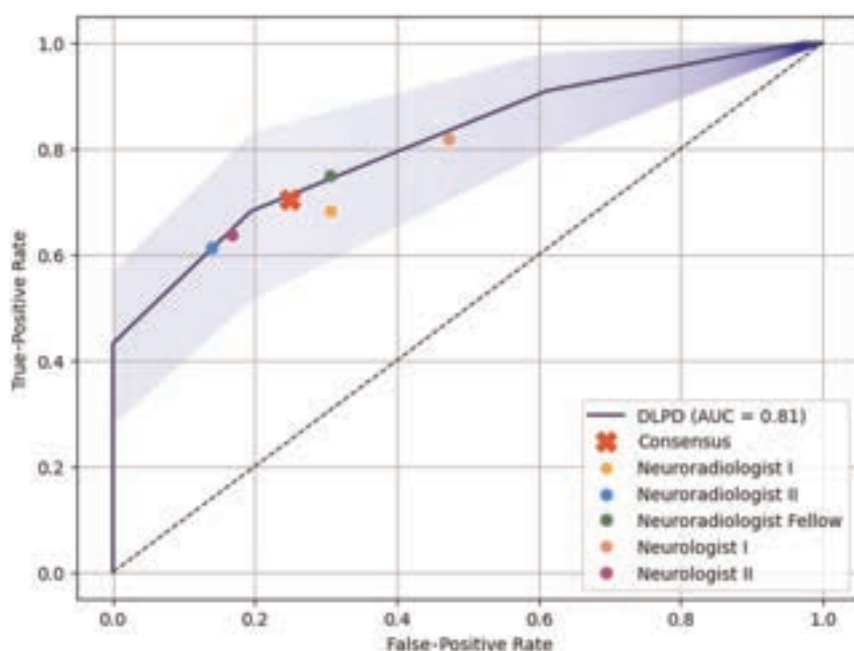
<sup>a</sup> The data in the parentheses represent the 95% confidence interval. The P value is for the noninferiority test between the consensus clinical read and the DLPD (predefined margin, .05).

compared with the prior work, which may be attributed to several factors. These include the following: 1) natural variations in patient demographics and disease presentations across cohorts; and 2) the single-institution cohort of the current study likely offering more uniform treatment protocols, imaging techniques, and patient management strategies, unlike the multi-institution data sets of our prior work, which presented greater variability. Additionally, this work uniquely applies DL to predict stroke outcomes, achieving an ordinal mRS accuracy rate of 36%, nearly triple the rate of random guessing. Prior studies tackling this task typically had lower accuracy using methods such as linear regression and random forest machine learning with hand-crafted features.<sup>1,8</sup>

The DLPD also provides practical advantages. By eliminating the need for specialized neurologic expertise, it may be useful for facilities that lack immediate access to neurologists or neuroradiologists. Consequently, it broadens the scope of quality care by making sophisticated prognostic information more accessible across diverse health care environments. The DLPD model uses readily accessible imaging and clinical variables and can be easily integrated into the current clinical workflow for predicting 90-day mRS. This model requires minimal preprocessing steps, with the primary requirement being the normalization of DWI and B0 images to a standard template. Unlike approaches that depend on existing radiologic features—potentially introducing added complexity, human effort, and subjectivity—our method automatically uses information from imaging, thereby lending greater objectivity to our outcome-prediction model. For example, a recent multivariate ordinal mRS regression model required inclusion of 19 separate variables, including some derived from imaging, for which interobserver reproducibility has not been reported.<sup>17</sup> In contrast, the current model relies heavily on objective, data-derived imaging features, with only 7 standard clinical measurements.

This study represents the first report of clinical expert performance for 90-day mRS prediction and allows us to





**FIG 3.** The AUC of the DL-based predictive model for predicting unfavorable outcomes (mRS > 2) is shown alongside data points representing the performance of individual clinicians and the consensus of clinicians. The translucent blue region denotes the 95% confidence interval for the ROC curve, constructed using bootstrapping.

conduct a comparative analysis with the automated model. Thus, the results can act as a benchmark for future studies and further contextualize the results of the DLPD. One prior study<sup>11</sup> collected outcome-prediction data from treating providers before endovascular therapy and showed relatively poor performance (44% accuracy to predict into mRS bins of 0–2, 3–4, and 5–6). The performance of the readers in the current study was better, probably because they made their assessments after treatment was provided. Also, these authors stressed that the premorbid mRS was an important predictive feature, but it was often unavailable or inaccurately estimated compared with later retrospective assessment. This issue emphasizes the value of using the entire image, which can incorporate both acute and pre-existing lesions to improve prediction.

Potential reasons behind the noninferior performance of the DLPD in outcome prediction are manifold. First, the DLPD may identify and learn from patterns within complex, multimodal data, similar to or better than how physicians apply their medical knowledge and experience.<sup>18,19</sup> It emulates or improves on human readers by evaluating imaging and clinical data in a data-driven manner, considering all available information, from obvious clinical signs to subtle imaging hints. Additionally, these models can discern and understand nonlinear relationships and interactions among numerous variables, mirroring physicians' multidimensional thinking when assessing patient conditions and outcomes. Furthermore, DLPD models may bring additional advantages. Their inherent ability to process and learn from vast amounts of data offers unprecedented scalability. With the growth in available data, the performance of the model can potentially increase, indicating a cycle of

continuous advancement that may be challenging to match solely with human expertise.

Our study has the following limitations. First, our patient testing cohort was sourced from a single registry. While this feature mimics how the tool might be used in real practice, performance in other cohorts with different characteristics of severity and age is difficult to assess. However, the model has been previously applied to 2 other clinical cohorts with diverse levels of severity and demonstrated similar performance.<sup>14</sup> Second, the mRS served as our primary outcome measure. While the 90-day mRS is widely used to assess chronic disability severity, its subjective determination of categories and variability in reproducibility among different examiners presents notable challenges. Third, the imaging data used in our study were obtained at least 24 hours after the initial baseline imaging; this timeframe was chosen to minimize the effects of any acute interventions, which

were completed at the time of imaging. Future studies could consider using initial, pretreatment imaging combined with different therapies to predict outcomes, potentially informing treatment decision-making. Fourth, including imaging sequences beyond DWI and B0 could yield more detailed insight, though this would require more resources and image postprocessing. Fifth, despite the small number of patients in our study, we want to emphasize several factors: Our evaluation was conducted by 5 clinical readers, ensuring a thorough and nuanced assessment, being particularly noteworthy given the complex nature of the reader study and the demanding schedules of the clinicians.

Reader studies are inherently time-intensive, and coordinating such effort among 5 busy clinicians poses substantial challenges. Nonetheless, our findings demonstrate that the performance of the DL model aligns with that of their clinical evaluations, underscoring its potential for clinical application even within a limited patient cohort. This agreement reinforces our hypothesis that the model can operate at a level comparable with that of humans. Last, it is critical to recognize that outcomes may diverge significantly from predictions due to the multifaceted interplay of medical conditions, social determinants, and systemic health care factors that are not entirely predictable by our algorithms. While our model demonstrates robustness, it is not configured to anticipate every acute medical event or the full range of sociodemographic variables that may substantially affect the clinical course. Therefore, there is a clear need to continuously refine predictive methodologies, possibly incorporating a wider set of variables that capture the complexities of patient trajectories; although again, this suggestion comes with drawbacks related to the complexity of the models and the need to collect information that may be difficult to obtain.



## CONCLUSIONS

We demonstrated that a DLPD model that leverages brain MR imaging and routinely obtained clinical information to predict long-term outcomes in patients with AIS generalizes well to another clinical cohort. We have further provided a benchmark of human expert performance on this task and show that the DLPD model is noninferior to predictions made by neuroradiologists and stroke neurologists.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

1. Xie Y, Jiang B, Gong E, et al. **Use of gradient boosting machine learning to predict patient outcome in acute ischemic stroke on the basis of imaging, demographic, and clinical information.** *AJR Am J Roentgenol* 2019;212:44–51 CrossRef Medline
2. Nichols-Larsen DS, Clark PC, Zeringue A, et al. **Factors influencing stroke survivors' quality of life during subacute recovery.** *Stroke* 2005;36:1480–84 CrossRef Medline
3. Langhorne P, Bernhardt J, Kwakkel G. **Stroke rehabilitation.** *Lancet* 2011;377:1693–702 CrossRef Medline
4. Heo J, Yoon JG, Park H, et al. **Machine learning-based model for prediction of outcomes in acute stroke.** *Stroke* 2019;50:1263–65 CrossRef Medline
5. Brugnara G, Neuberger U, Mahmutoglu MA, et al. **Multimodal predictive modeling of endovascular treatment outcome for acute ischemic stroke using machine-learning.** *Stroke* 2020;51:3541–51 CrossRef Medline
6. Ospel JM, Hill MD, Menon BK, et al; ESCAPE-NA1 Investigators. **Strength of association between infarct volume and clinical outcome depends on the magnitude of infarct size: results from the ESCAPE-NA1 trial.** *AJNR Am J Neuroradiol* 2021;42:1375–79 CrossRef Medline
7. Albers GW, Marks MP, Kemp S, et al; DEFUSE 3 Investigators. **Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging.** *N Engl J Med* 2018;378:708–18 CrossRef Medline
8. Zhang MY, Mlynash M, Sainani KL, et al. **Ordinal prediction model of 90-day modified Rankin scale in ischemic stroke.** *Front Neurol* 2021;12:727171 CrossRef Medline
9. Shen D, Wu G, Suk HI. **Deep learning in medical image analysis.** *Annu Rev Biomed Eng* 2017;19:221–48 CrossRef Medline
10. Liu Y, Zheng H, Liang Z, et al. **Textured-based deep learning in prostate cancer classification with 3t multiparametric MRI: comparison with PI-RADS-based classification.** *Diagnostics (Basel)* 2021;11:1785 CrossRef Medline
11. Fargen KM, Kittel C, Curry BP, et al; Satin Research Group. **Mechanical thrombectomy decision making and prognostication: Stroke treatment Assessments prior to Thrombectomy In Neurointervention (SATIN) study.** *J Neurointerv Surg* 2023;15:e381–87 CrossRef Medline
12. Banks JL, Marotta CA. **Outcomes validity and reliability of the modified Rankin scale: implications for stroke clinical trials: a literature review and synthesis.** *Stroke* 2007;38:1091–96 CrossRef Medline
13. Broderick JP, Adeoye O, Elm J. **Evolution of the modified Rankin scale and its use in future stroke trials.** *Stroke* 2017;48:2007–12 CrossRef Medline
14. Liu Y, Yu Y, Ouyang J, et al. **Functional outcome prediction in acute ischemic stroke using a fused imaging and clinical deep learning model.** *Stroke* 2023;54:2316–27 CrossRef Medline
15. Gwet KL. **Computing inter-rater reliability and its variance in the presence of high agreement.** *Br J Math Stat Psychol* 2008;61:29–48 CrossRef Medline
16. StataCorp. *Stata Statistical Software: Release 15.* College Station, Texas: StataCorp LLC. 2017.
17. Chalos V, Venema E, Mulder MJ, et al; HERMES Collaborators; MR CLEAN Registry Investigators. **Development and validation of a post-procedural model to predict outcome after endovascular treatment for ischemic stroke.** *JAMA Neurol* 2023;80:940–48 CrossRef Medline
18. Kamnitsas K, Ledig C, Newcombe VF, et al. **Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation.** *Med Image Anal* 2017;36:61–78 CrossRef Medline
19. Esteva A, Kuprel B, Novoa RA, et al. **Dermatologist-level classification of skin cancer with deep neural networks.** *Nature* 2017;542:115–18 CrossRef Medline

# Predictors of In-Stent Stenosis Following the Implantation of Pipeline Embolization Devices for the Treatment of Aneurysms Located at or beyond the Circle of Willis in the Anterior Circulation

Yajing Ma, Xin Deng, Junfan Chen, Feng Fan, Kaihao Han, Sheng Guan, and Xinbin Guo

## ABSTRACT

**BACKGROUND AND PURPOSE:** In-stent stenosis is commonly observed after stent implantation. There is no consensus on the contributing factors for in-stent stenosis, especially for aneurysms located at or beyond the circle of Willis in the anterior circulation. This study aimed to investigate the morbidity and determinants of in-stent stenosis in distal anterior circulation aneurysms following the implantation of Pipeline Embolization Devices.

**MATERIALS AND METHODS:** Patients who underwent Pipeline Embolization Device treatment at our center between January 1, 2018, and June 15, 2023, were enrolled. Distal anterior circulation aneurysms were defined as those occurring at or beyond the circle of Willis, including anterior communicating artery aneurysms, anterior cerebral artery aneurysms, and MCA aneurysms. Baseline information, aneurysm characteristics, and follow-up data of patients were analyzed. Patients were divided into 2 groups: the in-stent stenosis group (patients with a loss of >25% of the lumen diameter of the parent artery) and the non-in-stent stenosis group. Binary logistic regression and restricted cubic spline curves were used to explore risk factors.

**RESULTS:** We included 85 cases of 1213 patients treated with flow-diverter devices at our hospital. During an average follow-up period of 9.07 months, the complete occlusion rate was 77.64%. The overall incidence of in-stent stenosis was 36.47% (31/85), of which moderate stenosis accounted for 9.41% (8/85), and severe stenosis, 5.88% (5/85) (triglyceride-glucose index  $\geq 8.95$ ; OR = 6.883,  $P = .006$ ). The difference in diameters between the stent and parent artery of  $\geq 0.09$  mm (OR = 6.534,  $P = .015$ ) and 55 years of age or older (OR = 3.507,  $P = .036$ ) were risk factors for in-stent stenosis. The restricted cubic spline curves indicated that the risk of in-stent stenosis increased as the difference in diameter between stent and parent artery and the triglyceride-glucose index increased.

**CONCLUSIONS:** Compared with the on-label use of Pipeline Embolization Devices, the rate of in-stent stenosis did not obviously increase when treating distal anterior circulation aneurysms with these devices. The incidence of in-stent stenosis was 36.47% when defined as a lumen diameter loss of >25%, and 15.2% when defined as a lumen diameter loss of >50%. Stent-size selection and biochemical indicators can potentially impact the incidence of in-stent stenosis.

**ABBREVIATIONS:** d = parent artery diameter; Dd = difference in diameter between stent and parent artery; ISS = in-stent stenosis; PED = Pipeline Embolization Device; TyG = triglyceride-glucose

Positive outcomes have been reported for Pipeline Embolization Device (PED; Medtronic) treatment for appropriately selected patients with intracranial aneurysms.<sup>1,2</sup> Recently, there have also been several studies investigating the safety and efficacy of the off-label uses of the PED, one of which is in the treatment of distal

circulation aneurysms,<sup>3-7</sup> referring to aneurysms occurring at the circle of Willis or beyond.

Distal circulating aneurysms are difficult to treat due to several factors, such as the small diameter of the parent artery (d), increased vascular curvature, and the presence of perforating arteries, which also increase the risk of clinical complications. The use of small devices ( $\leq 3$  mm in diameter) in the treatment of these aneurysms is safe and does not increase the incidence of in-stent stenosis (ISS) or other clinical complications.<sup>8</sup> However, Liou and Li<sup>9</sup> suggested that ISS may be more noteworthy in smaller PEDs because there is less room for stenosis to occur. Although most ISS stenoses are reported to be asymptomatic, studies have shown an increased incidence of long-term complications in patients with

Received October 26, 2023; accepted after revision December 11.

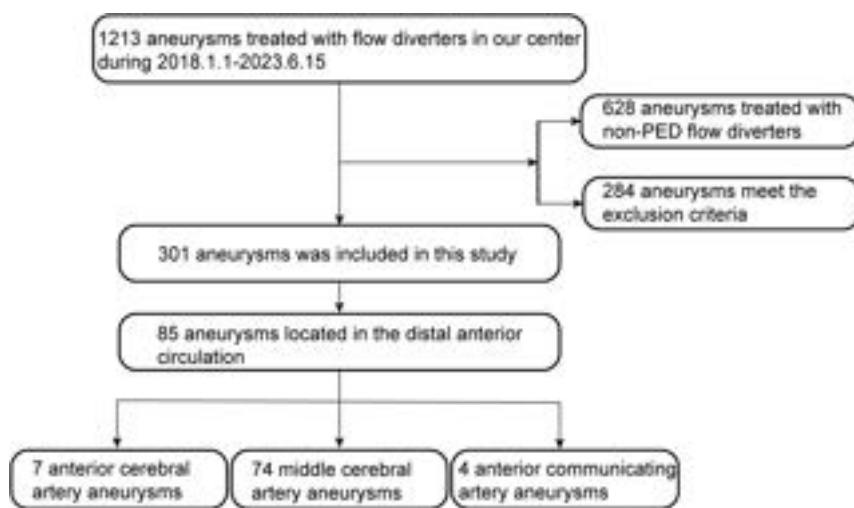
From the Department of Interventional Neuroradiology (Y.M., X.D., J.C., F.F., K.H., S.G., X.G.), The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan Province, China; and Neurointerventional Engineering Research Center of Henan Province (S.G.), Henan Province, China.

Please address correspondence to Xinbin Guo, PhD, The First Affiliated Hospital of Zhengzhou University; e-mail: gxb3906080@sina.com



Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8144>



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

ISS,<sup>10,11</sup> which can be fatal. Therefore, we aimed to perform a retrospective study to investigate the incidence of and factors associated with ISS following the treatment of distal anterior circulation aneurysms with PEDs to accurately identify high-risk patients and ensure timely interventions. This article follows the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) reporting guidelines.

## MATERIALS AND METHODS

### Study Design

This retrospective study included 85 aneurysms treated with a Pipeline Flex (Medtronic) at our center from January 1, 2018, to June 15, 2023. The exclusion criteria were as follows: 1) younger than 18 years of age; 2) acutely ruptured aneurysms; 3) inadequate information for analysis because of incomplete clinical or imaging data; and 4) patients who were noncompliant and did not adhere to medication instructions. More details are summarized in Fig 1. The study was approved by the Ethics Committee of Scientific Research and Clinical Trial of The First Affiliated Hospital of Zhengzhou University. All study participants provided informed consent.

### Data Collection and Definitions

The following information was collected for each participant: 1) baseline information (sex, age, body mass index, alcohol use, smoking status, blood pressure, major comorbidities, preoperative and postoperative mRS score, laboratory index, and so forth); 2) aneurysm-related information (location, neck, maximum diameter, proximal and distal parent artery diameters, the presence of multiple aneurysms, incorporation with a branch, and history of stent placement at the target site); and 3) surgical method, stent size, and follow-up imaging data, evaluated according to the O'Kelly-Marotta grading system.<sup>12</sup>

Distal anterior circulation aneurysms refer to aneurysms that occur at or beyond the circle of Willis, including aneurysms in the anterior communicating artery, anterior cerebral artery, and MCA. Venous blood samples were collected from participants after overnight fasting. The triglyceride-glucose (TyG) index was calculated as

follows: [fasting triglyceride (mg/dL)  $\times$  fasting glucose (mg/dL) / 2].<sup>13</sup> The parent artery diameter (d) was defined as the average diameter of the arteries on both sides of the aneurysm neck. The difference in diameter between the stent and parent artery (Dd) is the diameter of the stent. The maximum diameter was defined as the maximum distance from the midpoint of the neck to the dome of the aneurysm.

All measurements were performed independently by 2 neurointerventionalists with >5 years of experience on 3D DSA. If the difference between the 2 measurements was  $\leq 0.5$  mm, the average was used as the final data; in the case of measurement discrepancies of >0.5 mm, a third neurointerventionalist

with >10 years of endovascular treatment experience evaluated the case. A consensus was then reached by a panel of 3 physicians.

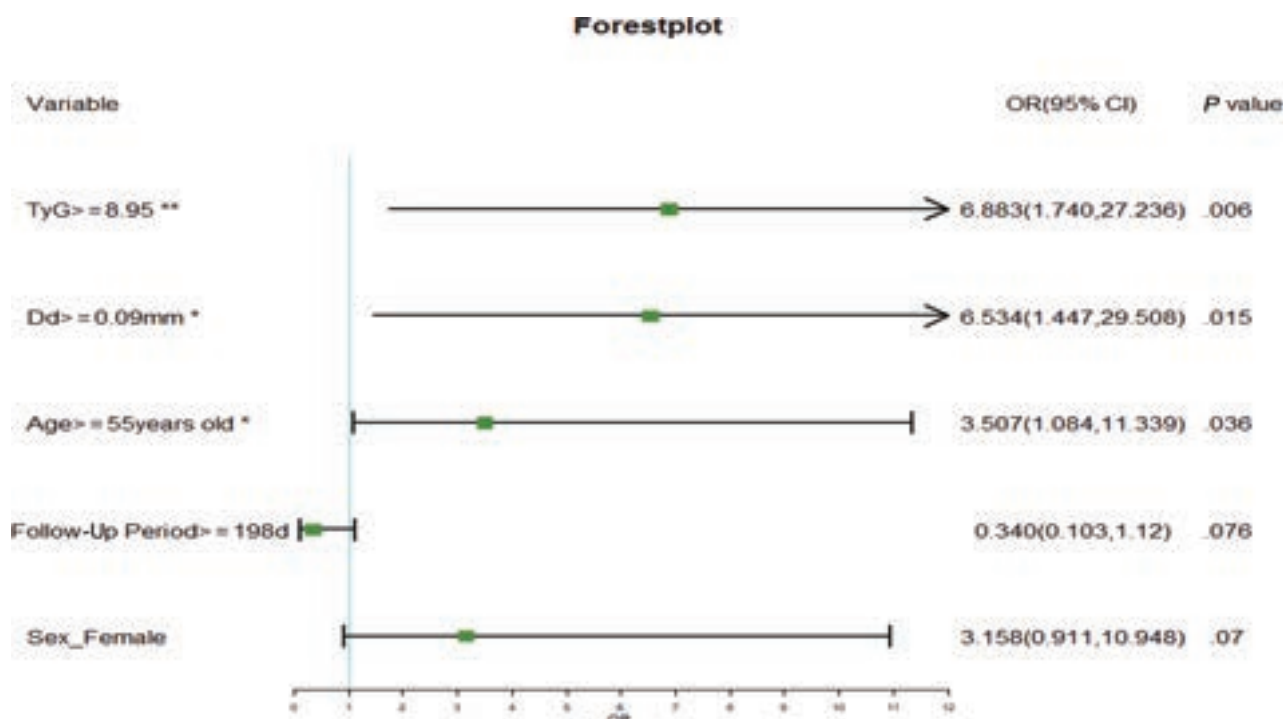
### Perioperative Management and Follow-Up

Patients with unruptured aneurysms routinely received 5–7 days of conventional dual antiplatelet therapy (aspirin, 100 mg once daily, + clopidogrel, 75 mg once daily) preoperatively. The effectiveness of the treatment was determined by thromboelastographic monitoring. For those who did not meet the standard of antiplatelet therapy, the dosage, adjusted according to the results of the genetic assessment; or aspirin, 100 mg once daily, + ticagrelor, 90 mg twice daily, was administered. Patients were monitored for ischemic/hemorrhagic complications, and the drug dosage was promptly adjusted as needed. They were also evaluated by thromboelastography before discharge, and if arachidonic acid and adenosine phosphate levels did not reach standard values (arachidonic acid >50% and adenosine phosphate >30%), clopidogrel was replaced with ticagrelor (90 mg/day, twice daily). Dual antiplatelet therapy was postoperatively administered for 6 months. Then, all patients were re-admitted to the hospital for a DSA examination; on the basis of these findings, a decision was made to transition to monotherapy.

Neurointerventionalists with >10 years of experience in treating aneurysms determined the stent selection. All patients were treated with Pipeline Flex stents. For achieving good apposition, the stent diameter was usually similar to the proximal d of the aneurysm. Balloon angioplasty was commonly used to correct severe stent malapposition and/or incomplete apposition.

### Outcomes

As the primary outcome, ISS was defined as >25% loss of vessel diameter at the stent deployment site or vessel occlusion within 5 mm on either side of the stent. Shrinkage was defined as mild if it was >25%, moderate if >50%, and severe if >75%. Regarding the secondary outcome, complete occlusion of the aneurysm was defined as grade D on the O'Kelly-Marotta scale.<sup>12</sup> Clinical outcomes were evaluated using mRS scores, and complications



**FIG 2.** Results of the multivariable logistic regression analysis shown by a forest plot. \* indicates  $P < .05$ ; \*\*,  $P < .01$ .

included hemorrhage, thrombosis, SAH, stroke, TIA, and access site-related complications.

### Statistical Analysis

Statistical analyses were performed using SPSS, Version 25 (IBM) and R statistical and computing software (Version 4.2.2; <http://www.r-project.org>). Continuous variables with normal distribution are presented as mean (SD), and non-normally distributed data, as median (interquartile range). Differences were compared using the  $\chi^2$  test for categorical variables and the Mann-Whitney  $U$  test for continuous variables.

According to the principle of the maximal Youden index, age, the TyG index, and Dd were converted into binary data. In addition, the d was divided into  $<2.5$  mm and  $\geq 2.5$  mm according to the definition of a small artery, and the body mass index was divided into  $<28$  and  $\geq 28$  kg/m<sup>2</sup> (Online Supplemental Data). The confounders included in the multivariable model were selected using a cutoff  $P$  value of  $<.1$  in univariate analysis. A multivariable logistic regression model (backward-stepwise) was used to analyze the association between confounders and ISS. Additionally, restricted cubic spline curves were performed to examine the associations among the TyG index, Dd, and ISS. The reference points for the TyG index and Dd were their corresponding medians. To balance the best fit and overfitting in restricted cubic spline, we used the Akaike information criterion to select the optimal number of knots. Two-tailed tests were performed, and a  $P$  value  $< .05$  was considered statistically significant.

## RESULTS

### Baseline Information

A total of 85 patients were enrolled, of whom 53 patients (62.3%) were female. The overall incidence of ISS was 36.47% (31/85), of

which moderate stenosis was 9.41% (8/85) and severe stenosis was 5.88% (5/85). Compared with the non-ISS group, the ISS group had a higher mean age (58.90 [SD, 2.06] years versus 54.59 [SD, 1.65] years,  $P = .111$ ), TyG index (8.65 [SD, 0.09] versus 8.41 [SD, .07],  $P = .052$ ), proportion of female patients (77.4% versus 53.7%,  $P = .03$ ), rate of stent malapposition (16.1% versus 1.9%,  $P = .042$ ), and a smaller parent artery diameter (2.37 mm [SD, 0.07] versus 2.58 mm [SD, 0.07],  $P = .049$ ) (Online Supplemental Data). The results of multivariable logistic regression analysis can be found in the Online Supplemental Data and Fig 2 (forest plot).

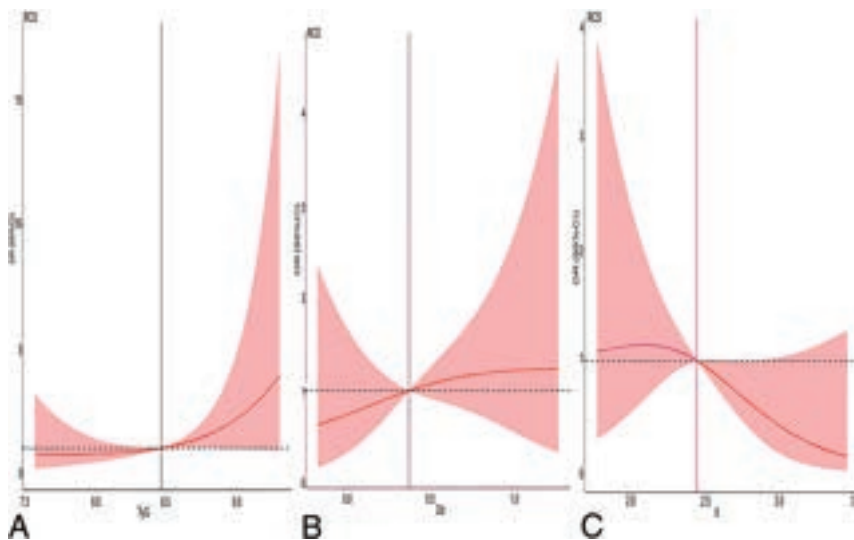
### Factors Associated with ISS

**TyG Index.** The risk of ISS in the group with a TyG index of  $\geq 8.95$  was 6.883 times higher than that in the group with a TyG index of  $<8.95$  ( $P = .006$ ). When the TyG index was treated as a continuous variable in a restricted cubic spline after adjusting for age, Dd, sex, and the follow-up period, the risk of ISS increased with the increase in TyG index (Fig 3A).

**Dd.** The risk of ISS in the  $Dd \geq 0.09$  mm group was 6.534 times higher than that in the  $Dd < 0.09$  mm group. After we adjusted for age, TyG index, sex, and follow-up period, the risk of ISS increased as Dd increased (Fig 3B).

**Diameter of the Parent Artery.** The rate of ISS in the  $d \geq 2.5$ -mm group was 21.05%, with 1 case (2.63%) presenting with moderate stenosis and no cases of severe stenosis. Meanwhile, the rate of ISS in the  $d < 2.5$ -mm group was 48.94%, with 7 cases (14.89%) of moderate stenosis and 5 (10.64%) of severe stenosis. According to the restricted cubic spline curve, the risk of ISS decreased as the d increased (Fig 3C).





**FIG 3.** A, Restricted cubic splines for the adjusted dose-response association of the TyG index for ISS. Data were fitted with a regression model using restricted cubic splines with 3 knots at the 10th, 50th, and 90th percentiles of the TyG index. The y-axis represents the OR, and the *shadow regions* are 95% confidence intervals.  $P$  nonlinear = .346,  $P$  overall = .0265; TyG = 8.47, OR = 1. B, Restricted cubic splines for the adjusted dose-response association of the Dd for ISS. Data were fitted with a logistic regression model using restricted cubic splines with 3 knots at the 10th, 50th, and 90th percentiles of the Dd. The y-axis represents the OR, and the *shadow regions* are 95% confidence intervals.  $P$  nonlinear = .697,  $P$  overall = .034; Dd = 0.38 mm, OR = 1. C, Restricted cubic splines for the dose-response association of the d for ISS. Data were fitted with a logistic regression model using restricted cubic splines with 3 knots at the 10th, 50th, and 90th percentiles of the d. The y-axis represents the OR, and the *shadow regions* are 95% confidence intervals.  $P$  nonlinear = .298,  $P$  overall = .149; d = 2.45 mm, OR = 1. The restricted cubic splines indicate that the risk of ISS increases with increasing TyG index and Dd, and decreases with increasing parent artery diameter.

**Age.** The risk of ISS was 3.507 times higher in patients 55 years of age or older than in those younger than 55 years of age (46.67% versus 25%,  $P = .036$ ). The rates of coexisting parent artery stenosis, hypertension, diabetes, and cardiovascular disease in the 55 years of age or older group were 13.3%, 77.78%, 17.78%, and 11.1%, respectively, whereas the rates in the younger than 55 years of age group were 7.5%, 40%, 2.5%, and 0%, respectively. The rates of balloon use were 4.44% and 2.5% in the 55 years of age and older and the younger than 55 years of age groups, respectively.

**Follow-Up Period.** The median follow-up period was shorter in the ISS group than in the non-ISS group (197 days versus 238.5 days,  $P = .208$ ). The percentage of patients with a follow-up period of  $\geq 198$  days was 48.4% in the ISS group and 68.5% in the non-ISS group ( $P = .067$ ). In multivariate logistic regression, there was no significant difference in the follow-up period between the 2 groups ( $P = .076$ ).

**Others.** Female patients accounted for 62.35% (53/85) of the cohort and had a higher incidence of ISS (45.28% versus 21.88%); however, there was no significant difference in multivariate logistic regression analysis ( $P = .070$ ). We also analyzed variables such as balloon use, stent length, difference in length between the stent and aneurysm neck size, aneurysm morphology, d, parent artery stenosis ( $>50\%$ ), and smoking status, but no significant differences were found.

## DISCUSSION

We included 85 cases of 1213 patients treated in our hospital, with an average follow-up of 9.07 months and a complete occlusion rate of 77.64%. The overall incidence of ISS was 36.47%. The TyG index of  $\geq 8.95$ , a Dd of  $\geq 0.09$  mm, and 55 years of age or older were risk factors for ISS. The restricted cubic spline curves indicated that the risk of ISS increased as the Dd and TyG index increased.

As the indications for PEDs continue to expand, neurointerventionalists worldwide are attempting to use these stents in the treatment of distal anterior circulation aneurysms. Compared with ICA aneurysms, distal anterior circulation aneurysms are usually irregularly shaped, often at bifurcations, and may be incorporated with important branch arteries. The parent arteries are tortuous and relatively small in diameter, making treatment difficult and complications more likely. There is significant heterogeneity in the reported incidence of ISS due to the differences in factors such as grading criteria, aneurysm location and size, and patient baseline information across studies. The Pipeline for Uncoilable or

Failed Aneurysms (PUFS) study considered any degree of lumen diameter loss within the stent as stenosis and found an ISS rate of 100%;<sup>2</sup> the corresponding probability was 6.6% when defining  $>25\%$  lumen diameter loss as ISS. However, all included aneurysms were in the ICA. John et al<sup>14</sup> reported an incidence of ISS of 9.8% with a mean follow-up of 12.5 months, with only 3 cases of distal aneurysms. In a large study conducted in China, the rate of ISS was found to be 10.03%,<sup>10</sup> with a mean d of  $>3.5$  mm. In anterior circulation aneurysms, distal aneurysms accounted for only 3.5%. According to a recent study by Wei et al,<sup>11</sup> the incidence of moderate and severe ISS was 15.2%, the same as ours, whereas distal arterial aneurysms accounted for only 2.3%.

To date, no study has investigated the incidence rate and associated factors of ISS after PED placement for distal anterior circulation aneurysms. Given the critical function of the distal anterior circulation, identifying patients at high risk of ISS and providing appropriate medical prophylaxis are critical. The associated factors included the TyG index, arterial and stent diameters, age, sex, and other factors.

The TyG index is a composite indicator calculated from fasting blood glucose and triglycerides, each of which represents the degree of insulin resistance in the liver and fat cells.<sup>15</sup> Cardiovascular studies have shown that a higher TyG index is associated with an increased incidence of coronary artery disease and an unfavorable prognosis.<sup>16-18</sup> Individuals with a higher TyG index may also have a higher risk of stroke.<sup>19</sup> As a marker of insulin resistance, the influence of TyG on blood



vessels can be explained by the mechanisms associated with insulin resistance. Hyperglycemia increases the production of reactive oxygen species, resulting in insulin resistance with impaired insulin-stimulated translocation of glucose transporter type 4 and glucose uptake.<sup>20</sup> In addition, reactive oxygen species activate proinflammatory signaling.<sup>21</sup> Moreover, increased free fatty acid levels directly inhibit glucose transport by causing mitochondrial dysfunction,<sup>22,23</sup> which enhances glucotoxicity. Insulin can stimulate the production of nitric oxide in endothelial cells<sup>24</sup>; an elevated TyG index may interfere with insulin signaling, reduce insulin sensitivity, and thereby affect the local microenvironment of the blood vessels and alter hemodynamics. Insulin can additionally inhibit cell migration and neointimal growth following arterial damage,<sup>25</sup> and the mechanism of ISS is caused by excessive growth of the intima. These factors may be the mechanisms behind the higher incidence of ISS in patients with higher TyG index.

Regarding the d and the difference between the diameters of the stent and parent artery, typically, the small diameter and tortuosity of the parent arteries increases the difficulty of intraoperative procedures and may cause vascular injury. Endothelial damage leads to platelet adhesion, activation, aggregation, and even thrombus formation, causing local hemodynamic disturbance. Additionally, the increased reactive oxygen species induced by hyperglycemia can promote nitrous oxide clearance, thereby restricting vasodilation.<sup>26</sup> The increased expression of proinflammatory genes caused by reactive oxygen species, combined with the weakened anti-inflammatory effect of nitrous oxide, activates the inflammation mechanism. In clinical practice, neurointerventionists tend to select stents slightly larger than the parent artery diameter to avoid ischemic events in the presence of important perforating arteries such as the lenticulostriate and Huebner recurrent arteries. Our restricted cubic spline curve shows an increased risk of ISS as Dd increases: when Dd = 0.38 mm, the OR = 1. Therefore, we suggest that when a stent covers the perforating arteries in A1 and M1, a stent slightly larger (preferably not more than 0.5 mm) than the diameter of the parent arteries should be selected to prevent the development of ISS.

Studies of age and sex show that younger patients are more likely to develop ISS due to a stronger repair mechanism.<sup>27</sup> However, we showed that patients aged older than 55 years of age are more prone to developing ISS, which may be due to the higher incidence of comorbidities and increased vascular tortuosity associated with aging, as well as a relatively higher prevalence of concomitant stenosis of the parent arteries and a weaker anti-inflammatory mechanism.

Estrogen can activate endothelial nitric oxide synthase in endothelial cells via the phosphatidylinositol-3-kinase-dependent signaling pathway, thereby increasing nitrous oxide levels and regulating local blood flow, while exerting anti-inflammatory effects.<sup>28</sup> However, as estrogen levels decline with aging and menopause, the regulating effect disappears. Additionally, estrogen can stimulate the retention of water and sodium. Decreased levels of estrogen increase the viscosity of the blood to a certain extent.

Additionally, ISS peaks at 12 months postprocedure with minimal progression thereafter.<sup>29</sup> However, ISS can be reversed within 24 months, implying that intraluminal hyperplasia may

undergo remodeling, leading to a decrease in lost luminal diameter. Recent studies on animals have suggested that ISS may be associated with endothelial thrombosis.<sup>30</sup> Incomplete stent apposition may result in greater wall shear stress in the corresponding area, leading to increased intimal hyperplasia and ultimately resulting in ISS.<sup>31</sup> We included a follow-up period, stent malapposition, and parent arteries with >50% stenosis in a multivariable logistic regression analysis, but none showed statistical significance. Studies have reported the association between smoking status, balloon angioplasty, fusiform aneurysm, non-aspirin use, and ISS after PED implantation, which may be related to endothelial damage, inflammatory mechanisms, and other factors. We found no significant differences between the 2 groups in these factors.

This study has certain limitations. First, it is a retrospective study, which could result in some selection bias. Second, it may have overestimated stent restenosis rates because of a median follow-up of 9.07 months and the lack of long-term follow-up. Third, strict dual antiplatelet management at our center may have limited the investigation of drug-related effects. Finally, treatment of distal anterior circulation aneurysms with the PED is an off-label use and has been used in only a few clinical cases. Large prospective studies may be required to validate the results of this study.

## CONCLUSIONS

A major difference in diameter between the stent and the parent artery, a higher TyG index, and 55 years of age or older are associated with an increased rate of ISS. Selecting a stent of similar or slightly larger diameter than the parent artery to avoid ISS is recommended. Managing comorbidities should also be emphasized in clinical practice.

**Disclosure forms** provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

- Hanel RA, Kallmes DF, Lopes DK, et al. **Prospective study on embolization of intracranial aneurysms with the Pipeline device: the PREMIER study 1 year results.** *J Neurointerv Surg* 2020;12:62–66 CrossRef Medline
- 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
- Yavuz K, Geyik S, Saatci I, et al. **Endovascular treatment of middle cerebral artery aneurysms with flow modification with the use of the Pipeline Embolization Device.** *AJNR Am J Neuroradiol* 2014; 35:529–35 CrossRef Medline
- Primiani CT, Ren Z, Kan P, et al. **A2, M2, P2 aneurysms and beyond: results of treatment with Pipeline Embolization Device in 65 patients.** *J Neurointerv Surg* 2019;11:903–07 CrossRef Medline
- 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
- 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 CrossRef Medline
- Lin N, Lanzino G, Lopes DK, et al. **Treatment of distal anterior circulation aneurysms with the Pipeline Embolization Device: a US multicenter experience.** *Neurosurgery* 2016;79:14–22 CrossRef Medline

8. Martin AR, Cruz JP, O'Kelly C, et al. **Small pipes: preliminary experience with 3-mm or smaller Pipeline flow-diverting stents for aneurysm repair prior to regulatory approval.** *AJNR Am J Neuroradiol* 2015;36:557–61 CrossRef Medline
9. Liou TM, Li YC. **Effects of stent porosity on hemodynamics in a sidewall aneurysm model.** *J Biomech* 2008;41:1174–83 CrossRef Medline
10. Turhon M, Kang H, Liu J, et al. **In-stent stenosis after Pipeline Embolization Device in intracranial aneurysms: incidence, predictors, and clinical outcomes.** *Neurosurgery* 2022;91:943–51 CrossRef Medline
11. Wei D, Deng D, Gui S, et al. **Machine learning to predict in-stent stenosis after Pipeline Embolization Device placement.** *Front Neurol* 2022;13:92984 CrossRef Medline
12. O'Kelly CJ, Krings T, Fiorella D, et al. **A novel grading scale for the angiographic assessment of intracranial aneurysms treated using flow diverting stents.** *Interv Neuroradiol* 2010;16:133–37 CrossRef Medline
13. Park K, Ahn CW, Lee SB, et al. **Elevated TyG index predicts progression of coronary artery calcification.** *Diabetes Care* 2019;42:1569–73 CrossRef Medline
14. John S, Bain MD, Hui FK, et al. **Long-term follow-up of in-stent stenosis after Pipeline flow diversion treatment of intracranial aneurysms.** *Neurosurgery* 2016;78:862–67 CrossRef Medline
15. Tian X, Chen S, Zhang Y, et al. **Time course of the triglyceride glucose index accumulation with the risk of cardiovascular disease and all-cause mortality.** *Cardiovasc Diabetol* 2022;21:183 CrossRef Medline
16. Luo E, Wang D, Yan G, et al. **High triglyceride-glucose index is associated with poor prognosis in patients with acute ST-elevation myocardial infarction after percutaneous coronary intervention.** *Cardiovasc Diabetol* 2019;18:150 CrossRef Medline
17. Guo X, Shen R, Yan S, et al. **Triglyceride-glucose index for predicting repeat revascularization and in-stent restenosis in patients with chronic coronary syndrome undergoing percutaneous coronary intervention.** *Cardiovasc Diabetol* 2023;22:43 CrossRef Medline
18. Ding X, Wang X, Wu J, et al. **Triglyceride-glucose index and the incidence of atherosclerotic cardiovascular diseases: a meta-analysis of cohort studies.** *Cardiovasc Diabetol* 2021;20:76 CrossRef Medline
19. Wang A, Wang G, Liu Q, et al. **Triglyceride-glucose index and the risk of stroke and its subtypes in the general population: an 11-year follow-up.** *Cardiovasc Diabetol* 2021;20:46 CrossRef Medline
20. Kim J, Montagnani M, Koh KK, et al. **Reciprocal relationships between insulin resistance and endothelial dysfunction: molecular and pathophysiological mechanisms.** *Circulation* 2006;113:1888–904 CrossRef Medline
21. Kamata H, Honda SI, Maeda S, et al. **Reactive oxygen species promote TNF $\alpha$ -induced death and sustained JNK activation by inhibiting MAP kinase phosphatases.** *Cell* 2005;120:649–61 CrossRef Medline
22. Savage DB, Petersen KF, Shulman GI. **Mechanisms of insulin resistance in humans and possible links with inflammation.** *Hypertension* 2005;45:828–33 CrossRef Medline
23. Lowell BB, Shulman GI. **Mitochondrial dysfunction and type 2 diabetes.** *Science* 2005;307:384–87 CrossRef Medline
24. Zeng G, Nystrom FH, Ravichandran LV, et al. **Roles for insulin receptor, PI3-kinase, and Akt in insulin-signaling pathways related to production of nitric oxide in human vascular endothelial cells.** *Circulation* 2000;101:1539–45 CrossRef Medline
25. Breen DM, Chan KK, Dhaliwall JK, et al. **Insulin increases reendothelialization and inhibits cell migration and neointimal growth after arterial injury.** *Arterioscler Thromb Vasc Biol* 2009;29:1060–66 CrossRef Medline
26. Salt IP, Morrow VA, Brandie FM, et al. **High glucose inhibits insulin-stimulated nitric oxide production without reducing endothelial nitric-oxide synthase Ser1177 phosphorylation in human aortic endothelial cells.** *J Biol Chem* 2003;278:18791–97 CrossRef Medline
27. Sweid A, Starke RM, Herial N, et al. **Predictors of complications, functional outcome, and morbidity in a large cohort treated with flow diversion.** *Neurosurgery* 2020;87:730–43 CrossRef Medline
28. Simoncini T, Hafezi-Moghadam A, Brazil DP, et al. **Interaction of oestrogen receptor with the regulatory subunit of phosphatidylinositol-3-OH kinase.** *Nature* 2000;407:538–41 CrossRef Medline
29. Gui S, Chen X, Wei D, et al. **Long-term outcomes and dynamic changes of in-stent stenosis after Pipeline embolization device treatment of intracranial aneurysms.** *J Neurointerv Surg* 2023;15:1187–93 CrossRef Medline
30. Monteiro A, Lopes DK, Aghaebrahim A, et al. **Optical coherence tomography for elucidation of flow-diversion phenomena: the concept of endothelialized mural thrombus behind reversible in-stent stenosis in flow-diverters.** *Interv Neuroradiol* 2021;27:774–80 CrossRef Medline
31. Sindeev S, Prothmann S, Frolov S, et al. **Intimal hyperplasia after aneurysm treatment by flow diversion.** *World Neurosurg* 2019;122:e577–83 CrossRef Medline

# Initial Experiences with the Trenza Embolization Device for the Treatment of Wide-Neck Intracranial Aneurysms: A 12-Patient Case Series

Rahul Raj and Jussi Numminen



## ABSTRACT

**SUMMARY:** The Trenza embolization device is a frame coil implant with flow-disruption properties and is a new alternative to treat challenging mid-to-large-sized broad-neck bifurcation or sidewall aneurysms. We conducted an observational single-center retrospective study of 12 consecutive patients treated for 10 unruptured and 2 ruptured 6- to 12-mm broad-neck bifurcation or sidewall aneurysms with the Trenza device during 2022–2023. The median patient age was 64 years (interquartile range, 59–70 years), 58% were women, the median largest aneurysm diameter was 9.6 mm (interquartile range, 7.5–11.9 mm), the median dome-to-neck ratio was 1.8 (interquartile range, 1.6–1.9), the most common aneurysm locations were the anterior communicating artery (33%) and basilar artery tip (33%). After a median follow-up of 6.5 months, adequate aneurysm occlusion was achieved in 83%. There were 3 major ischemic complications (25%), leading to 2 permanent neurologic deficits (17%) and 1 transient neurologic deficit (8%). There was 1 fatal rupture of a treated aneurysm 1.6 months after the index treatment. Two patients were retreated (17%). Ischemic complications occurred in patients after a too-dense coil packing at the base of the aneurysm. No technical issues related to the device were encountered. In summary, an adequate aneurysm occlusion rate was achieved using the Trenza-assisted coiling technique for otherwise challenging mid-to-large-sized broad-neck aneurysms. Ischemic complications seemed to occur following overdense coiling at the base of the aneurysm.

**ABBREVIATIONS:** ASA = acetylsalicylic acid; DAPT = double antiplatelet treatment; IQR = interquartile range

Endovascular coiling or balloon-assisted coiling of broad-neck mid-sized intracranial aneurysms without the use of additional stents is associated with an up to 20%–30% probability of incomplete occlusion<sup>1–5</sup> but with a rather low risk of thromboembolic complications.<sup>6</sup> Stent-assisted coiling improves the occlusion rate but is associated with an increased risk of thromboembolic events, especially when treating ruptured intracranial aneurysms or when using complex stent configurations.<sup>7–10</sup> Thus, to keep the occlusion rate of stent-assisted coiling but to minimize the risk of stent-associated thromboembolic events, especially in ruptured aneurysms, endosaccular flow-disruption devices such as the Woven EndoBridge device (WEB; Sequent Medical), the Artisys intrasaccular device (Medtronic), and the Contour device (Cerus Endovascular) have been introduced.<sup>11–16</sup> Yet, these intrasaccular

devices have geometric and size restrictions, which limit their use in certain aneurysms.<sup>17</sup>

Recently, the Trenza coil implant (Stryker) was introduced as an alternative to treat 6- to 12-mm broad-neck bifurcation or sidewall aneurysms. The Trenza device is an intrasaccular braided ribbon-like frame coil implant, with flow-disruption properties, which aims to create a stable basket in the aneurysm that is then filled with standard coils. It is hypothesized that due to the endosaccular nature, the flow-disruption properties of the Trenza device together with a normal coiling procedure would lead to higher occlusion rates but a lower risk of thromboembolic complications compared with simple coiling alone or stent-assisted coiling.

Here we summarize our initial experience treating 12 patients with the Trenza device with the aim of transparently reporting radiologic and neurologic complications as well as clinical and 6-month angiographic results.

## MATERIALS AND METHODS

We conducted a single-center retrospective analysis of all consecutive patients with intracranial aneurysms treated with the Trenza embolization device at Helsinki University Hospital between May 2022 and January 2023. The Trenza device is CE approved and was available for routine clinical use in Finland

Received November 9, 2023; accepted after revision December 20.

From the Departments of Neurosurgery (R.R.) and Radiology (J.N.), Helsinki University Hospital and University of Helsinki, Helsinki, Finland.

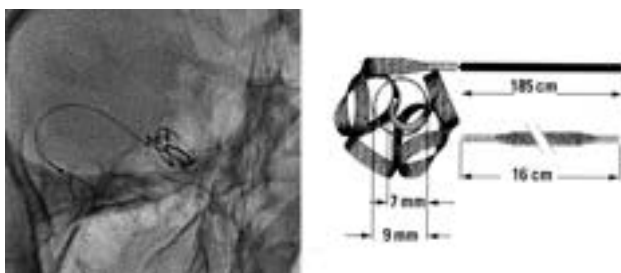
This study was supported by the state funding for university-level health research in Finland (TYH2023330, Helsinki University Hospital).

Please address correspondence to Rahul Raj, MD, PhD, Bridge Hospital, 3rd floor, Room STA3.082, Haartmaninkatu 4, PO Box 320, 00029 HUS, Helsinki, Finland; e-mail: rahul.raj@hus.fi; @rahulbrraj



Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8153>



**FIG 1.** Native image showing the Trezza embolization device (left) and an illustration of a 9-mm Trezza device.

during the study. All consecutive patients were treated before the start of the Trezza Embolization Device for Intracranial Aneurysm Treatment study (TREAT, clinicaltrials.gov identifier: NCT04380350; recruitment started in Helsinki in December 2022). We included patients with ruptured and unruptured aneurysms. The decision to treat the aneurysm was made by a multidisciplinary neurovascular team. The decision to treat the aneurysm with the Trezza embolization device was made by the operating interventionalist. Aneurysms considered for treatment with the Trezza device had to be 6- to 12-mm bifurcation or sidewall aneurysms with a wide neck (defined as neck of  $\geq 4$  mm or dome-to-neck ratio of  $< 2$ ). Aneurysms fulfilling the criteria were carefully reviewed before deciding to use the Trezza device. Aneurysms that were considered to have a high risk of treatment-related complications using other techniques (eg, stent-assisted coiling or flow diversion) and suitable for the Trezza device were selected. Although the aspect ratio of the aneurysm does not per se affect the usability of the Trezza, given that the other measurement features are fulfilled, we did not choose aneurysms with a marked discrepancy in aneurysm dimensions.

The institutional review board approved the study (research approval HUS/313/2022). Due to the retrospective design of the study, national legislation does not warrant informed consent because the Trezza device was used as part of routine care. The study was reported with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.

### Trezza Device

The Trezza embolization device is a specially designed coil implant that creates a stable  $\omega$ -shaped basket to be filled with coils until the aneurysm is occluded. The coil implant is designed to have flow-disruption properties to enhance aneurysm thrombosis. The Trezza device is available in 6- to 12-mm sizes, with 1-mm intervals. Each size has a fixed length (eg,  $6 \times 11$  mm,  $7 \times 13$  mm,  $8 \times 15$  mm, and so forth). One or several Trezza devices can be used to create a stable basket after which the recommendation is to fill the basket with Target XL (Stryker Neurovascular) coils until the aneurysm is occluded. The device is attached to a flexible delivery wire, and detachment of the device is electrothermal, similar to several other neurovascular implant-delivery systems (Fig 1). The device is sized as the mean diameter of the aneurysm in 3 planes, irrespective of the neck size, with no plane being smaller than 5 mm without accounting for secondary lobules.

The indications to use the Trezza embolization device are broad-based saccular bifurcation or sidewall aneurysms (defined

as having a dome-to-neck ratio of  $< 2$  or a neck ratio of  $\geq 4.0$  mm) between 6 and 12 mm in size.

### Treatment Setup

For treatment, we used a coaxial system (femoral or radial access, 8F-long sheath, distal-access catheter) with the Excelsior 1018 microcatheter (Stryker). All patients were heparinized during the treatment, and heparinization levels were followed using active clotting time measurements. Generally, we aimed for an active clotting time value of approximately 2 times the baseline level.

In general, our pretreatment antiplatelet regimen consisted of double antiplatelet treatment (DAPT), consisting of acetylsalicylic acid (ASA), 100 mg once a day, and prasugrel, 10 mg once a day, or clopidogrel, 75 mg once a day, starting 7 days before the intervention. Our posttreatment antiplatelet regimen was ASA, 100 mg once a day for 3 months, either with or without prasugrel (favoring ASA + prasugrel with increased experience with the Trezza device). We tested the preintervention antiplatelet medication response using the VerifyNow method (Werfen).

### Definition of Radiologic Complications, Neurologic Complications, Clinical Outcome, and Radiologic Outcome

We defined radiologic complications as hemorrhagic or ischemic. Hemorrhagic and ischemic complications had to be symptomatic and verified by CT or MR imaging within 6 months of the treatment and located in the vascular territory of the treated aneurysm. All patients underwent postoperative head CT on the first postoperative day. Further CT or MR imaging scans were obtained at the discretion of the treating clinicians. Patients with ruptured aneurysms were imaged at the discretion of the treating clinicians. We further noted whether the patient was retreated or scheduled for retreatment during the study period.

We defined symptomatic neurologic complication as a new neurologic deficit with supportive imaging findings related to the treated aneurysm occurring within 6 months of the treatment. If the neurologic symptom resolved within 6 months, we classified it as being transient.

We assessed clinical outcomes using the mRS (range, 0 [no symptoms] to 6 [death]) 6 months after the index treatment at the time of the control DSA.

We performed a routine DSA follow-up at approximately 6 months for all patients. We assessed aneurysm occlusion rates using the modified Raymond-Roy classification.<sup>18</sup>

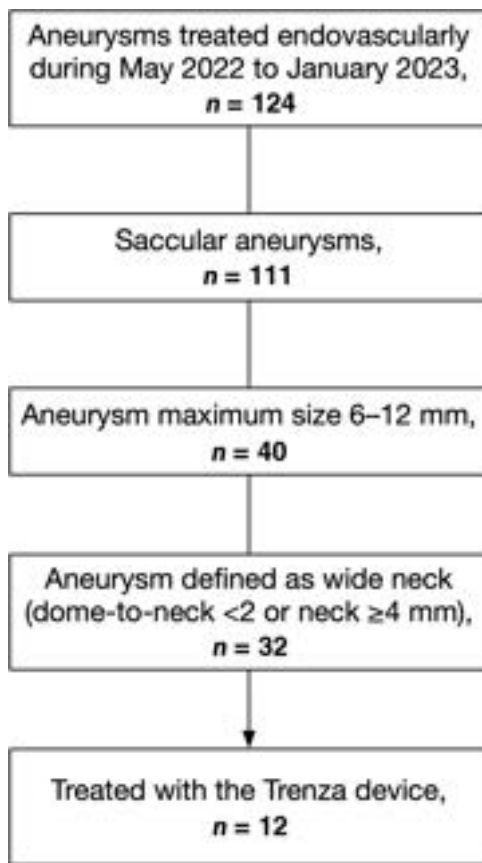
### Statistical Analysis

Given the small sample size, we used only descriptive statistics. Categorical data were presented as number (percentage); parametric data, as mean (SD); and nonparametric data, as median (interquartile range [IQR]). STATA (StataCorp) was used for the statistical analyses.

## RESULTS

### Baseline Characteristics

Of 32 eligible aneurysms, a total of 12 aneurysms in 12 patients were treated from May 2022 to January 2023 with the Trezza embolization device (Fig 2). The median age was 64 years, 58%



**FIG 2.** Flow chart describing the patient population.

were women, 42% were never smokers, 58% had a history of hypertension, and 17% (2 patients) had acutely ruptured aneurysms (Table 1).

Nine patients were DAPT-premedicated, and 1 patient was prasugrel-only pretreated due to ASA allergy. Two patients were acutely ruptured cases and did not receive any pretreatment antithrombotic medication. The posttreatment antiplatelet regime was DAPT (ASA + prasugrel) in 4 patients, ASA only in 6 patients, ASA + apixaban in 1 patient (due to concomitant atrial fibrillation), and prasugrel only in 1 patient (due to ASA allergy).

Anterior communicating artery and basilar artery aneurysms were the most frequent (33% for both). The median largest aneurysm diameter was 9.6 mm, median neck length was 5.4 mm, and the median dome-to-neck ratio was 1.8 (Table 1).

#### Treatment Characteristics

The median intervention time was (from puncture to closure) 121 minutes (IQR, 100–170 minutes). In 1 procedure, an additional stent (NeuroForm Atlas; Stryker) was used to protect the lumen of the parent artery (case 8 in the Online Supplemental Data); in 1 case, a Comaneci device (Rapid Medical) was used as support during the embolization (case 2 in the Online Supplemental Data); in 1 case, a balloon was used during the implantation of the Trezza device (case 11 in the Online Supplemental Data); and in 1 case an unruptured right-sided multilobular anterior cerebral artery, 1 segment

**Table 1: Patient and aneurysm characteristics**

Variable	Patients (n = 12)
<b>Patient characteristics</b>	
Age (median) (IQR) (yr)	64 (59–70)
Female sex	7 (58%)
Smoking	
Never	5 (42%)
Ex-smoker, stopped >6 months ago	5 (42%)
Active smoker or stopped <6 months ago	2 (16%)
Hypertension	7 (58%)
Diabetes	2 (17%)
<b>Aneurysm characteristics</b>	
Acutely ruptured	2 (17%)
Location	
AcomA	4 (33%)
Basilar artery tip	4 (33%)
ICA communicating segment	3 (25%)
MCA	1 (9%)
Height (median) (IQR) (mm)	8.2 (6.3–8.9)
Dome (median) (IQR) (mm)	9.6 (7.4–11.9)
Neck (median) (IQR) (mm)	5.4 (3.7–7.2)
Largest diameter (median) (IQR) (mm)	9.6 (7.5–11.9)
Aspect ratio (median) (IQR)	0.83 (0.78–1.09)
Dome-to-neck ratio (median) (IQR)	1.8 (1.6–1.9)

**Note:**—AcomA indicates anterior communicating artery.

(A1), aneurysm was coiled in the same session (case 5 in the Online Supplemental Data).

One Trezza device was used in 5 cases (42%), and 2 Trezza devices were used in 7 cases (58%). There were no technical failures related to the Trezza device.

All pre-embolization DSA images and follow-up DSA images are shown in the Online Supplemental Data.

#### Radiologic and Neurologic Complications

Two patients (17%) had an immediate treatment-related symptomatic ischemic complication, and 1 patient's aneurysm bled 1.6 months after the index treatment, which lead to death. Both patients with symptomatic ischemic complications were pretreated with DAPT. The clinical consequences of the ischemic complications were 2 (17%) permanent neurologic deficits (Table 2). There were no delayed ischemic complications.

#### Clinical Outcome

The median follow-up time was 6.5 months (IQR, 6.2–6.8 months). Of the patients treated for an unruptured aneurysm, 7 of 10 (70%) had an unchanged mRS of 0. One patient died from a posttreatment bleed from the treated aneurysm, and 2 patients' mRS scores worsened to 2.

Both patients treated for a ruptured aneurysm (initial SAH severity, World Federation of Neurosurgical Societies I and IV, respectively) had good clinical outcomes (mRS 0 and 1, respectively).

#### Radiologic Outcome and Retreatments

Eight patients (67%) had a complete occlusion, and 10 patients (83%) had an adequate occlusion of the treated aneurysm after a median of 6.5 months (IQR, 6.2–6.8 months). Two patients had an inadequately obliterated aneurysm, with contrast opacification between the Trezza and coils and the aneurysmal wall. One caused a rupture 1.6 months from the index treatment, and the



**Table 2: Aneurysm occlusion rate and outcomes<sup>a</sup>**

Characteristic	Patients (n = 12)
Aneurysm occlusion <sup>b</sup>	
Complete obliteration	8 (67%)
Residual neck	2 (16.5%)
Residual aneurysm with contrast filling within the coils	0
Residual aneurysm with contrast filling between the coils and the aneurysm wall	2 (16.5%)
Complications	
Ischemic complications	2 (17%)
Hemorrhagic complications	1 (8%)
Symptomatic neurologic complications	
Permanent	2 (17%)
Transient	0 (0%)
Death	1 (7.5%)
Permanent symptomatic neurologic deficit or death	3 (25%)

<sup>a</sup> All variables presented as No. (%).

<sup>b</sup> Median time to imaging follow-up 6.5 months (IQR, 6.2–6.8 months).

patient died after an acute re-intervention (case 9 in the Online Supplemental Data). The other re-treated patient was re-treated 7 months after the index treatment (case 10 in the Online Supplemental Data). Thus, 2 of 12 patients were re-treated (17%).

## DISCUSSION

In this case series, we evaluated 12 patients with broad-based mid-to-large-sized aneurysms (median diameter of nearly 10 mm, median dome-to-neck ratio of 1.8) who underwent endovascular treatment using the Trezza coil implant device complemented by additional coiling. The six-month follow-up demonstrated an occlusion rate of 83%. This rate is particularly notable given the challenging nature of the aneurysms treated, characterized by their size and broad necks—factors traditionally associated with lower occlusion rates and higher procedural risks.<sup>2</sup>

The use of the Trezza device was also associated with a reduction in the need for stent deployment, with 11 of the 12 patients avoiding stents altogether. This outcome not only simplifies the treatment but may also lead to a lower risk of ischemic complications, which are more prevalent in patients with larger aneurysms and wider necks.<sup>6,19,20</sup> Furthermore, given the challenging nature of the treated aneurysms, an assisting support device during the coiling was only needed in 2 cases. The avoidance of support devices and stents may contribute to a lower incidence of ischemic complications, especially in ruptured aneurysms, though this result remains to be elucidated in comparative studies. Thus, the ability of the Trezza device to achieve high rates of occlusion in challenging aneurysms without the adjunctive use of stents is noteworthy.

Despite the high occlusion rate, there was a 17% incidence of ischemic complications within this cohort. This figure, while not negligible, must be interpreted with caution due to the small sample size and the complex anatomy of the aneurysms treated. Specifically, there were 2 ischemic complications in 2 patients with basilar tip aneurysms measuring 12 and 14 mm, where the coiling at the base was excessively dense, leading to compromised flow in the posterior cerebral arteries (as shown in cases 8 and 10 of the Online Supplemental Data). Thus, our preliminary data

suggest that for large bifurcation aneurysms with broad necks, a strategy of less dense coiling at the base may be advantageous, preserving the patency of the bifurcation. This approach is facilitated by the flow-disruption capabilities of the Trezza device, which appears to promote effective aneurysm occlusion even with less dense packing at the neck. The efficacy of the Trezza device in flow disruption and subsequent aneurysm occlusion is particularly evident in cases 2, 3, and 11 (Online Supplemental Data), in which, despite the aneurysms being loosely coiled at their bases, complete occlusion was achieved at follow-up.

In this cohort, 92% of patients had favorable functional outcomes after 6 months. One patient had a fatal rupture of the treated unruptured ICA aneurysm (maximum diameter, 12 mm) 1.6 months after treatment (case 9 in the Online Supplemental Data). Considering the aneurysm size, the short time span from treatment to rupture, and the point of extravasation being seen from the densely coiled part of the aneurysm, this complication is likely related to the instability of the large aneurysm instead of being related to the device itself.

Technically, the use of the Trezza device is similar to that of a normal coiling procedure with some exceptions. The Excelsior 1018 microcatheter (0.019 inch) must be used, which is somewhat larger and stiffer than other coiling microcatheters but smaller than some microcatheters used for other devices. The Trezza device is stiffer than normal framing coils and may cause recoil kickback of the microcatheter, highlighting the need for good proximal catheter support. Furthermore, if the Trezza basket does not cover the aneurysm optimally, it may be challenging to coil the space between the Trezza and the aneurysm wall (for example, case 5 in the Online Supplemental Data). However, complete or dense coiling of the aneurysm might not always be necessary to achieve complete aneurysm occlusion (case 11 in the Online Supplemental Data).

There have been several intrasaccular devices designed to treat wide-neck aneurysms. These include, for example, the Medina device (Medtronic), the Luna aneurysm embolization device (Medtronic), the WEB device, and the Contour device. Of these various intrasaccular devices, the Trezza and the Medina share similar engineering principles in that both are coil-like implants with flow-disrupting properties. Reports on the Medina device have shown adequate occlusion rates of 71%–83% (83% in our study), retreatment rates of 8%–38% (17% in our study), and complication rates of 8%–23% (17% in our study).<sup>21–23</sup> Studies on the Medina device have included smaller aneurysms with wider necks; thus, the results are not directly comparable. Of the intrasaccular devices, the WEB is the most established, with long-term adequate occlusion rates of approximately 90%,<sup>24</sup> long-term retreatment rates of 12%,<sup>25</sup> and an ischemic complication rate of 7.5%.<sup>26</sup> Data regarding the Luna and Contour devices are scarcer, but reports have shown adequate occlusion and ischemic complication rates of 85% and 2.5%–12.5%, respectively, for the Contour<sup>15,27</sup> and 78% and 5%, respectively, for the Luna.<sup>28</sup> Nevertheless, comparative long-term outcome studies of the various intrasaccular devices are needed of the optimal device for specific patients and aneurysms.

In conclusion, the Trezza coil implant device has potential as a valuable tool in the endovascular armamentarium for the

treatment of broad-neck aneurysms measuring 6–12 mm. Nevertheless, the findings from this series should be interpreted as preliminary. Larger, controlled studies are warranted to corroborate these results and to refine the coiling strategies that optimize clinical outcomes. Furthermore, comparative studies with other endovascular techniques are essential to delineate the role of the Trenza device in the broader context of aneurysm management.

### Limitations

This is a case series of only 12 patients, so there are some inherent limitations. First, the small sample size limits the generalizability of the results, and further larger studies are needed to really assess the effectiveness and safety of the Trenza device. Second, because the device is new, there are no data regarding the optimal antiplatelet treatment strategy. Because the treated aneurysms were broad-neck and the need for stents is challenging to exclude before treatment, we pretreated most patients with DAPT and continued with DAPT or single antiplatelet therapy after the intervention for 3 months. The use of DAPT is somewhat contradictory because one of the main advantages of intrasaccular over parent artery stent placement is the avoidance of DAPT. Pretreatment with DAPT was used as a safety measure in case of bailout stent placement, something needed in only 1 of the 12 cases. DAPT was continued for 3 months due to the relatively high metal coverage of a wide neck, again, used as a safety measure that may not be necessary. For example, only posttreatment aspirin was used in the 2 ruptured cases, with neither patient exhibiting any signs of device-related ischemic complications. Third, because we had no clinical experience in using the device (except ex vivo simulations), the duration of the interventions was longer than normal (median, 2 hours). It is likely that with increased experience, intervention times will go down, possibly reducing the risk of ischemic complications. Furthermore, with increased experience, a better understanding of how to place the Trenza device to achieve complete occlusion is likely, possibly increasing the occlusion rate and decreasing the risk of ischemic complications. The learning curve is expected to be low due to the similarities to a normal, simple coiling procedure. Fourth, we report results after a median of 6 months. More long-term data are needed to establish the safety and effectiveness of the Trenza device.

### CONCLUSIONS

An adequate aneurysm occlusion rate was achieved using the Trenza-assisted coiling technique for otherwise challenging mid-to-large-sized broad-neck aneurysms. Ischemic complications seemed to occur following overly dense coiling at the base of the aneurysm.







Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

### REFERENCES

- Raymond J, Klink R, Chagnon M, et al. **Hydrogel versus bare platinum coils in patients with large or recurrent aneurysms prone to recurrence after endovascular treatment: a randomized controlled trial.** *AJNR Am J Neuroradiol* 2017;38:432–41 CrossRef Medline
- Pierot L, Barbe C, Thierry A, et al. **Patient and aneurysm factors associated with aneurysm recanalization after coiling.** *J Neurointerv Surg* 2022;14:1096–101 CrossRef Medline
- Raymond J, Guilbert F, Weill A, et al. **Long-term angiographic recurrences after selective endovascular treatment of aneurysms with detachable coils.** *Stroke* 2003;34:1398–403 CrossRef Medline
- Pierot L, Cognard C, Anxionnat R, et al; CLARITY Investigators. **Endovascular treatment of ruptured intracranial aneurysms: factors affecting midterm quality anatomic results—analysis in a prospective, multicenter series of patients (CLARITY).** *AJNR Am J Neuroradiol* 2012;33:1475–80 CrossRef Medline
- Ferns SP, Sprengers ME, van Rooij WJ, et al. **Coiling of intracranial aneurysms: a systematic review on initial occlusion and reopening and retreatment rates.** *Stroke* 2009;40:e523–29 CrossRef Medline
- Pierot L, Barbe C, Nguyen HA, et al. **Intraoperative complications of endovascular treatment of intracranial aneurysms with coiling or balloon-assisted coiling in a prospective multicenter cohort of 1088 participants: Analysis of Recanalization after Endovascular Treatment of Intracranial Aneurysm (ARETA) study.** *Radiology* 2020;296:E130–33 CrossRef Medline
- Piotin M, Blanc R, Spelle L, et al. **Stent-assisted coiling of intracranial aneurysms: clinical and angiographic results in 216 consecutive aneurysms.** *Stroke* 2010;41:110–15 CrossRef Medline
- Bsat S, Bsat A, Tamim H, et al. **Safety of stent-assisted coiling for the treatment of wide-necked ruptured aneurysm: a systematic literature review and meta-analysis of prevalence.** *Interv Neuroradiol* 2020;26:547–56 CrossRef Medline
- Hong Y, Wang YJ, Deng Z, et al. **Stent-assisted coiling versus coiling in treatment of intracranial aneurysm: a systematic review and meta-analysis.** *PLoS One* 2014;9:e82311 CrossRef Medline
- Bartolini B, Blanc R, Pistocchi S, et al. **“Y” and “X” stent-assisted coiling of complex and wide-neck intracranial bifurcation aneurysms.** *AJNR Am J Neuroradiol* 2014;35:2153–58 CrossRef Medline
- Kabbasch C, Goertz L, Siebert E, et al. **WEB embolization versus stent-assisted coiling: comparison of complication rates and angiographic outcomes.** *J Neurointerv Surg* 2019;11:812–16 CrossRef Medline
- Naamani KE, Chen CJ, Abbas R, et al. **Woven EndoBridge versus stent-assisted coil embolization of cerebral bifurcation aneurysms.** *J Neurosurg* 2022;137:1786–93 CrossRef Medline
- Arthur AS, Molyneux A, Coon AL, et al; WEB-IT Study investigators. **The safety and effectiveness of the Woven EndoBridge (WEB) system for the treatment of wide-necked bifurcation aneurysms: final 12-month results of the pivotal WEB Intrasaccular Therapy (WEB-IT) Study.** *J Neurointerv Surg* 2019;11:924–30 CrossRef Medline
- Pierot L, Moret J, Barreau X, et al. **Safety and efficacy of aneurysm treatment with WEB in the cumulative population of three prospective, multicenter series.** *J Neurointerv Surg* 2018;10:553–59 CrossRef Medline
- Liebig T, Killer-Oberpfalzer M, Gal G, et al. **The safety and effectiveness of the Contour Neurovascular System (Contour) for the treatment of bifurcation aneurysms: the CERUS Study.** *Neurosurgery* 2022;90:270–77 CrossRef Medline
- Piotin M, Fahed R, Redjem H, et al. **The ARTISSE intrasaccular device for intracranial aneurysm treatment: short-term, mid-term and long-term clinical and angiographic results.** *J Neurointerv Surg* 2022;14:957–61 CrossRef Medline
- Goyal N, Hoit D, DiNitto J, et al. **How to WEB: a practical review of methodology for the use of the Woven EndoBridge.** *J Neurointerv Surg* 2020;12:512–20 CrossRef Medline
- 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
- Pierot L, Spelle L, Vitry F, ATENA Investigators. **Immediate clinical outcome of patients harboring unruptured intracranial aneurysms treated by endovascular approach: results of the ATENA study.** *Stroke* 2008;39:2497–504 CrossRef Medline

20. Pierot L, Cognard C, Anxionnat R, et al; CLARITY Investigators. **Ruptured intracranial aneurysms: factors affecting the rate and outcome of endovascular treatment complications in a series of 782 patients (CLARITY study).** *Radiology* 2010;256:916–23 CrossRef Medline
21. Bhogal P, Brouwer PA, Yeo L, et al. **The Medina Embolic Device: Karolinska experience.** *Interv Neuroradiol* 2018;24:4–13 CrossRef Medline
22. Sourour NA, Vande Perre S, Maria FD, et al. **Medina® embolization device for the treatment of intracranial aneurysms: safety and angiographic effectiveness at 6 months.** *Neurosurgery* 2018;82:155–62 CrossRef Medline
23. Haffaf I, Clarençon F, Shotar E, et al. **Medina embolization device for the treatment of intracranial aneurysms: 18 months' angiographic results.** *J Neurointerv Surg* 2019;11:516–22 CrossRef Medline
24. Hassankhani A, Ghozy S, Amoukhteh M, et al. **Long-term outcomes of the Woven EndoBridge device for treatment of intracranial aneurysms: a systematic review and meta-analysis.** *Interv Neuroradiol* 2023 Jun 26. [Epub ahead of print] CrossRef Medline
25. Pierot L, Szikora I, Barreau X, et al. **Aneurysm treatment with the Woven EndoBridge (WEB) device in the combined population of two prospective, multicenter series: 5-year follow-up.** *J Neurointerv Surg* 2023;15:552–57 CrossRef Medline
26. Dmytriw AA, Diestro JD, Dibas M, et al. **International study of intracranial aneurysm treatment using Woven EndoBridge: results of the WorldWideWEB Consortium.** *Stroke* 2022;53:e47–49 CrossRef Medline
27. Hecker C, Broussalis E, Pfaff JA, et al. **Comparison of the Contour Neurovascular System and Woven EndoBridge device for treatment of wide-necked cerebral aneurysms at a bifurcation or side-wall.** *J Neurosurg* 2023;139:563–72 CrossRef Medline
28. Pötin M, Biondi A, Sourour N, et al. **The LUNA aneurysm embolization system for intracranial aneurysm treatment: short-term, mid-term and long-term clinical and angiographic results.** *J Neurointerv Surg* 2018;10:e34 CrossRef Medline

# Evaluation of Conditions for the Development of Cryogenic Spinal Cord Injury Using a Canine Model: An Experimental Study on the Safety of Cryoablation for Metastatic Spinal Tumors

Motoya Kobayashi,  Satoshi Kato,  Satoru Demura,  Noriaki Yokogawa,  Akira Yokka,  Yusuke Nakade,  Ryohei Annen, Toshifumi Gabata, and Hiroyuki Tsuchiya



## ABSTRACT

**BACKGROUND AND PURPOSE:** Although the application of cryoablation to metastatic spinal tumors has been attempted, spinal cryoablation has the unique complication of cryogenic spinal cord injury. This study aimed to elucidate the conditions for the development of cryogenic spinal cord injury.

**MATERIALS AND METHODS:** Fifteen canines were used in this study. A metal probe was inserted into the 13th thoracic vertebral body. Cryoablation was performed for 10 minutes by freezing the probe in liquid nitrogen. The control canine underwent probe insertion only. Spinal cord monitoring, epidural temperature measurement, motor function assessment, and pathologic examination of the spinal cord were performed.

**RESULTS:** During the 10 minutes of cryoablation, the epidural temperature decreased and reached the lowest epidural temperature (LET) at the end of cryoablation. The LETs (degrees celsius [°C]) of each canine were −37, −30, −27, −8, −3, −2, 0, 1, 4, 8, 16, 18, 20, and 25, respectively. As the epidural temperature decreased, waveform amplitudes also decreased. At the end of cryoablation (10 minutes after the start of cryoablation), abnormal waves were observed in 92.9% (13/14) of canines. With epidural rewarming, the amplitude of the waveforms tended to recover. After epidural rewarming (2 hours after the start of cryoablation), abnormal waves were observed in 28.6% (4/14) of canines. The LETs (°C) of the canines with abnormal waves after epidural rewarming were −37, −30, −27, and −8. None of the canines with normal waves after epidural rewarming had any motor impairment. In contrast, all canines with remaining abnormal waves after epidural rewarming had motor impairment. In the pathologic assessment, cryogenic changes were found in canines with LETs (°C) of −37, −30, −27, −8, 0, and 1.

**CONCLUSIONS:** This study showed that 10-minute spinal cryoablation with LETs (°C) of −37, −30, −27, −8, 0, and 1 caused cryogenic spinal cord injury. There was no evidence of cryogenic spinal cord injury in canines with LET of ≥4°C. The epidural temperature threshold for cryogenic spinal cord injury is between 1 and 4°C, suggesting that the epidural temperature should be maintained above at least 4°C to prevent cryogenic spinal cord injury.

**ABBREVIATIONS:** CMAP = compound muscle action potential; LET = lowest epidural temperature; SCEP = spinal cord-evoked potential

Metastatic spinal tumors can significantly compromise a patient's quality of life because they can cause severe pain, pathologic vertebral fractures, and spinal cord compression.<sup>1</sup> Therefore, local control of metastatic spinal lesions is essential to maintain a patient's quality of life, particularly in patients with oligometastatic disease, which has a better prognosis.

Currently, the standard treatments for metastatic spinal tumors are based on surgical and radiation therapy. Surgical treatment is an

effective procedure that can remove or reduce metastatic spinal lesions and provide spinal stability with instrumentation.<sup>2</sup> However, surgery is usually not performed in patients with extensive metastatic disease burden and patients with poor functional status because of surgical invasiveness and the risk of developing complications.<sup>3</sup> Therefore, the current mainstay of treatment for metastatic spinal tumors is radiation therapy. In addition, the emerging technique of stereotactic body radiation therapy is reported to be effective even for radioresistant tumors.

Recently, cryoablation using extremely cold conditions to destroy tumor cells has been widely used in tumor treatment for


Received September 16, 2023; accepted after revision December 18.

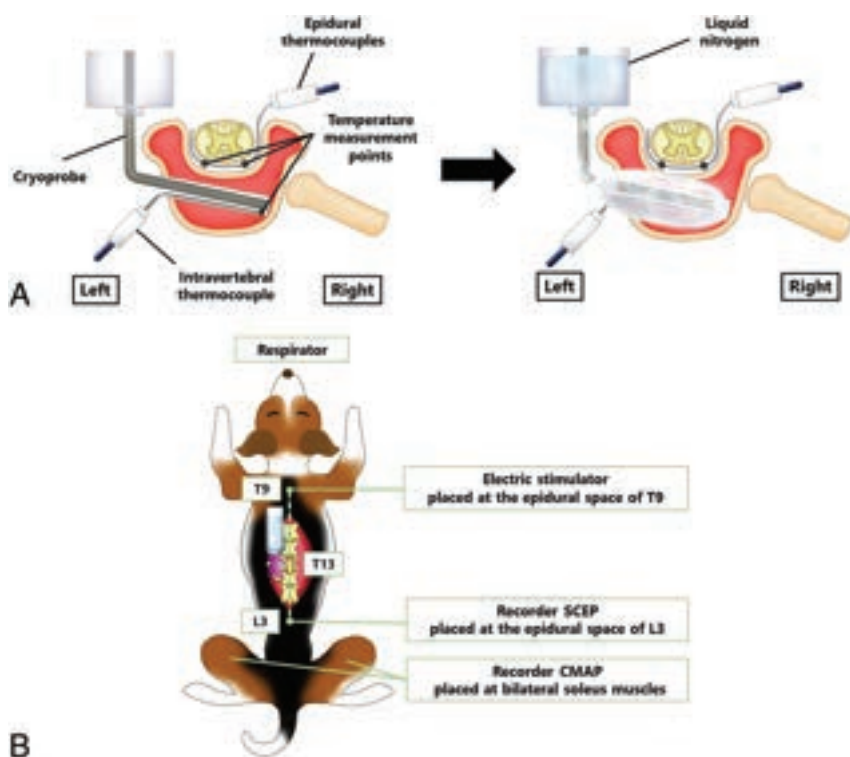
From the Department of Orthopedic Surgery (M.K., S.K., S.D., N.Y., R.A., H.T.), Graduate School of Medical Sciences, Department of Radiology (A.Y., T.G.) Graduate School of Medical Sciences, and Department of Clinical Laboratory (Y.N.), Kanazawa University Hospital, Kanazawa, Ishikawa, Japan.

This research was supported by a Grant-in-Aid for Scientific Research, grant No. 22K09303.

This funder did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Please address correspondence to Satoshi Kato, MD, PhD, Department of Orthopedic Surgery, Graduate School of Medical Sciences Kanazawa University 13-1 Takara-machi, Kanazawa, Ishikawa, 920-8641, Japan; e-mail: skato323@gmail.com

 Indicates article with online supplemental data.  
<http://dx.doi.org/10.3174/ajnr.A8151>



**FIG 1.** A, Details of cryoablation and thermal monitoring. B, Details of spinal cord monitoring.

various organs. Cryoablation is increasingly being performed in musculoskeletal metastases and has been reported to provide sustainable local control.<sup>4,5</sup> The advantages of cryoablation include direct visualization of the ablation zone, decreased periprocedural pain, and the ability to use multiple probes in variable configurations to create tailored ablation zones.<sup>6</sup> Moreover, the efficacy of cryoablation has been confirmed in radiation therapy-resistant tumors such as renal cancer.<sup>7</sup> In clinical practice, cryoablation has been used for metastatic spinal tumors.<sup>8</sup>

However, cryoablation of the spine may result in cryogenic spinal cord injury, a serious complication occurring when cooling causes irreversible damage to the spinal cord.<sup>9</sup> Moreover, the conditions for developing cryogenic spinal cord injury remain unclear, so there are limited indications for the use of cryoablation in metastatic spinal tumors. To establish the safety of spinal cryoablation, it is important to identify the conditions for developing cryogenic spinal cord injury. This study aimed to determine the threshold epidural temperature for developing cryogenic spinal cord injury using a canine spinal cryoablation model.

## MATERIALS AND METHODS

Experiments were performed on 15 canines (female Beagles) between 9 and 10 kg and 80 and 90 cm in length. All canines were acclimatized under the same conditions at the Institute for Animal Experiments, where they were housed individually in metal cages and provided with food and water ad libitum.

### Anesthetic Technique

After the administration of medetomidine (30  $\mu$ g/kg) and midazolam (0.3 mg/kg) into the paravertebral muscles, a catheter was

inserted into the radial vein for drug administration and fluid replacement. The canines were intubated and artificially ventilated with 50% nitrous oxide and 50% oxygen. They were placed in the prone position on an operating table under IV anesthesia with propofol (0.2 mg/kg/min). Muscle relaxants were not used. A cannula (2.0-mm diameter) was inserted into the right femoral artery for repeat blood sampling and continuous blood pressure monitoring. Body temperature was maintained between 36°C and 37°C using a heating pad as necessary and monitored with a rectal temperature probe. Arterial blood gases were measured at 90-minute intervals. The metabolic and respiratory acid-base balance was controlled with supplemental IV NaHCO<sub>3</sub> to maintain a pH of approximately 7.40. PaO<sub>2</sub> was maintained above 100 mm Hg, and PaCO<sub>2</sub> was maintained between 30 and 45 mm Hg by adjusting the respiratory volume. Concentrations of nitrous oxide and oxygen were constant throughout the procedure. Painful stimulation did

not cause any increase in blood pressure during steady-state anesthesia. Systemic blood pressure was maintained between 90 and 120 mm Hg for all procedures. In cases with a hypotensive trend, the infusion solution was increased or the propofol flow volume was slightly decreased.

### Freezing Device

A freezing device was created for our experiment on the basis of a previous report.<sup>10</sup> A 10-mL disposable syringe was incorporated into a 20-mL syringe to create a layer of air and prevent heat conduction from the syringe surface. A metal wire was placed at the tip of the syringe to form the cryoprobe. The metal wire tip was rapidly cooled via thermal conduction by injecting liquid nitrogen into the device (Fig 1A).

### Surgical and Cryoablation Procedures

After exposing the laminae from T11 to L2 through a posterior midline longitudinal incision, laminectomy at T13 and resection of the left proximal 13th rib were performed. After we exposed the left lateral aspect of the T13 vertebral body, a hole for the insertion of a cryoprobe was made using a 3-mm steel bar and a pedicle probe. A cryoprobe was inserted into the T13 vertebra, and 10-minute cryoablation was performed (Fig 1A). The intensity of cryoablation varied for each canine by changing the diameter and material of the metal probe (diameter range of probe: 1.2–2.0 mm; material of probe: copper or iron). An increase in probe size led to higher freezing intensity. Furthermore, changing the probe material from iron to copper amplified the freezing intensity due to differences in thermal conductivity (Online Supplemental Data). The control canine underwent probe insertion



as a sham operation. In this procedure, 1 cycle of 10 minutes of cooling and passive thawing was performed with epidural temperature measurement and spinal cord monitoring performed until 120 minutes after the start of cryoablation. Ablated vertebrae were evaluated by a CT apparatus for small experimental animals (Model LaTheta LCT-200; Hitachi-Aloka) to confirm that the probe hole was in the appropriate position 7 days after the cryoablation procedure.

**Thermal Monitoring**

Two spatula-type thermocouples (MF-SP-K; AS ONE; range: -50 degrees celsius [°C] to 200°C) were inserted into the ventral epidural space of T13 on the same plane as the cryoprobe, measuring the temperature of the bilateral ventral epidural space. In addition, a sheath-type thermocouple (S1K05 × 300-2; TOHO ELECTRONICS; range: -50°C to 200°C) was attached to the tip of the cryoprobe to monitor the intravertebral temperature (Fig 1A). Each temperature was recorded every 10 seconds until 2 hours after the start of cryoablation and recorded using a data logger (GL240-SD; GRAPTEC).

**Compound Muscle Action Potentials and Spinal Cord–Evoked Potentials**

Spinal cord function was evaluated using compound muscle action potentials (CMAPs)<sup>11,12</sup> and spinal cord–evoked potentials (SCEPs)<sup>11,13</sup> produced by spinal cord stimulation (Fig 1B).

CMAPs were recorded from the bilateral soleus muscles with each pair of needle electrodes (NE-110B; Nihon Kohden) using the tendon-belly method after spinal cord stimulation.<sup>14</sup> The electrodes for stimulation consisted of 18-ga tubes with 2 fine platinum coil terminals covered with polyethylene insulation (USY-100-2PMC; Unique Medical). The stimulating electrode was placed on the posterior midline of the dura mater at the level of T9, introduced from the laminectomy site at T13. The stimulus condition had a rectangular waveform, a pulse duration of 0.5ms, an interstimulus interval of 2.0 ms, and a train of 5 pulses per stimulation at 1 Hz to evoke muscle action potentials. The stimulation intensity ranged from 2 to 3 mA and was set to 10% above the level that elicited the maximal amplitude. A band-pass filter was selected at 10–3000 Hz, and the mean summation was 10. A Neuropack  $\Sigma$ EMG-SYSTEM (MEB-5504; Nihon Kohden) was used to record the signals. CMAPs were recorded before cryoablation (control) and at 2.5, 5, 7.5, 10, 30, 60, and 120 minutes after the start of cryoablation. Amplitude was measured between the maximum positive and negative peaks. Decreases of >70% of the control CMAP amplitudes were considered abnormal.<sup>15</sup> The CMAP reflects the activity of the corticospinal tract (mainly the lateral column) and the anterior horn cells of the spinal gray matter.

In the SCEP study, the stimulating electrode (USY-100-2PMC) was placed on the posterior midline of the dura mater at the level of T9, introduced from the laminectomy site at T13. A recording electrode (USY-100-2PMC) was placed on the posterior midline of the dura mater at the level of L3, introduced from the laminectomy site at T13. The spinal cord stimulation used a rectangular waveform for 0.2-ms duration at 5 Hz. Furthermore, stimulation intensity was set at a supramaximal strength. The band-pass filter was set at 10–3000 Hz, and the mean summation

**Modified Tarlov scale**

Grade	
0	Complete paraplegia with no hind extremity function
I	Minor joint movements
II	Major joint movements
III	Animal can stand
IV	Animal can walk
V	Animal can climb a 20° inclined plane

at 50. A Neuropack  $\Sigma$ EMG-SYSTEM (MEB-5504) was used to record the signals. SCEPs were recorded before cryoablation (control) and at 2.5, 5, 7.5, 10, 30, 60, and 120 minutes after the start of cryoablation. Amplitude was measured between the maximum negative and positive peaks of the first potential. Decreases of >50% in control SCEP amplitudes were considered abnormal.<sup>13,16</sup> The SCEP mainly reflects the combined activity of the long tract in the dorsal and lateral columns.

**Postprocedural Neurologic Examination**

Postprocedural neurologic function was evaluated using a modified Tarlov scale<sup>17,18</sup> at 1 day and 1 week after the cryoablation procedure. The scale used is presented in the Table.

**Pathologic Examination**

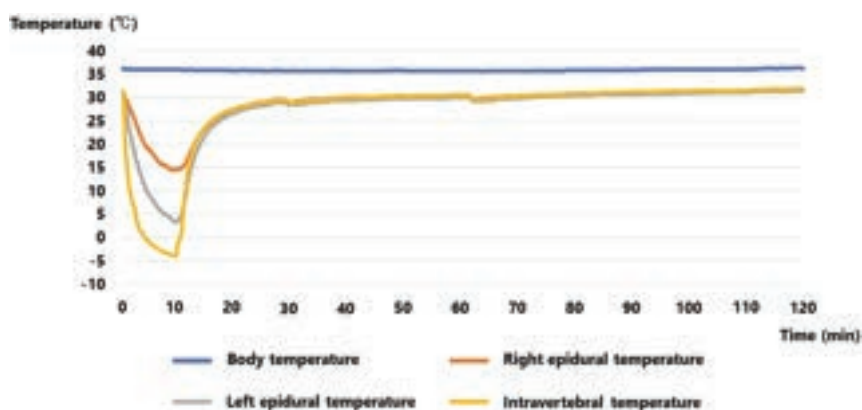
Seven days after the cryoablation procedure,<sup>18,19</sup> the canines were euthanized using an IV potassium chloride bolus. The T13 vertebra, including the spinal cord, was removed and fixed in 10% formalin. The mean length of the T13 vertebral body (maximal anterior-posterior diameter of the vertebral body), width (maximal transverse diameter of the vertebral body), and height (central height of the vertebral body) were 7.8 (SD, 0.56) mm, 15.2 (SD, 0.69) mm, and 15.3 (SD, 0.85) mm, respectively. The specimens were demineralized with ethylenediaminetetraacetic acid and stained with H&E for pathologic evaluation. Observations were performed using a microscope (BZ-9000; KEYENCE).

**RESULTS**

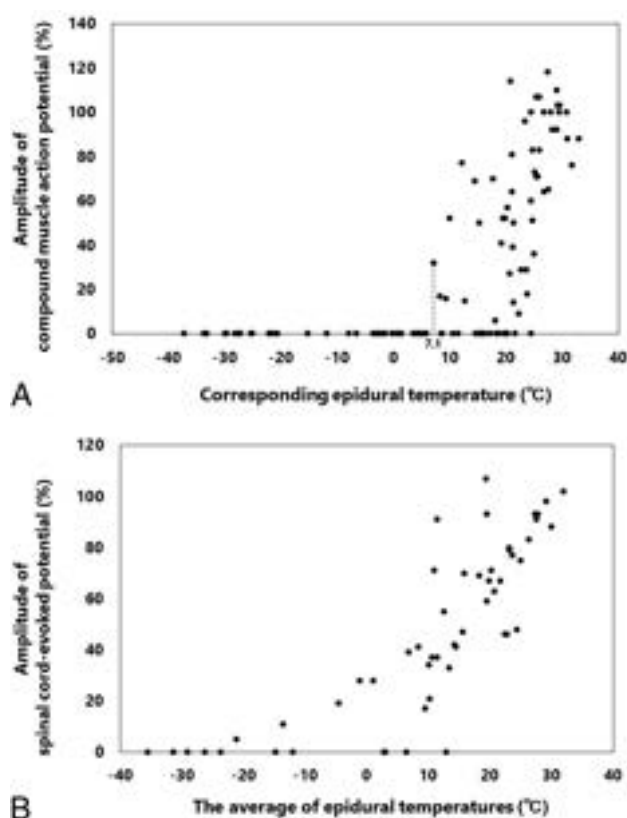
During the 10 minutes of cryoablation, the epidural temperature continued to decrease until the end of cryoablation. Overall, the left epidural measurement point had a lower temperature curve than the right. The left epidural temperature at the end of cryoablation (10 minutes after the start of cryoablation) was the lowest in each experiment. The lowest epidural temperatures (LETs) (°C) of each canine were -37, -30, -27, -8, -3, -2, 0, 1, 4, 8, 16, 18, 20, and 25. The epidural temperature increased after the end of cryoablation and almost recovered within 20 minutes after the end of cryoablation in all canines (Fig 2).

The relationship between the amplitude of spinal cord monitoring during cryoablation and the corresponding epidural temperature is illustrated in scatterplots (Fig 3). As the epidural temperature decreased, the waveform amplitudes of spinal cord monitoring also decreased. At the end of cryoablation, abnormal waves in CMAPs and SCEPs were observed in 85.7% (12/14) and 78.6% (11/14) of the canines, respectively.

Thirteen canines had abnormal waveforms in either the CMAPs or SCEPs at the end of cryoablation; however, the amplitude of the waveforms recovered in 9 canines with the rewarming



**FIG 2.** Temperature curves of the representative case with the lowest epidural temperature of 4°C (case 6).



**FIG 3.** A, Scatterplots illustrate the relationship between the left and right CMAPs at 2.5, 5, 7.5, and 10 minutes after the start of cryoablation and the corresponding epidural temperatures. B, Scatterplots illustrate the relationship between the SCEPs at 2.5, 5, 7.5, and 10 minutes after the start of cryoablation and an average of the corresponding left and right epidural temperatures.

of the epidural temperature (Fig 4A). After epidural rewarming (2 hours after the start of cryoablation), abnormal waves of CMAPs and SCEPs were observed in 28.6% (4/14) and 21.4% (3/14) of the canines, respectively. In cases with LETs of  $-37^{\circ}\text{C}$ ,  $-30^{\circ}\text{C}$ , and  $-27^{\circ}\text{C}$ , abnormal waves of CMAPs and SCEPs remained after epidural rewarming. In the canine with an LET of  $-8^{\circ}\text{C}$ , abnormal waves of CMAPs remained after epidural rewarming (Fig 4B).

All canines with normal waves on spinal cord monitoring after epidural rewarming had no hindlimb motor impairment with grade V on the modified Tarlov scale on day 1 and week 1 after the procedure. In contrast, all canines that showed abnormal waves in spinal cord monitoring after epidural rewarming had grade III or lower hindlimb motor impairment on the modified Tarlov scale on day 1 and week 1 after the procedure.

During the pathologic assessment of the spinal cord, cryogenic changes were detected in the canines with LETs ( $^{\circ}\text{C}$ ) of  $-37$ ,  $-30$ ,  $-27$ ,  $-8$ ,  $0$ , and  $1$ .

In the lesion of cryogenic change, there were characteristic findings of acute spinal cord injury, including the proliferation of glial cells, neuronal vacuolation (spongiosis), and dilated myelin sheaths containing hypereosinophilic swollen axons (spheroids) (Fig 5).

A summary of the results, including intravertebral and epidural temperatures at the end of cryoablation, amplitude change of spinal cord monitoring, postprocedural modified Tarlov scale, and pathologic findings, are described in the Online Supplemental Data.

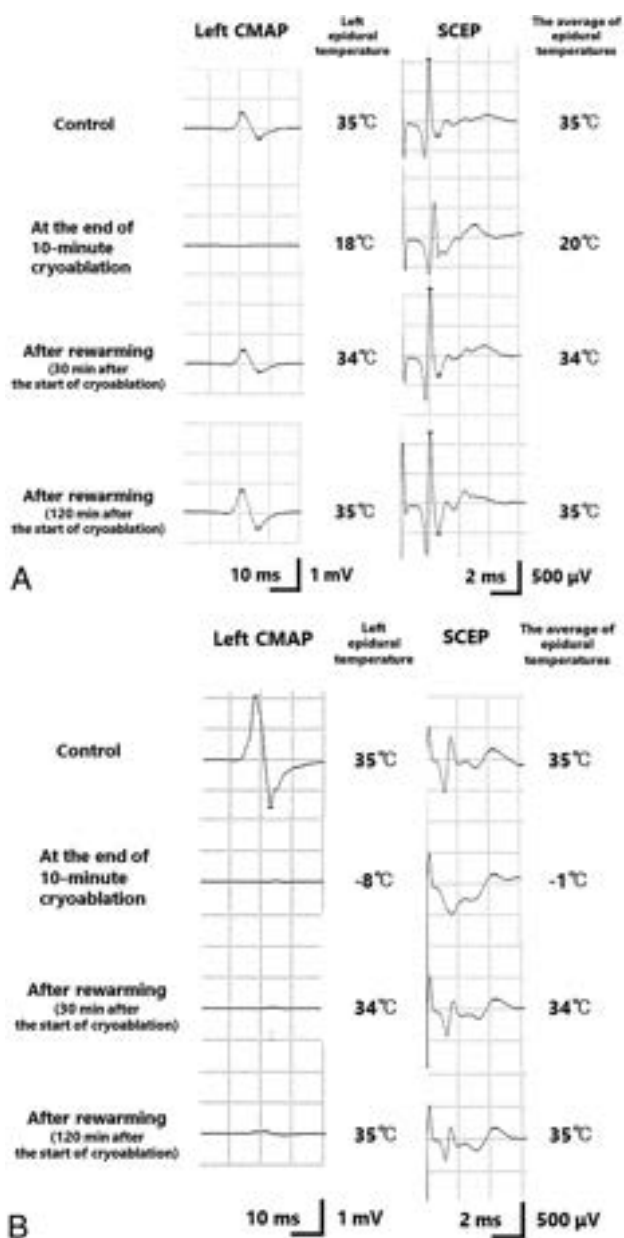
The sensitivity and specificity for detecting pathologic spinal cord injury in spinal cord monitoring combined with CMAPs and SCEPs during cryoablation were 100% (6/6) and 12.5% (1/8), respectively. After epidural rewarming, the sensitivity and specificity in spinal cord monitoring for detecting pathologic spinal cord injury were 66.7% (4/6) and 100% (8/8), respectively. The sensitivity and specificity of detecting postprocedural motor impairment in spinal cord monitoring after epidural rewarming were 100% (4/4) and 100% (10/10), respectively.

## DISCUSSION

This study was the first experimental study to focus on elucidating the conditions under which cryogenic spinal cord injury develops. This study included conventional assessments, such as an observation of motor function and pathologic evaluation of the spinal cord, as well as an electrophysiologic assessment of the spinal cord. The effects of cryoablation on the spinal cord were investigated in detail by real-time monitoring of the spinal cord function.

To date, 2 experimental studies have been conducted on the safety of spinal cryoablation. Wallace et al<sup>20</sup> examined the thermal protective capacity of the vertebral cortex and the accuracy of MR imaging for monitoring the ablation area using a sheep model. de Freitas et al<sup>19</sup> performed cryoablation of porcine vertebrae and verified the feasibility of cryoablation of the spine. However, these 2 studies did not examine the conditions under which cryogenic spinal cord injury develops.

In the current study, findings of cryogenic spinal cord injury were observed in all cases with an LET of  $-8^{\circ}\text{C}$  or lower. These results may be associated with extracellular ice formation. Water in the extracellular space crystallizes as the temperature drops to approximately  $-7^{\circ}\text{C}$ .<sup>21</sup> The extracellular fluid becomes hypertonic



**FIG 4.** Change of the CMAP and SCEP waveform in representative cases with the lowest epidural temperature of 18°C (case 3) (A) and -8°C (case 11) (B). A, The CMAP and SCEP waveforms decreased as the epidural temperature dropped. However, the waveforms recovered with rewarming of the epidural temperature. B, The CMAP and SCEP waveforms decreased as the epidural temperature dropped. The waveform of the CMAP did not completely recover after rewarming of the epidural temperature.

because of extracellular ice formation, causing cell dehydration and desiccation, which results in cell death. An experimental study using a rat model showed extensive axonal damage with loosening and edematous changes in interstitial tissue in the spinal cord after freezing at -8°C for 15 minutes,<sup>22</sup> supporting our hypothesis that the withdrawal of water because of extracellular ice formation is associated with cryogenic spinal cord injury. Hence, spinal cryoablation with an LET of -8°C or lower, with extracellular ice crystal formation, may have a considerable risk of cryogenic spinal cord injury.

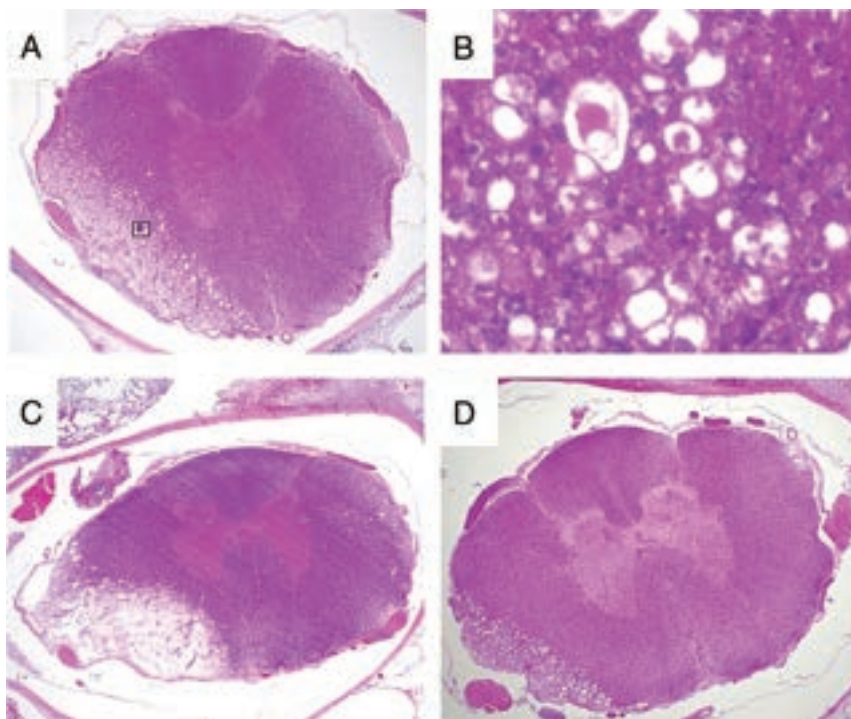
In all cases with an LET of -3°C or higher, intraprocedural spinal cord monitoring waveforms recovered after epidural rewarming, and the postprocedural motor function was normal. However, the pathologic evaluation revealed cryogenic spinal cord injury in 2 cases with LETs of 0°C and 1°C (Fig 5C, -D). In these 2 cases, cryogenic spinal cord injury was observed in a temperature range in which extracellular ice formation did not occur.

There is a similar phenomenon in the peripheral nerves called nonfreezing cold nerve injury, in which cryogenic nerve injury develops at a temperature above the freezing point. A rat model study demonstrated pathologic cryogenic nerve injury and an irreversible reduction in nerve blood flow after local cooling of the sciatic nerve for 3 hours at 1°C to 5°C, showing that ischemia plays a principal role in the nonfreezing cold nerve injury.<sup>23</sup> Moreover, the association between cooling with near-freezing temperatures and ischemia has been reported in the spinal cord. A canine experimental model reported that spinal cord blood flow decreased to 50% of the normothermic values during the cooling of the posterior dura mater to 3°C.<sup>24</sup> Considering these results, cooling to near-freezing temperatures may pose a risk of spinal cord injury including spinal cord ischemia, even if it does not reach the freezing range of extracellular fluid.

There were no cryogenic spinal cord injury findings in all cases with LETs of 4°C or higher. This finding was consistent with the results of a canine experimental model in which no cryogenic spinal cord injury was reported after selective spinal cord cooling at a dorsal column temperature of 5°C to 6°C for 100 minutes.<sup>25</sup> These results show that maintaining the epidural temperature above at least 4°C prevents cryogenic spinal cord injury during cryoablation for metastatic spinal tumors.

To the best of our knowledge, this study is the first reported experiment to visualize the electrical activity of the spinal cord during cryoablation using spinal cord monitoring. The application of spinal cord monitoring for cryoablation of metastatic spinal tumors has been attempted to compensate for the disadvantage of conventional ice ball monitoring, namely that only frozen areas can be visualized.<sup>8</sup> The findings of this study may contribute to establishing optimal protocols for spinal cord monitoring during cryoablation of metastatic spinal tumors.

The current study showed that the sensitivity for detecting cryogenic spinal cord injury in spinal cord monitoring combined with CMAPs and SCEPs during cryoablation was relatively high, 100% (6/6), and there were no false-negatives. Therefore, if the normal waveform of spinal cord monitoring is confirmed during cryoablation, the development of cryogenic spinal cord injury can be excluded, providing useful intraoperative information. However, the specificity for detecting cryogenic spinal cord injury in spinal cord monitoring during cryoablation was low at 12.5% (1/8), because spinal cord monitoring during cryoablation shows transient waveform reduction because of cooling. Hence, setting a lower threshold for abnormal waveforms should be considered when using spinal cord monitoring to reduce the rate of false-positives. In the present study, if even the slightest waveform of CMAP was observed during cryoablation, the ventral epidural temperature was confirmed to be at least 7.1°C, with a low risk of developing cryogenic spinal cord injury (Fig 3A). Therefore, during cryoablation, defining waveform disappearance as an abnormality in CMAPs may be practical.



**FIG 5.** The pathologic findings of the spinal cord in H&E staining. *A*, The pathologic specimen of the spinal cord in the representative case with the lowest epidural temperature of  $-8^{\circ}\text{C}$  (case 11) illustrates the cryogenic change in the left lateral and anterior columns ( $2\times$  magnification). *B*, In lesions of cryogenic change, there is the proliferation of glial cells, neuronal vacuolation (spongiosis), and dilated myelin sheaths that contain hypereosinophilic swollen axons (spheroids) ( $60\times$  magnification). *C* and *D*, The pathologic specimens of the spinal cord in cases 7 (*C*) and 8 (*D*). In these 2 cases with the lowest epidural temperatures of  $1^{\circ}\text{C}$  (case 7) and  $0^{\circ}\text{C}$  (case 8), the pathologic evaluation shows cryogenic spinal cord injuries, though spinal cord monitoring after epidural rewarming and the postprocedural motor function were normal. The specimens of the spinal cord in cases 7 (*C*) and 8 (*D*) illustrate the cryogenic changes in the left lateral and anterior columns ( $2\times$  magnification).

In contrast, after epidural rewarming, spinal cord monitoring showed better diagnostic accuracy for detecting cryogenic spinal cord injury (sensitivity: 66.7% [4/6]; specificity: 100% [8/8]). However, 2 false-negative cases (cases 7 and 8) resulted in a slightly low sensitivity. In both false-negative cases, most areas of cryogenic injury were located in the anterolateral part of the spinal cord (Fig 5C, -D), possibly making detecting the cryogenic damage difficult in spinal cord monitoring combined with CMAPs (which mainly reflect electrical activity in the lateral column of the spinal cord) and SCEPs (which mainly reflect electrical activity in the lateral and posterior columns of the spinal cord). Additionally, as shown in case 11 in which the waveform recovered after epidural rewarming in SCEPs but not CMAPs, SCEPs are considered to have a lower LET cutoff value for waveform recovery after epidural rewarming than CMAPs. This finding may be because CMAPs are generally the most sensitive method among the various spinal cord monitoring procedures.<sup>26</sup> Another possible reason may be that SCEPs mainly reflect electrical activity in the lateral and posterior columns of the spinal cord, making it difficult to detect cryogenic damage in the anterior part of the spinal cord accurately, compared with CMAPs.

Notably, for detecting postprocedural motor impairment, the sensitivity and specificity in spinal cord monitoring after epidural

rewarming were quite high at 100% (4/4) and 100% (10/10), respectively. Intraoperative prediction for postprocedural motor impairment is crucial because motor impairment can directly affect activities of daily living and quality of life.<sup>27</sup> Using spinal cord monitoring would contribute to preventing postprocedural motor impairment, which is associated with great benefits in cryoablation for metastatic spinal tumors.

Although spinal cord monitoring is useful for detecting cryogenic spinal cord injury, it has the disadvantage of many false-positives at low epidural temperatures during cryoablation. Therefore, if a certain degree of epidural temperature reduction is expected, such as in cryoablation of spinal metastases surrounding the spinal cord, the combined use of epidural temperature measurement would be desirable. Previous studies have demonstrated the feasibility and safety of thermal ablation with real-time monitoring of the epidural temperature for treating spinal metastases adjacent to the spinal cord.<sup>28,29</sup> A thermocouple in contact with the spinal cord provides continuous thermal monitoring of the surrounding spinal cord. Precise prediction of the development of spinal cord injury is possible by combining this real-time thermal monitoring

with the findings on the conditions for the development of cryogenic spinal cord injury.

However, accurate placement of the thermocouple for measuring epidural temperature is challenging using a percutaneous procedure. Additionally, the preventive measure applied to avoid cryogenic spinal cord injury, such as carbon dioxide epidural injection, may interfere with the accuracy of the measurement of epidural temperature. A previous study recommended the placement of several thermocouples, because a single thermocouple may not be able to reflect the area of the largest temperature change.<sup>30</sup> On the contrary, when cryoablation is combined with open decompression surgery for spinal metastases, thermocouples and insulation material for protecting the spinal cord can be directly placed in an appropriate epidural position because the spinal cord is exposed intraoperatively. We believe that this combined therapy can not only enhance local tumor control and functional prognosis but also improve the safety of the cryoablation procedure.

This study had several limitations. First, only 1 cycle of cooling and thawing was performed, and 2 cycles of cooling and thawing are commonly performed in clinical spinal cryoablation.<sup>8</sup> We focused on 1 cycle of cooling and thawing in this study because in an experiment with 2 cycles, we cannot determine



whether the first, second, or both cycles caused the cryogenic spinal cord injury, making it difficult to identify the accurate epidural temperature at which cryogenic spinal cord injury develops. A future study is required investigating the relationship between the cooling and thawing cycles and the conditions for the development of cryogenic spinal cord injury. Second, this study evaluated the LET but did not examine the cooling rate. The cooling rate is a factor that determines the degree of cryogenic tissue damage,<sup>31</sup> and this should be assessed in future studies. Third, this study sample was not strictly a spinal cryoablation model. The intravertebral temperature must be below  $-40^{\circ}\text{C}$ , when cryogenic tumor death occurs, to establish an accurate spinal cryoablation model.<sup>32</sup> However, to prepare the cryoablation models with various ranges of epidural temperatures, it was necessary to include the models with intravertebral temperatures above  $-40^{\circ}\text{C}$ , because creating a significant difference between the intravertebral and epidural temperatures was difficult due to the small size of the canine model in this study. Fourth, healthy vertebrae were used for this cryoablation model. The cooling distribution may be different between healthy vertebrae and tumor vertebrae due to mixed tissue with tumor and bone marrow, lytic destruction, and extraosseous extension into the canal. This limitation needs to be considered when applying the results of the current study to cases of spinal metastasis in clinical practice. Fifth, in this experiment, the lateral approach was used for probe placement because with the use of the transpedicular approach, there was a significant risk of probe deviation into the spinal canal from the canine's thin pedicle. This experimental system was designed to focus on investigating the conditions in which the cryogenic spinal cord injury develops, and this experimental system does not directly reflect clinical spinal cryoablation in which the transpedicular approach is commonly performed.

Despite these limitations, the findings of this study on the conditions for the development of cryogenic spinal cord injury contribute to establishing suitable monitoring methods and appropriate preventive measures for cryogenic spinal cord injury.

## CONCLUSIONS

This study showed that spinal cryoablation with a duration of 10 minutes and LETs ( $^{\circ}\text{C}$ ) of  $-37$ ,  $-30$ ,  $-27$ ,  $-8$ ,  $0$ , and  $1$  caused cryogenic spinal cord injury. In contrast, there were no cryogenic spinal cord injury findings in the cases with LETs of  $4^{\circ}\text{C}$  or higher. Therefore, the epidural temperature threshold for cryogenic spinal cord injury was considered between  $1^{\circ}\text{C}$  and  $4^{\circ}\text{C}$ , suggesting that the epidural temperature should be maintained above  $4^{\circ}\text{C}$  to prevent cryogenic spinal cord injury during cryoablation for metastatic spinal tumors.

## ACKNOWLEDGMENTS

We are grateful to Suga Chikazawa for her assistance with the pathologic examination. We would like to thank Yoshihiro Mozumi for technical assistance with the experiments. We are grateful to Kotoe Kobayashi for her assistance with creating the figures and to the members of the Institute for Animal Experiments for their important contributions to the experiments.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

1. Nater A, Martin AR, Sahgal A, et al. **Symptomatic spinal metastasis: a systematic literature review of the preoperative prognostic factors for survival, neurological, functional and quality of life in surgically treated patients and methodological recommendations for prognostic studies.** *PLoS One* 2017;12:e0171507 CrossRef Medline
2. Laufer I, Rubin DG, Lis E, et al. **The NOMS framework: approach to the treatment of spinal metastatic tumors.** *Oncologist* 2013;18:744–51 CrossRef Medline
3. Tarawneh AM, Pasku D, Quraishi NA. **Surgical complications and re-operation rates in spinal metastases surgery: a systematic review.** *Eur Spine J* 2021;30:2791–99 CrossRef Medline
4. Callstrom MR, Charboneau JW, Goetz MP, et al. **Image-guided ablation of painful metastatic bone tumors: a new and effective approach to a difficult problem.** *Skeletal Radiol* 2006;35:1–15 CrossRef Medline
5. Kurup AN, Morris JM, Callstrom MR. **Ablation of musculoskeletal metastases.** *AJR Am J Roentgenol* 2017;209:713–21 CrossRef Medline
6. Kwak K, Yu B, Lewandowski RJ, et al. **Recent progress in cryoablation cancer therapy and nanoparticles mediated cryoablation.** *Theranostics* 2022;12:2175–204 CrossRef Medline
7. Ismail M, Nielsen TK, Lagerveld B, et al. **Renal cryoablation: multi-disciplinary, collaborative and perspective approach.** *Cryobiology* 2018;83:90–94 CrossRef Medline
8. Tomasian A, Wallace A, Northrup B, et al. **Spine cryoablation: pain palliation and local tumor control for vertebral metastases.** *AJNR Am J Neuroradiol* 2016;37:189–95 CrossRef Medline
9. Tan LA, Deutsch H. **Thermal injury of thoracic spinal cord after percutaneous cryoablation of spinal tumor—when needles are more dangerous than the knife.** *Br J Neurosurg* 2015;29:443 CrossRef Medline
10. Annen R, Kato S, Demura S, et al. **Tumor-specific immunoenhancing effects after local cryoablation for metastatic bone tumor in a mouse model.** *Int J Mol Sci* 2022;23:9445 CrossRef Medline
11. Iwasaki H, Tamaki T, Yoshida M, et al. **Efficacy and limitations of current methods of intraoperative spinal cord monitoring.** *J Orthop Sci* 2003;8:635–42 CrossRef Medline
12. Machida M, Weinstein SL, Yamada T, et al. **Monitoring of motor action potentials after stimulation of the spinal cord.** *J Bone Joint Surg Am* 1988;70:911–18 CrossRef Medline
13. Tamaki T, Noguchi T, Takano H, et al. **Spinal cord monitoring as a clinical utilization of the spinal evoked potential.** *Clin Orthop Relat Res* 1984;58–64 CrossRef Medline
14. Nakagawa Y, Tamaki T, Yamada H, et al. **Discrepancy between decreases in the amplitude of compound muscle action potential and loss of motor function caused by ischemic and compressive insults to the spinal cord.** *J Orthop Sci* 2002;7:102–10 CrossRef Medline
15. Kobayashi S, Matsuyama Y, Shinomiya K, et al. **A new alarm point of transcranial electrical stimulation motor evoked potentials for intraoperative spinal cord monitoring: a prospective multicenter study from the Spinal Cord Monitoring Working Group of the Japanese Society for Spine Surgery and Related Research.** *J Neurosurg Spine* 2014;20:102–07 CrossRef Medline
16. Shinomiya K, Furuya K, Sato R, et al. **Electrophysiologic diagnosis of cervical OPLL myelopathy using evoked spinal cord potentials.** *Spine (Phila Pa 1976)* 1988;13:1225–33 CrossRef Medline
17. Tarlov IM, Klinger H. **Spinal cord compression studies, II: time limits for recovery after acute compression in dogs.** *AMA Arch Neurol Psychiatry* 1954;71:271–90 Medline
18. Kato S, Kawahara N, Tomita K, et al. **Effects on spinal cord blood flow and neurologic function secondary to interruption of bilateral segmental arteries which supply the artery of Adamkiewicz: an experimental study using a dog model.** *Spine (Phila Pa 1976)* 2008;33:1533–41 CrossRef Medline
19. de Freitas RM, Andrade CS, Caldas JG, et al. **Image-guided cryoablation of the spine in a swine model: clinical, radiological, and**



- pathological findings with light and electron microscopy. *Cardiovasc Intervent Radiol* 2015;38:1261–70 CrossRef Medline
20. Wallace AN, Hillen TJ, Friedman MV, et al. **Percutaneous spinal ablation in a sheep model: protective capacity of an intact cortex, correlation of ablation parameters with ablation zone size, and correlation of postablation MRI and pathologic findings.** *AJNR Am J Neuroradiol* 2017;38:1653–59 CrossRef Medline
  21. Gage AA, Baust J. **Mechanisms of tissue injury in cryosurgery.** *Cryobiology* 1998;37:171–86 CrossRef Medline
  22. Collins GH, West NR, Parmely JD, et al. **The histopathology of freezing injury to the rat spinal cord: a light microscope study, I: early degenerative changes.** *J Neuropathol Exp Neurol* 1986;45:721–41 CrossRef Medline
  23. Jia J, Pollock M. **The pathogenesis of non-freezing cold nerve injury: observations in the rat.** *Brain* 1997;120:631–46 CrossRef Medline
  24. Hansebout RR, Lamont RN, Kamath MV. **The effects of local cooling on canine spinal cord blood flow.** *Can J Neurol Sci* 1985;12:83–87 CrossRef Medline
  25. Romero-Sierra C, Hansebout R, Sierhuis A, et al. **A new method for localised spinal-cord cooling.** *Med Biol Eng* 1974;12:188–93 CrossRef Medline
  26. Ito Z, Matsuyama Y, Ando M, et al. **What is the best multimodality combination for intraoperative spinal cord monitoring of motor function? A multicenter study by the monitoring committee of the Japanese Society for Spine Surgery and Related Research.** *Global Spine J* 2016;6:234–41 CrossRef Medline
  27. Brekke MF, la Cour K, Brandt Å, et al. **The association between ADL ability and quality of life among people with advanced cancer.** *Occup Ther Int* 2019;2019:2629673 CrossRef Medline
  28. Nakatsuka A, Yamakado K, Takaki H, et al. **Percutaneous radiofrequency ablation of painful spinal tumors adjacent to the spinal cord with real-time monitoring of spinal canal temperature: a prospective study.** *Cardiovasc Intervent Radiol* 2009;32:70–75 CrossRef Medline
  29. Buy X, Tok CH, Szwarc D, et al. **Thermal protection during percutaneous thermal ablation procedures: interest of carbon dioxide dissection and temperature monitoring.** *Cardiovasc Intervent Radiol* 2009;32:529–34 CrossRef Medline
  30. Kastler A, Krainik A, Sakhri L, et al. **Feasibility of real-time intra-procedural temperature control during bone metastasis thermal microwave ablation: a bicentric retrospective study.** *J Vasc Interv Radiol* 2017;28:366–71 CrossRef Medline
  31. Bischof J, Christov K, Rubinsky B. **A morphological study of cooling rate response in normal and neoplastic human liver tissue: cryosurgical implications.** *Cryobiology* 1993;30:482–92 CrossRef Medline
  32. Gage AA. **What temperature is lethal for cells?** *J Dermatol Surg Oncol* 1979;5:459–60, 464 CrossRef Medline

# Automated Detection of Cervical Spinal Stenosis and Cord Compression via Vision Transformer and Rules-Based Classification

David L. Payne, Xuan Xu, Farshid Faraji, Kevin John, Katherine Ferra Pradas, Vahni Vishala Bernard, Lev Bangiyev, and Prateek Prasanna

## ABSTRACT

**BACKGROUND AND PURPOSE:** Cervical spinal cord compression, defined as spinal cord deformity and severe narrowing of the spinal canal in the cervical region, can lead to severe clinical consequences, including intractable pain, sensory disturbance, paralysis, and even death, and may require emergent intervention to prevent negative outcomes. Despite the critical nature of cord compression, no automated tool is available to alert clinical radiologists to the presence of such findings. This study aims to demonstrate the ability of a vision transformer (ViT) model for the accurate detection of cervical cord compression.

**MATERIALS AND METHODS:** A clinically diverse cohort of 142 cervical spine MRIs was identified, 34% of which were normal or had mild stenosis, 31% with moderate stenosis, and 35% with cord compression. Utilizing gradient-echo images, slices were labeled as no cord compression/mild stenosis, moderate stenosis, or severe stenosis/cord compression. Segmentation of the spinal canal was performed and confirmed by neuroradiology faculty. A pretrained ViT model was fine-tuned to predict section-level severity by using a train:validation:test split of 60:20:20. Each examination was assigned an overall severity based on the highest level of section severity, with an examination labeled as positive for cord compression if  $\geq 1$  section was predicted in the severe category. Additionally, 2 convolutional neural network (CNN) models (ResNet50, DenseNet121) were tested in the same manner.

**RESULTS:** The ViT model outperformed both CNN models at the section level, achieving section-level accuracy of 82%, compared with 72% and 78% for ResNet and DenseNet121, respectively. ViT patient-level classification achieved accuracy of 93%, sensitivity of 0.90, positive predictive value of 0.90, specificity of 0.95, and negative predictive value of 0.95. Receiver operating characteristic area under the curve was greater for ViT than either CNN.

**CONCLUSIONS:** This classification approach using a ViT model and rules-based classification accurately detects the presence of cervical spinal cord compression at the patient level. In this study, the ViT model outperformed both conventional CNN approaches at the section and patient levels. If implemented into the clinical setting, such a tool may streamline neuroradiology workflow, improving efficiency and consistency.

**ABBREVIATIONS:** AI = artificial intelligence; AUC = area under the curve; CNN = convolutional neural network; ER = emergency room; GRE = gradient-echo; NPV = negative predictive value; PPV = positive predictive value; ViT = vision transformer

Cervical spinal cord compression, a clinical entity which results from severe narrowing of the spinal canal by an extrinsic source, is widely recognized as a clinical emergency, and can lead to severe clinical consequences, including muscular weakness, pain, sensory disturbances, paralysis, and even death. Early medical and surgical intervention in the setting of severe spinal canal

stenosis and spinal cord compression has been shown to improve long-term neurologic outcomes.<sup>1</sup> Although CT has its own advantages in evaluation of the spine in terms of bone integrity, MR imaging is the reference standard for evaluation of the soft tissues, intervertebral discs, and spinal cord. MR imaging examinations of the cervical spine are increasingly utilized in the acute or emergent setting and are frequently interpreted by general radiologists.<sup>2,3</sup>

A continuous growth in frequency of cross-sectional imaging performed has led to explosive growth in demands on interpreting radiologists, who are now often required to interpret an average of one image every 3–4 seconds to meet workload demands.<sup>4</sup> There is extensive academic research and private investment aiming to utilize artificial intelligence (AI) as a tool to augment radiologist productivity, improve accuracy, and reduce cognitive strain.

Received October 3, 2023; accepted after revision December 15.

From the Departments of Radiology (D.L.P., F.F., K.J., K.F.P., V.V.B., L.B.), and Biomedical Informatics (D.L.P., X.X., F.F., K.J., P.P.), Stony Brook University, Stony Brook, New York.

This work was partially supported by NIH 1R21CA258493-01A. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Please address correspondence to David L. Payne, MD, 101 Nicolls Rd, Stony Brook, NY 11794; e-mail: David.payne@stonybrookmedicine.edu; @DavidLPayneMD  
<https://dx.doi.org/10.3174/ajnr.A8141>

Additionally, radiology AI has shown promise as a triaging tool to preview examinations as soon as they are performed, reorganizing these examinations to the top of the clinical workload and alerting radiologists to possible abnormal findings.<sup>5-7</sup>

Previous work in this domain has demonstrated the ability of convolutional neural network (CNN) algorithms to detect cervical cord compression at the section level in patients with known diagnosis of cervical myelopathy.<sup>8</sup> Additional studies have shown ability of CNNs to detect patient-level spinal pathology; however, most of these have focused exclusively on degenerative disc disease, the lumbar spine, or have utilized nonclinically representative cohorts.<sup>9-14</sup> To our knowledge, no study has leveraged individual section-level predictions to generate an overall patient-level severity prediction, a method of classification that mimics the manner in which a practicing radiologist would view cervical spine MR imaging.

Additionally, while previously published work has demonstrated impressive results utilizing CNNs, none to our knowledge has leveraged vision transformers (ViTs), a cutting-edge deep learning architecture, for the detection of cervical spinal pathology. ViTs have gained prominence in the medical/radiologic image analysis field since 2020 for their facility in segmentation and classification, frequently demonstrating comparable or superior results to CNNs.<sup>15-19</sup> In addition to their measurable performance, ViTs hold great promise as a tool in medical imaging analysis because of their ability to capture long-range, global dependencies within an image compared with the local receptive fields of CNNs.<sup>20,21</sup>

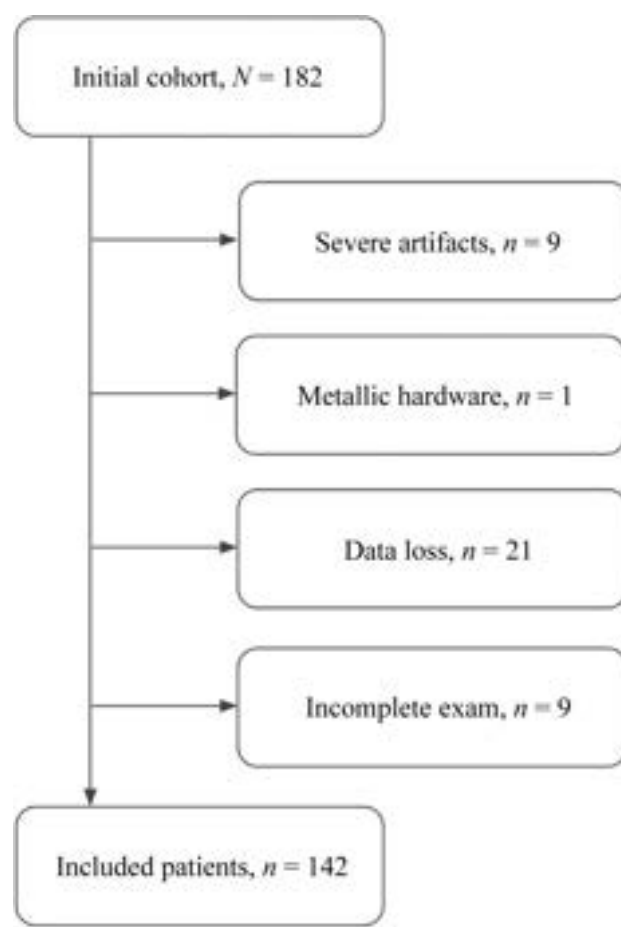
ViT brought forth a novel approach by considering images as sequences comprising patches, subsequently processing these patches by using transformer blocks. The self-attention mechanism present in ViT enables the capture of comprehensive global contextual information by facilitating long-range interactions between different regions or patches within the images.<sup>22</sup> As such, ViT offers a promising solution for encoding meaningful and intricate representations of complex structures, wherein both shape and spatial arrangement are pivotal elements in cord compression classification.

Developing a reliable and validated tool to detect cord compression could meaningfully streamline the work of practicing emergency room (ER), general, and neuroradiologists by helping to triage their workload. Without such a tool, the practicing radiologist may have a cervical spine MR imaging on their workload with critical cord compression but no warning until they physically open the examination. This study demonstrates the ability of a ViT combined with simple rules-based classification to accurately evaluate cervical spine MRIs for cord compression with high accuracy. This study also compares the performance of the proposed ViT model to 2 conventional CNN architectures.

## MATERIALS AND METHODS

### Patient Selection and Image Preprocessing

Following institutional review board approval for this retrospective study, an initial cohort was identified utilizing an internal proprietary search engine that indexes radiology reports performed at our institution. Search terms included “MR imaging cervical spine,” “severe stenosis,” “cord compression,” “moderate



**FIG 1.** Flow chart of patient inclusion.

stenosis,” and “mild stenosis.” Examination reports were then reviewed chronologically preceding from 2022 backward until a sufficient number of examinations was identified, culminating in a study period extending from 2017–2022.

To ensure a balanced representation of the disease severity categories at the patient level, the cohort was constructed such that 63 of the patients had radiologic reports describing normal examinations or mild stenosis, 64 with up to moderate stenosis at least 1 or more spinal level, and 55 had severe stenosis or cord compression at 1 or more levels, comprising a total of 182 studies.

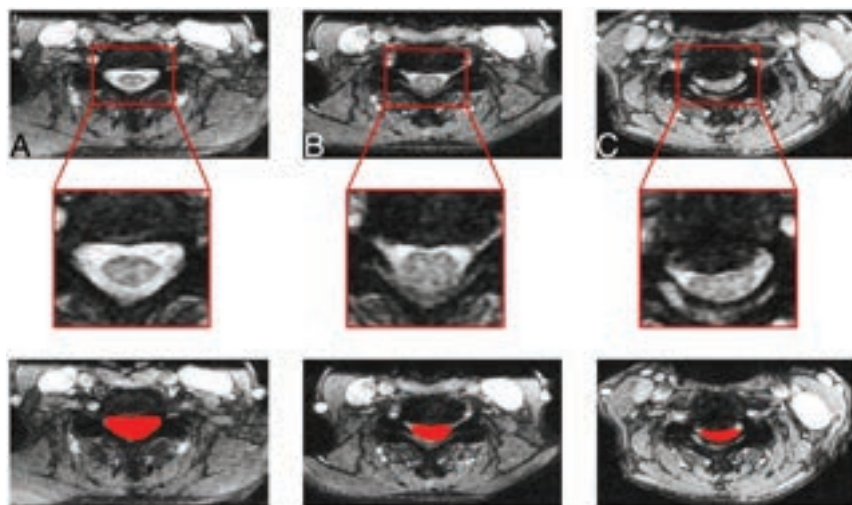
A total of 40 of these examinations were excluded secondary to severe artifacts (9 patients), presence of metallic hardware (1 patient), data quality (21 patients), or a lack of an axial gradient-echo (GRE) sequence (9 patients) (Fig 1). Of the resulting cohort of 142 patients, 34% were described in their final radiology reports as normal or with up to mild stenosis, 31% with up to moderate stenosis, and 35% with cord compression.

A total of 27% of studies were obtained on a 3T Siemens magnet, and the remainder were obtained on 1.5T GE Healthcare and Siemens magnets. Typical scanner parameters for the axial GRE sequence included TR of 538–761 ms, TE of 13 ms, flip angle of 30°, field of view of 165–180 mm, number of excitations of 1–3, section thickness of 3 mm, section gap of 0 mm, matrix size of 320–384 × 128–192, 34–64 slices, and sequence acquisition time

of 2.5–4.0 minutes. Gradient-echo imaging was selected as our sequence of interest because of prominent CSF flow artifact on a large number of spin-echo T2 images that were initially reviewed, as well as internal clinical protocols specific to our institution.

These studies were then anonymized and converted to NIfTI format. Axial gradient-echo slices were labeled as 0 = normal/mild stenosis, 1 = moderate stenosis, and 2 = cord compression utilizing a classification scheme simplified from Kang et al, and adapted to the axial plane.<sup>22</sup> Specifically, mild stenosis was defined as loss of subarachnoid space up to 50%, moderate stenosis with loss of subarachnoid space greater than

50%, and cord compression with complete loss of subarachnoid space and frank cord deformity, with or without T2 hyperintensity within the cord. Additionally, the spinal canal was manually segmented by using ITK-SNAP segmentation software (Fig 2). These tasks were performed by 3 senior radiology resident physicians (PGY-4 and PGY-5 levels) and further confirmed by senior neuroradiology faculty with > 15 years of experience. Section-level ground truth severity was initially judged by the resident team, with all cases reviewed and grading modified as needed by the above criteria by fellowship-trained neuroradiology faculty before any analysis was performed.

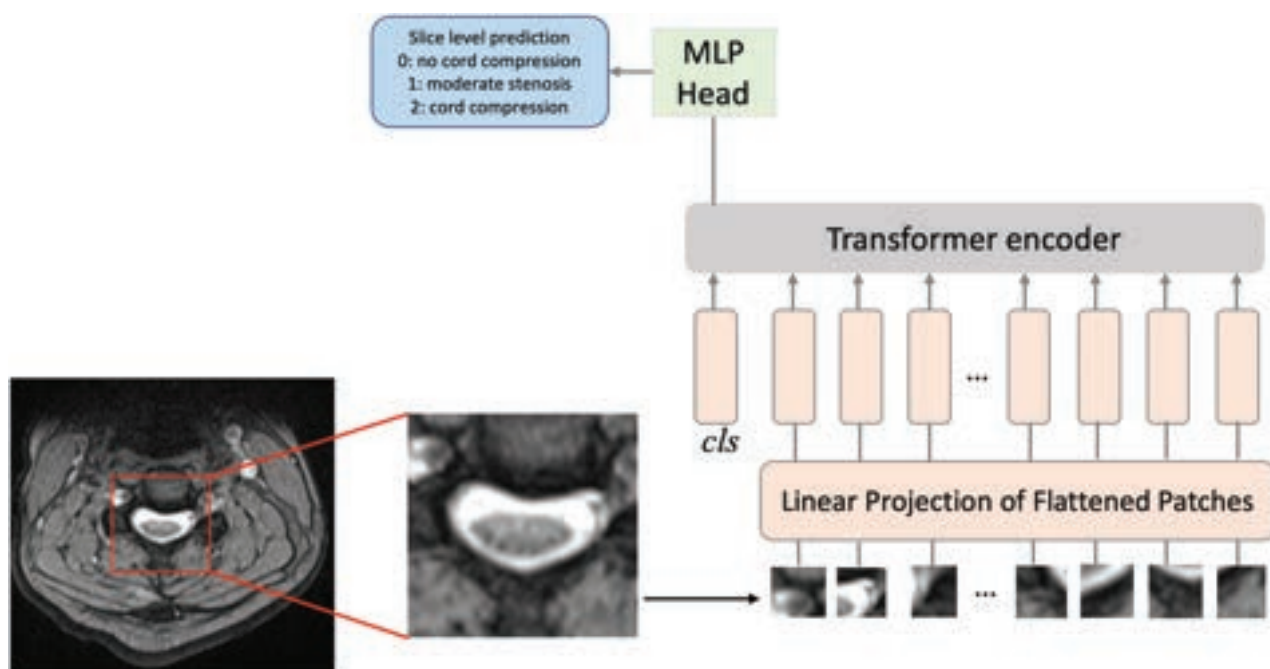


**FIG 2.** Representative GRE images of the cervical spine (top) with insets (middle) and examples of spinal canal segmentations (bottom) performed by trained radiology physicians demonstrating examples of no/mild stenosis (A), moderate stenosis (B), and severe stenosis/cord compression (C).

### ViT Model and Patient-Level Prediction

For our study, a ViT model pretrained on ImageNet was fine-tuned by using PyTorch (<https://pytorch.org/>) to predict section-level severity by using a train:validation:test split of 60:20:20, a batch size of 16, and 200 epochs.

Experiments were performed to assess model accuracy at 3-level prediction (normal/mild stenosis versus moderate stenosis versus cord compression) by using a forced square centered on the manual segmentation of the spinal canal as the area of interest (Fig 3). Patient-level severity was judged based on the highest degree of section severity within each examination, with a patient/examination was judged as having been categorized as “severe” if  $\geq 1$  section was predicted



**FIG 3.** Vision transformer analysis pipeline. Images are divided into flattened patches, which are then fed into the transformer encoder along with positional encoding. For each section, a prediction of severity is generated.

as having severe disease. Accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated based on a comparison of the patient-level predictions with the ground-truth patient-level severity.

In addition to the ViT model, our study also encompassed the evaluation of 2 popular CNN architectures, ResNet50 and DenseNet121. These models were chosen because of their recognized efficacy and widespread popularity within radiology AI, including within studies of cervical spine pathology.<sup>12,23</sup> Performance of the 3 models (ViT, ResNet, DenseNet) was compared, both at the section level and patient level.

RESULTS

Patient Characteristics

Patient characteristics are summarized in Table 1. No significant differences were present when comparing differences in patient age, sex, acute history of trauma, or clinical setting (emergency room versus inpatient versus outpatient) between the training/validation cohort and the testing cohort utilizing *t*-test for continuous and  $\chi^2$  for categoric variables (all *P* > .05, Table 1). The most common etiologies of trauma were fall and motor vehicle crashes. The most common presenting histories in patients

without history of acute trauma were degenerative disc disease, malignancy, concern for infection, upper extremity sensory/motor disturbances, and multiple sclerosis.

Section-Level Characteristics

To avoid overfitting, the model was trained and validated on a balanced subset of the data, with a roughly equivalent number of severe/moderate and normal/mild slices. The training/validation data set consisted of 716 normal/mild images, 496 moderate images, and 234 images with cord compression and a total of 113 examinations.

To better reflect real-world conditions, the testing set, which was balanced at the examination/patient level (10 cases with compression, 9 with moderate disease, 10 with mild/no disease), was left unbalanced at the section level. This left a data set with 1241 normal/mild slices, 99 moderate slices, and 25 severe slices.

Image and Patient-Level Experiment Results

ViT outperformed ResNet50 and DenseNet121 at the section level, achieving 82% accuracy in the classification of spinal stenosis and cord compression. By comparison, ResNet50 achieved section-level accuracy of 72% and DenseNet121 yielded 78% accuracy. Significance of these differences in accuracy was confirmed by using pair-wise comparison (ViT versus ResNet50, ViT versus DenseNet121, ResNet50 versus DenseNet121) using the McNemar test (*P* < .001 comparing ViT versus ResNet50, *P* = .008 comparing ViT versus DenseNet 121, *P* < .001 comparing ResNet50 versus DenseNet121).

ViT also outperformed the CNNs at patient-level prediction, with sensitivity of 0.90, PPV of 0.90, specificity of 0.95, and NPV value of 0.95. Receiver operating characteristic curves were generated for each model's patient-level classification accuracy, with area under the curve (AUC) of 0.92 for ViT, 0.71 for ResNet50, and 0.66 for DenseNet121 (Fig 4). Pair-wise comparison was

Table 1: Patient demographics and clinical setting in the training/validation and testing cohorts

Characteristic	Training/Validation (n = 113)	Testing (n = 29)
Demographics		
Male	56 (49.5%)	11 (37.9%)
Age	57.5	53.1
Clinical setting		
ER	64 (56.6%)	18 (62.1%)
Inpatient	16 (14.2%)	3 (10.3%)
Outpatient	33 (29.2%)	8 (27.6%)
History of trauma	49 (43.4%)	14 (48.3%)

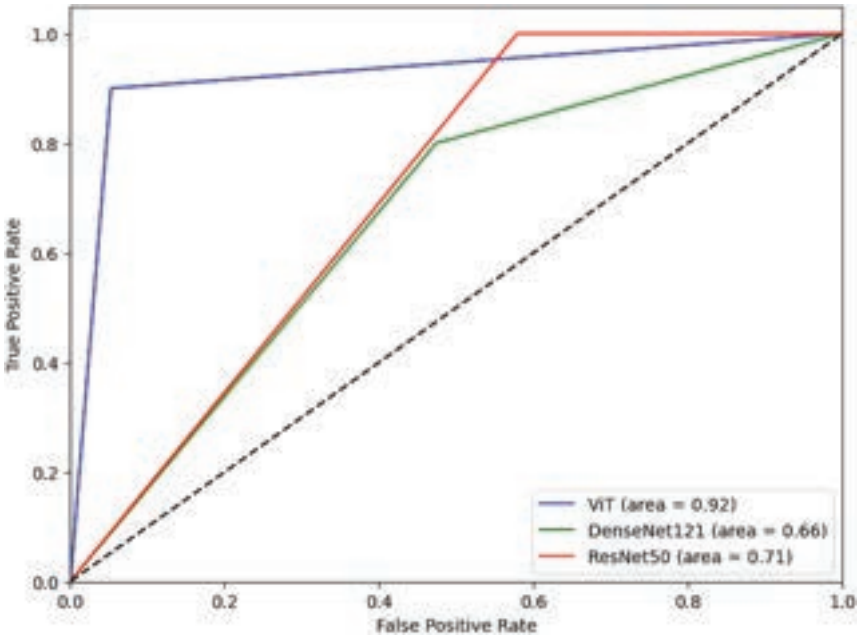
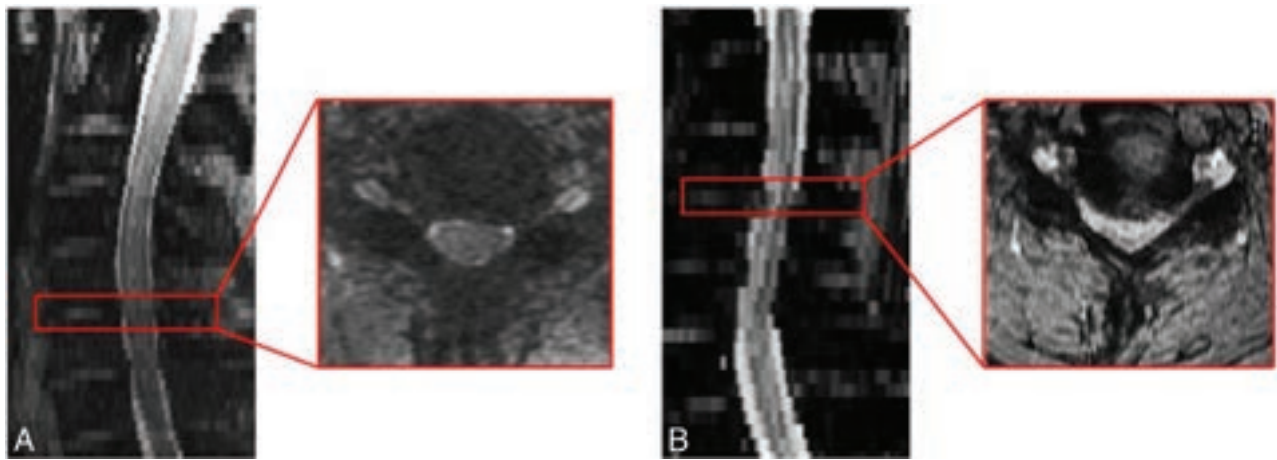


FIG 4. Receiver operating characteristic curves comparing patient-level performance of ViT, DenseNet121, and ResNet50 in detecting cord compression.





**FIG 5.** False-positive and -negative cases. *A*, False-positive: Diffuse congenital moderate stenosis with small superimposed annular bulge, incorrectly categorized as positive for cord compression. *B*, False-negative: Large annular bulge with severe spinal stenosis and cord compression and abnormal T2 hyperintense signal within the cord, incorrectly categorized as negative for cord compression.

**Table 2: Section-level and patient-level test results for ViT, ResNet50, and DenseNet121 models**

	ViT	ResNet50	DenseNet121
Section-level classification accuracy	82%	72%	78%
Patient-level accuracy	93%	62%	62%
Patient-level sensitivity	0.90	1.0	0.8
Patient-level specificity	0.95	0.42	0.52
Patient-level PPV	0.90	0.47	0.47
Patient-level NPV	0.95	1.0	0.83

performed for each model by using the DeLong test, which demonstrated significantly greater AUC of ViT compared with both DenseNet 121 and ResNet50 ( $P = .014$  comparing ViT versus ResNet50,  $P = .015$  comparing ViT versus DenseNet121), but no significant difference between ResNet50 and DenseNet121 ( $P = .66$ ).

Two cases within the testing set were misclassified: 1 falsely negative and the other falsely positive for cord compression (Fig 5). No differences in accuracy were observed when considering differences in sex, age, or clinical setting.

While the ResNet50 model demonstrated 100% sensitivity and NPV in patient-level classification, specificity and PPV were poor, at 0.42 and 0.47, respectively. DenseNet121-based patient-level classification underperformed both models at patient-level prediction, with sensitivity of 0.8, specificity of 0.52, PPV of 0.47, and NPV of 0.83. These results are summarized in Table 2.

Post hoc power analyses were performed for both section-level and patient-level classification by using commercially available sample size calculation software nQuery as well as the Cleveland Clinic Department of Quantitative Health Sciences Sample Size Calculator.<sup>24</sup> With both methods of power calculation, power met or exceeded 0.95 (both for section and patient classification), utilizing a null hypothesis of AUC equal to 0.5, analogous to random guessing. Additionally, 95% CIs were calculated for the ViT patient-level AUC of 0.92, yielding 0.80–1.00.

## DISCUSSION

In our study, we presented a ViT model to classify MR imaging images of the cervical spine as normal/mild stenosis, moderate

stenosis, or cord compression by using a diverse clinical cohort taken from emergency, inpatient, and routine outpatient settings. Based on the predictions of this model, we utilized a simple rules-based approach to label entire examinations as positive for cord compression or negative for cord compression. Based on this methodology, we achieved a high accuracy, sensitivity, specificity, PPV, and NPV. In our study, ViT outperformed 2 popular CNN architectures. To our knowledge, our work presented the first attempt at approximating the pattern of human interpretation of MR imaging cervical spine examinations by utilizing section-level predictions to generate an overall patient-level severity classification for cervical spinal pathology. Additionally, no other study has utilized ViT models for such a task within the cervical spine. We believe the results of our experiments are meaningful for several reasons.

First, while there is tremendous interest in utilizing AI-based classification tools in neuroimaging, applications in cervical spine pathology as depicted by MR imaging are relatively limited. Most AI-based research in spinal neuroimaging and pathology relates to lumbar spinal disease.<sup>11,13</sup> There are also several published studies demonstrating the ability of deep learning algorithms to detect cervical spinal fractures via CT.<sup>25,26</sup>

The few studies that exist relating to cervical spinal cord pathology differ from our study in meaningful ways. For example, Merali et al<sup>8</sup> demonstrate impressive performance of a deep learning model at section-level prediction within a cohort of patients with known cervical myelopathy. While such approaches have yielded excellent performance, they may not retain their accuracy when applied to our cohort, which represents a broader array of etiologies for spinal stenosis and cord compression. Tamai et al,<sup>9</sup> Hopkins et al,<sup>10</sup> and Ma et al<sup>12</sup> also differ from our methodology substantially by focusing solely on degenerative disc disease, section-level prediction, classifying images/patients into categories different from our study, utilizing a different deep learning methodology, and/or utilizing a cohort of patients with known cervical myelopathy.

By contrast, our study leverages a vision transformer model combined with simple rules-based logic dictating that an examination with predicted cord compression at any level equals an

examination positive for suspected cord compression. This methodology is inspired by working clinically with commercially available intracranial hemorrhage detection software that labels an examination with “suspected hemorrhage” if a section on a given CT demonstrates evidence of hemorrhage.

This methodology holds a distinct advantage given the clinical significance of the finding of cervical cord compression, as a false-positive is far less likely to cause patient harm than a false-negative. Therefore, any images flagged for possible compression should trigger immediate review by the radiology physician and/or the clinician who ordered the examination. Critically, however, the ViT model was also highly specific with a high PPV. Although the ResNet50 model achieved a sensitivity of 100% (ViT achieved 90%), given its poor performance in regard to specificity, PPV, and lower AUC compared with ViT, its clinical usefulness would likely be limited, given that greater than one-half of the examinations it predicted as positive for cord compression were false-positives.

Given the ever-increasing workload of radiology physicians, augmentation via AI is needed more than ever to improve workflow and speed, decrease cognitive load, and reduce misdiagnosis. If implemented in a clinical setting, a triaging tool for detecting cervical cord compression could be a meaningful aid to the practicing radiologist.

Limitations of our study include a relatively small sample size, single-center data set, and categoric/nonquantitative classification of disease. Given that our study has a relatively low  $n$  and is sourced from a single tertiary care academic medical center, our results should be considered preliminary until replicated across multiple additional sites. Additionally, there is increasing recognition of a significant proportion among healthy populations, particularly the elderly, of asymptomatic cervical cord compression.<sup>27</sup> Therefore, a tool similar to the one described in this manuscript, which flags examinations with potential cord compression, would likely be more useful in the setting of concern for acute pathology, such as trauma; less likely for outpatient examinations evaluating for chronic processes, such as degenerative disc disease. Notably, however, 58.2% of patients with severe spinal stenosis in our cohort had clinical signs and symptoms retrospectively referable to myelopathy.

Future directions will include expanding our data set to include a larger multi-institutional cohort, quantitative analysis (eg, spinal canal measurements), auto-segmentation, and possibly an evaluation of clinical implementation.

## CONCLUSIONS

This study demonstrates accurate detection of cervical spinal cord compression at the patient level by combining vision transformer section-level predictions with rules-based classification. If implemented into the clinical setting, this tool could serve as a meaningful aide to radiologists by flagging potentially abnormal examinations for expedited review.








Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

1. Fehlings MG, Vaccaro A, Wilson JR, et al. **Early versus delayed decompression for traumatic cervical spinal cord injury: results of the Surgical Timing in Acute Spinal Cord Injury Study (STASCIS).** *PLoS One* 2012;7:e32037 CrossRef Medline
2. Laur O, Nandu H, Titelbaum D, et al. **Nontraumatic spinal cord compression: MRI primer for emergency department radiologists.** *Radiographics* 2019;39:1862–80 CrossRef Medline
3. Parizel PM, Van Der Zijden T, Gaudino S, et al. **Trauma of the spine and spinal cord: imaging strategies.** *Eur Spine J* 2010;19 Suppl 1:S8–17 CrossRef Medline
4. 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
5. Tadavarthy Y, Vey B, Krupinski E, et al. **The state of radiology AI: considerations for purchase decisions and current market offerings.** *Radiol Artif Intell* 2020;2:e200004 CrossRef Medline
6. Kelly BS, Judge C, Bollard SM, et al. **Radiology artificial intelligence: a systematic review and evaluation of methods (RAISE).** *Eur Radiol* 2022;32:7998–8007 CrossRef Medline
7. Shin HJ, Han K, Ryu L, et al. **The impact of artificial intelligence on the reading times of radiologists for chest radiographs.** *NPJ Digit Med* 2023;6:82–88 CrossRef Medline
8. Merali Z, Wang JZ, Badhiwala JH, et al. **A deep learning model for detection of cervical spinal cord compression in MRI scans.** *Sci Rep* 2021;11:10473 CrossRef Medline
9. Tamai K, Terai H, Hoshino M, et al. **Deep learning algorithm for identifying cervical cord compression due to degenerative canal stenosis on radiography.** *Spine (Phila Pa 1976)* 2023;48:519–25 CrossRef Medline
10. Hopkins BS, Weber KA, Kesavabhotla K, et al. **Machine learning for the prediction of cervical spondylotic myelopathy: a post hoc pilot study of 28 participants.** *World Neurosurg* 2019;127:e436–42 CrossRef Medline
11. Ren G, Yu K, Xie Z, et al. **Current applications of machine learning in spine: from clinical view.** *Global Spine J* 2022;12:1827–40 CrossRef Medline
12. Ma S, Huang Y, Che X, et al. **Faster RCNN-based detection of cervical spinal cord injury and disc degeneration.** *J Appl Clin Med Phys* 2020;21:235–43 CrossRef Medline
13. Stephens ME, O’Neal CM, Westrup AM, et al. **Utility of machine learning algorithms in degenerative cervical and lumbar spine disease: a systematic review.** *Neurosurg Rev* 2022;45:965–78 CrossRef Medline
14. Yi W, Zhao J, Tang W, et al. **Deep learning-based high-accuracy detection for lumbar and cervical degenerative disease on T2-weighted MR images.** *Eur Spine J* 2023;32:3807–14 CrossRef Medline
15. Zhou G, Mosadegh B. **Distilling knowledge from an ensemble of vision transformers for improved classification of breast ultrasound.** *Acad Radiol* 2024;31:104–20 CrossRef
16. Liu Z, Lv Q, Yang Z, et al. **Recent progress in transformer-based medical image analysis.** *Comput Biol Med* 2023;164:107268 CrossRef Medline
17. Dosovitskiy A, Beyer L, Kolesnikov A, et al. **An image is worth 16x16 words: transformers for image recognition at scale.** *arXiv* 2021; <https://doi.org/10.48550/arXiv.2010.11929>. Accessed August 13, 2023
18. Chen CF, Fan Q, Panda R. **CrossViT: cross-attention multi-scale vision transformer for image classification.** *arXiv* 2021; <https://doi.org/10.48550/arXiv.2103.14899>. Accessed August 13, 2023
19. Gu J, Kwon H, Wang D, et al. **Multi-scale high-resolution vision transformer for semantic segmentation.** *arXiv* 2021; <https://doi.org/10.48550/arXiv.2111.01236>. Accessed August 13, 2023
20. Zhang Z, Lu X, Cao G, et al. **ViT-YOLO: transformer-based YOLO for object detection.** In: *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*. Montreal: IEEE; 2021:2799–2808 CrossRef
21. Vaswani A, Shazeer N, Parmar N, et al. **Attention is all you need.** *arXiv* 2023; <https://doi.org/10.48550/arXiv.1706.03762>. Accessed August 13, 2023

22. Kang Y, Lee JW, Koh YH, et al. **New MRI grading system for the cervical canal stenosis.** *AJR Am J Roentgenol* 2011;197:W134–40 CrossRef Medline
23. Moguilner S, Whelan R, Adams H, et al. **Visual deep learning of unprocessed neuroimaging characterises dementia subtypes and generalises across non-stereotypic samples.** *eBioMedicine* 2023;90:104540 CrossRef Medline
24. Wang X, Ji X. **Sample size estimation in clinical research: from randomized controlled trials to observational studies.** *Chest* 2020;158:S12–20 CrossRef Medline
25. Small JE, Osler P, Paul AB, et al. **CT cervical spine fracture detection using a convolutional neural network.** *AJNR Am J Neuroradiol* 2021;42:1341–47 CrossRef Medline
26. Voter AF, Larson ME, Garrett JW, et al. **Diagnostic accuracy and failure mode analysis of a deep learning algorithm for the detection of cervical spine fractures.** *AJNR Am J Neuroradiol* 2021;42:1550–56 CrossRef Medline
27. Smith SS, Stewart ME, Davies BM, et al. **The prevalence of asymptomatic and symptomatic spinal cord compression on magnetic resonance imaging: a systematic review and meta-analysis.** *Global Spine J* 2021;11:597–607 CrossRef Medline

# Identifying Patients with CSF-Venous Fistula Using Brain MRI: A Deep Learning Approach

 Shahriar Faghani,  Mana Moassefi,  Ajay A. Madhavan,  Ian T. Mark,  Jared T. Verdoorn,  Bradley J. Erickson, and  John C. Benson



## ABSTRACT

**BACKGROUND AND PURPOSE:** Spontaneous intracranial hypotension is an increasingly recognized condition. Spontaneous intracranial hypotension is caused by a CSF leak, which is commonly related to a CSF-venous fistula. In patients with spontaneous intracranial hypotension, multiple intracranial abnormalities can be observed on brain MR imaging, including dural enhancement, “brain sag,” and pituitary engorgement. This study seeks to create a deep learning model for the accurate diagnosis of CSF-venous fistulas via brain MR imaging.

**MATERIALS AND METHODS:** A review of patients with clinically suspected spontaneous intracranial hypotension who underwent digital subtraction myelogram imaging preceded by brain MR imaging was performed. The patients were categorized as having a definite CSF-venous fistula, no fistula, or indeterminate findings on a digital subtraction myelogram. The data set was split into 5 folds at the patient level and stratified by label. A 5-fold cross-validation was then used to evaluate the reliability of the model. The predictive value of the model to identify patients with a CSF leak was assessed by using the area under the receiver operating characteristic curve for each validation fold.

**RESULTS:** There were 129 patients included in this study. The median age was 54 years, and 66 (51.2%) had a CSF-venous fistula. In discriminating between positive and negative cases for CSF-venous fistulas, the classifier demonstrated an average area under the receiver operating characteristic curve of 0.8668 with a standard deviation of 0.0254 across the folds.

**CONCLUSIONS:** This study developed a deep learning model that can predict the presence of a spinal CSF-venous fistula based on brain MR imaging in patients with suspected spontaneous intracranial hypotension. However, further model refinement and external validation are necessary before clinical adoption. This research highlights the substantial potential of deep learning in diagnosing CSF-venous fistulas by using brain MR imaging.

**ABBREVIATIONS:** AUROC = area under the receiver operating characteristic curve; CET1 = contrast-enhanced T1-weighted; CTM = CT myelography; DL = deep learning; DSM = digital subtraction myelogram; SIH = spontaneous intracranial hypotension; SLEC = spinal longitudinal extradural fluid collection

Spontaneous intracranial hypotension (SIH) is a clinical condition that is usually caused by a spinal CSF leak.<sup>1</sup> SIH is associated with substantial morbidity. Patients classically present with orthostatic headaches but may also suffer from symptoms such as tinnitus, dizziness, and, in severe cases, long-term disability or even coma.<sup>2</sup> There are many subtypes of CSF leaks, though the role of CSF-venous fistulas has recently gained greater recognition.<sup>3</sup> Much of this recognition has come from considerable progress in the diagnosis

and treatment of CSF-venous fistulas, particularly with the rise of decubitus digital subtraction myelograms (DSMs) and the increased usage of decubitus CT myelography (CTM), which have led to the increased diagnosis of previously under-recognized CSF-venous fistulas.<sup>4</sup>

Multiple intracranial abnormalities have been reported on brain MR imaging in patients with SIH. These findings can largely be explained by the Monro-Kellie doctrine, in which the relative lack of CSF leads to a compensatory response in the brain parenchyma and vascular structures.<sup>5</sup> Namely, the pituitary, dura, and dural venous sinuses engorge. In some cases, the brain also “sags,” displaying both decreased mamillopontine distance and effacement of the suprasellar and prepontine cisterns. Together, these findings have been used to make a probabilistic scoring system (the “Bern score”) in which the probability of identifying a CSF leak can be calculated based on the presence or absence of these findings.<sup>6</sup>

Received September 27, 2023; accepted after revision December 12.

From the Radiology Informatics Lab (S.F., M.M., B.J.E.), and Department of Radiology, (A.A.M., I.T.M., J.T.V., J.C.B.), Mayo Clinic, Rochester, Minnesota.

Please address correspondence to John C. Benson, MD, Department of Radiology, Mayo Clinic, 200 1st St SW, Rochester, MN 55905; e-mail: benson.john3@mayo.edu; @ShahriarFaghani; @ManaMSF94; @iantmark; @Slowvak; @MayoAllLab; @MayoRadiology



Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8173>



Despite these advancements, there is still a need for a more precise method by which to identify patients with CSF-venous fistulas by using brain imaging, as up to 19% of patients with SIH can have normal brain MR imaging.<sup>7</sup> As of now, no deep learning (DL) algorithm has been specifically designed to analyze brain MRIs for this medical condition. This study aims to develop and internally validate such an algorithm (“SIHnet”) among a group of patients suspected to have SIH.

## MATERIALS AND METHODS

Institutional review board approval was obtained for this retrospective study. The requirement for informed consent was waived.

### Patient Cohort

A retrospective analysis was carried out, focusing on patients suspected of having SIH who underwent lateral decubitus DSM between December 30, 2021, and November 30, 2022. The exclusion criteria included the absence of pre-DSM brain MR imaging or substandard image quality. In addition, to focus on the SIH subpopulation associated with CSF-venous fistula, patients were excluded if a spinal longitudinal extradural fluid collection (SLEC) was noted on the pre-DSM spinal MR imaging. For any patients who had undergone a previous DSM (ie, before the study inclusion date), either 1) the initial DSM identified a CSF leak or 2) the first DSM performed at our institution (in patients for whom a CSF leak was not identified) and the most recent brain MR imaging preceding the DSM were used for analysis. The number of days between the pre-DSM brain MR imaging and analyzed DSM was recorded.

### Digital Subtraction Myelography and CT Myelography Technique

Our previously described DSM technique<sup>8</sup> spanned 2 days, with patients being positioned first in the right lateral decubitus position and subsequently in the left. DSMs were performed using biplane fluoroscopy machines, specifically, either the Artis Icono system (Siemens Healthineers) or the Allura Xper system (Philips Healthcare), with most cases using only a single anteroposterior plane. Patients were placed on a wedge-shaped cushion on a tiltable table such that their hips were higher than their shoulders. A 20- or 22-gauge spinal needle was inserted into the thecal sac, usually at the L2–3 or L3–4 level. Each side underwent 2 distinct DSM acquisitions, with one focusing on the upper spine and the other on the lower. A total of 11 mL of intrathecal Omnipaque 300 (GE Healthcare) was used. Post-DSM, the patients were transferred to a CT scanner for a full-spine same-side lateral decubitus CTM, which was conducted on a dual source Somatom Flash scanner (Siemens Healthineers). Both the DSM and the CT myelogram were analyzed as parts of a unified diagnostic assessment.

### MR Brain Imaging Protocol

Brain MRIs were performed by using either 1.5T or 3T scanners. Most image analyses relied on fat-saturated postcontrast 3D T1 fast spin echo sequences (TR = 600 ms, TE = 7.2 ms, flip angle = 120°, section thickness = 1 mm, FOV = 250 × 250 mm<sup>2</sup>).

### Image Analysis

Four neuroradiologists (I.T. M., A.A.M., J.C.B., and J.T.V.) with 2–8 years of posttraining subspecialty expertise in interpreting neuroradiologic examinations reviewed the DSM and pre-DSM MR brain images, with the entire cohort being split evenly among the reviewers. The reviewers were blinded to clinical information but not to the official reports or annotations in the picture archiving and communication system. Each reviewer independently evaluated all DSMs for the presence or absence of an identified CSF leak (specified as positive, negative, or indeterminate). MR imaging brain scans were assessed for multiple potential stigmata of SIH that must be measured to determine a Bern score, including the engorgement of venous sinuses, pachymeningeal enhancement, suprasellar cistern height, subdural fluid collection, prepontine cistern width, and mamillopontine distance. The Bern score was determined for each patient, and its discrimination performance was evaluated by calculating the area under the receiver operating characteristic curve (AUROC).

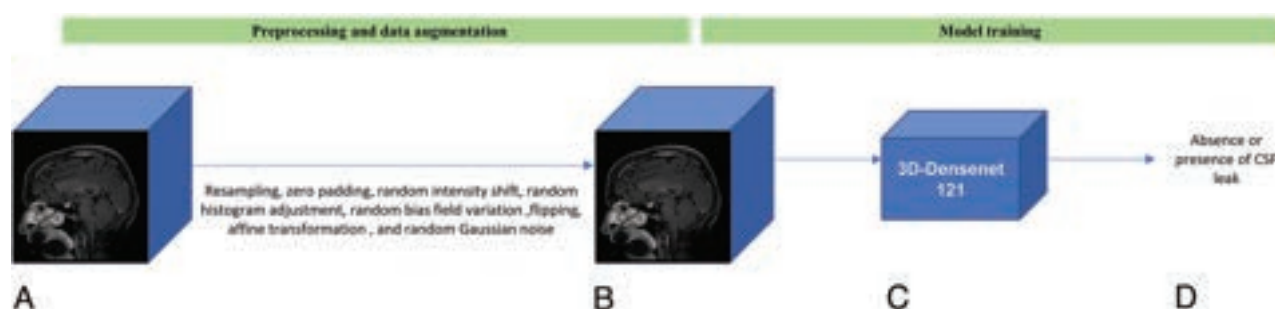
### Data Set Splitting

Participants whose DSM studies demonstrated an absence or definitive presence of CSF-venous fistula were included in the study. The data were split into 5 folds at the patient level by using the GroupKFold module from the scikit-learn package, version 1.2.0. Cross-validation of the final model was performed on all 5 groups to determine the robustness of the results.<sup>9</sup>

### Data Preprocessing and Model Development

Initially, for CSF-venous fistula detection, the process involved converting 2D sagittal contrast-enhanced T1-weighted (CET1) images from the DICOM format to the 3D Neuroimaging Informatics Technology Initiative format. Subsequently, the MR images were resampled to possess voxel dimensions of 1 × 1 × 1 mm through trilinear interpolation. To satisfy image size uniformity, the images were zero-padded to dimensions of 240, 260, and 260 voxels, corresponding to the largest dimensions observed across the data set.<sup>10</sup> To prevent overfitting, a range of data augmentation techniques were applied.<sup>11</sup> These encompassed random intensity shifts, random histogram adjustments, random bias field variations, flipping, random affine transformations, and random Gaussian noise. For further descriptions of these techniques, refer to the Online Supplemental Data.

We used a 3D-DenseNet-121 classifier sourced from the MONAI package (version 1.1.0), which showed promising results in classification tasks dealing with 3D medical imaging volumes.<sup>10,12,13</sup> DenseNet, a convolutional neural network, uses convolutions to extract meaningful information while connecting each layer to every other layer in a feed-forward manner. Within the architecture of DenseNet, a 1 × 1 convolution was introduced before each convolution layer to serve as a bottleneck layer to trim the count of feature mappings.<sup>14</sup> This approach to dimension reduction, implemented through both bottleneck layers and transition layers within each attenuated block, contributes to enhanced parameter efficiency and a reduction in model complexity. This design choice aids in mitigating the risk of overfitting. To enhance the model's generalizability and mitigate overfitting, we used the AdamW optimizer with a batch size of 4



**FIG 1.** Schematic illustration of the preprocessing, data augmentation, and model training pipeline. A, 3D contrast-enhanced T1 input volume. B, Preprocessed and augmented CET1 to mitigate overfitting. C, 3D-DenseNet model. D, Prediction of absence or presence of CSF leak.

**Table 1: Participant characteristics and Bern score status**

Subject Characteristics	All Subjects (n = 129)
Median age in years (interquartile range)	54 (20)
Age range in years	23–87
Female	84 (65.12%)
Male	45 (34.48%)
Bern score	
Low risk (Bern score 0–2)	47 (36.43%)
Intermediate risk (Bern score 3–4)	26 (20.15%)
High risk (Bern score >4)	56 (43.42%)

**Table 2: Summary of classifier performance per validation fold**

Fold Number	AUROC
1	0.8988
2	0.8284
3	0.8580
4	0.8580
5	0.8910

and a weight decay of 0.1.<sup>15</sup> This choice of optimizer integrates weight decay as a regularization technique. The scheduler started with an initial learning rate of 0.001 and ran for 450 epochs, and this was followed by a restart interval of 200 epochs. This approach was adopted to circumvent potential issues with converging to local minima.<sup>16</sup> Figure 1 illustrates the schematic pipeline for preprocessing and training.

To handle the slight class imbalance in the data set, we utilized a weighted cross-entropy loss function and assigned reverse class ratios as weights to the positive and negative classes to more effectively address the imbalance.<sup>17</sup> The model weights from the epoch with the greatest AUROC, indicating the highest discriminative value, were saved for every fold.<sup>18</sup> In addition to each fold, the mean and standard deviation of the AUROC across the folds were reported.

### Occlusion Interpretation Maps

In our study, we used an occlusion interpretation technique to investigate the significance of different regions within 3D CET1 volumes for model predictions. This approach systematically replaces portions of the input volume with zeros and measures the impact on the model's output logits to identify the most influential regions in the CET1 volume for the model's decision-making process. A 3D patch with a shape of  $12 \times 13 \times 13$  was slid across the input CET1 volume, and the change in the model's prediction score for the target label was recorded at each step.

These changes were accumulated to form an occlusion map, which served as a visual interpretation tool. The stride for moving the patch through the input CET1 volume was set to  $12 \times 13 \times 13$ , ensuring a comprehensive yet computationally efficient exploration of the 3D volume space.

Our model training was conducted on a cluster comprising 4 NVIDIA A100 GPUs. All image processing and model development were performed using PyTorch 1.12.0 and MONAI 1.1.0 with Python 3.10.4 on a GPU cluster of 4 GPUs (NVIDIA A100). All statistical analyses were performed using scikit-learn 1.2.0.

## RESULTS

There were 129 patients were included in this study. The median age was 54 years (interquartile range = 20 years). There were 66 (51.2%) patients with and 63 (48.8%) patients without CSF-venous fistula. Patient demographics and brain MR imaging Bern scores are summarized in Table 1.

The AUROC for the Bern score, which is used to differentiate between patients with CSF-venous fistula and those without a leak, was 0.8. The classifier demonstrated an average AUROC of 0.8668, with a standard deviation of 0.0254 across the folds, in determining the positive and negative cases of CSF leaks. Table 2 and Figure 2 summarize the model performance on each fold.

Figure 3 illustrates the occlusion map, highlighting the key regions upon which the model relies for its decision-making process.

## DISCUSSION

The present study set out to develop and internally validate a DL algorithm capable of identifying patients in whom DSM could localize a spinal CSF leak in the context of suspected SIH, based on sagittal CET1 brain MR imaging. The results indicate that the algorithm created in this study had a classifier with an average AUROC of 0.8668, surpassing the discriminative ability of the Bern score and indicating a promising level of performance.

The algorithm developed in this study has the potential to substantially impact the ability of clinicians to identify patients who would benefit from DSM imaging in the context of suspected SIH. Specifically, this algorithm could serve as a preliminary screening tool to aid both neurologists and neuroradiologists in the assessment of patients with clinically suspected SIH. By identifying which patients are considered likely to have a CSF-venous fistula identified on subsequent imaging, clinicians could use this

tool to better select who should undergo additional invasive testing, such as DSM or dynamic CTM imaging. Thus, the algorithm has the potential to both streamline the diagnostic process and reduce health care costs by minimizing the occurrence of unnecessary tests.

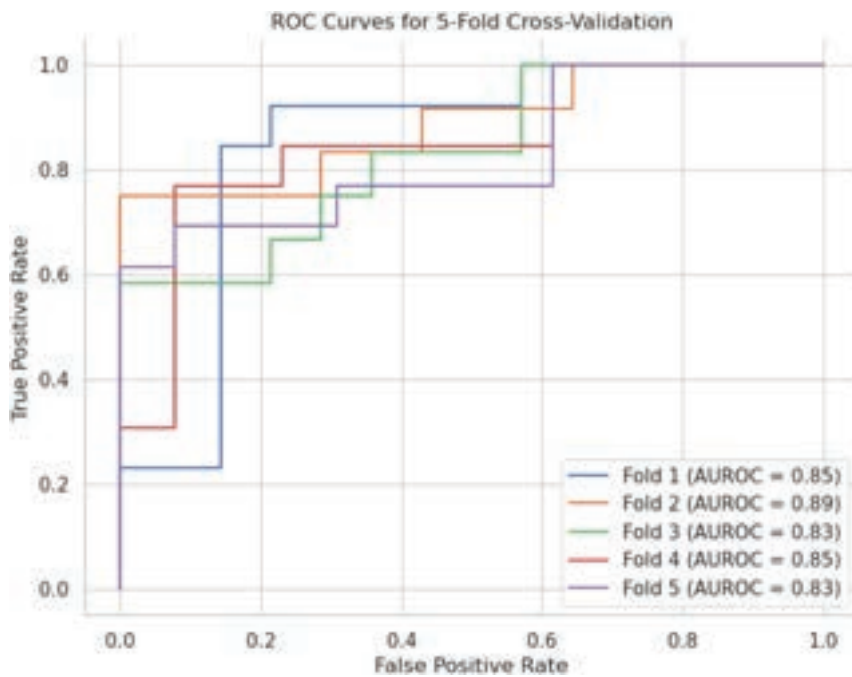
In addition, the results presented here are the first of their kind. While machine learning has made considerable strides in neuroradiology, there is a notable gap in its application for the diagnosis and prognosis of CSF leaks.<sup>19</sup> A study by Fu et al<sup>20</sup> used DL to quantify spinal CSF in whole-spine MR myelography. However, their focus was on a different aspect of CSF analysis, leaving a critical void in the utilization of machine learning for CSF leak diagnosis. Arnold et al<sup>21</sup> extracted radiomics features from brain MR imaging and then trained a support vector

machine to discriminate between patients with SIH and healthy controls. These results underscore the significance of the current study, which aims to address this unexplored area by applying end-to-end deep learning methods, specifically for the diagnosis of CSF leaks.

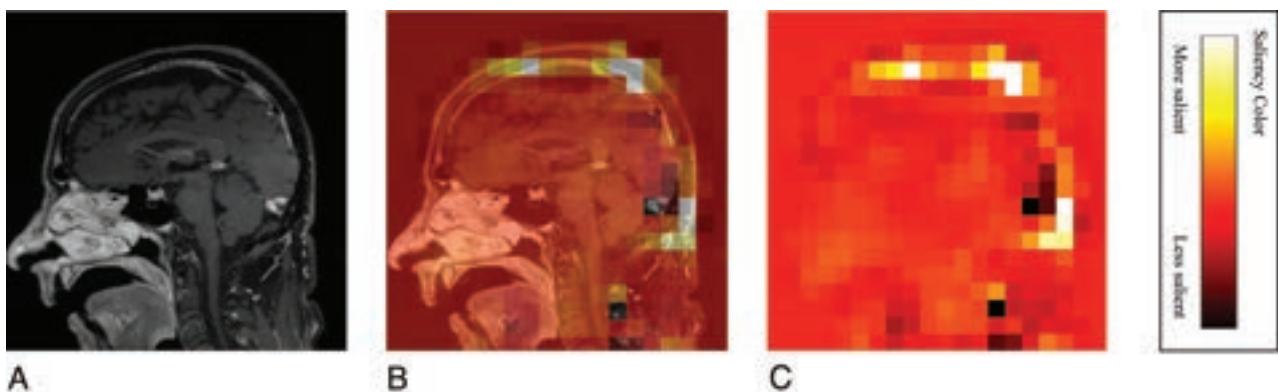
Over recent years, there has been considerable progress in our knowledge and diagnostic understanding of SIH and CSF leak imaging. Throughout these gains, the Bern score has remained a steadfast metric that is widely used to predict which patients with suspected SIH may have a spinal CSF leak.<sup>6</sup> Nevertheless, the Bern score does have some limitations.<sup>22</sup> It is heavily dependent on sub-mm measurements, which are imperfect on the PACS systems of most institutions. In addition, the cohort used for analysis in the initial Bern score study was composed entirely of

patients with “fast” CSF leaks, meaning those that are accompanied by a SLEC. However, it is now known that a substantial percentage of patients with SIH lack a SLEC. CSF leaks in such patients, if found, are often CSF-venous fistulas. The current study sought to address this gap in knowledge by using a cohort of patients without a SLEC on pre-DSM spinal MR imaging.

Interestingly, the occlusion maps produced by this study’s algorithm revealed SIHnet’s tendency to allocate greater weights to the superior and posterior regions of the brain. This allocation does not fully align with the established Bern criteria, which place greater emphasis on the brainstem, suprasellar cistern, and adjacent anatomic structures. Nevertheless, it is crucial to interpret these results judiciously. Occlusion maps and other interpretation methodologies are not definitive tools for understanding the underlying decision-making process of a model. While they may not elucidate “which” features are considered



**FIG 2.** Receiver operating characteristic (ROC) curves for 5-fold cross-validation. Each curve represents the performance of the model on a distinct validation fold. The curves demonstrate the model’s ability to distinguish between the absence and presence of CSF leaks from brain MR imaging scans.



**FIG 3.** Three-part representation of the regions that are crucial to the model’s decision-making process in detecting CSF leaks. A, Sagittal view of contrast-enhanced T1 brain MR imaging. B, Occlusion mask overlaid on the original contrast-enhanced T1 image, highlighting the regions that significantly influence the model’s predictions. C, Occlusion mask generated to identify regions of interest.

essential by a model, they do potentially indicate “where” the model is focusing its attention. Despite this limitation, the model may discern brain sagging by focusing on extracted information from the superior and posterior parts of the brain. It could potentially evaluate features such as pachymeningeal enhancement, venous engorgement, and subdural collections, as well as the dural geometry in these areas, and these could be evaluated in a quantitative manner rather than via a simple qualitative assessment of their presence or absence. The insights derived from interpretation maps could still be valuable for identifying imaging biomarkers in future studies.<sup>23,24</sup>

Our study is not without limitations. Most notably, the algorithm was trained and validated by using retrospective data from a single institution, albeit using multiple MR imaging scanner models and protocols. To generalize the applicability of our model, it is crucial to externally validate it by using data sets from various institutions with diverse imaging protocols. In addition, only patients without a SLEC were included in the final patient cohort. Additional validation by using patients with both “slow” and “fast” CSF leaks will be needed to allow for greater generalizability of the results generated here.

## CONCLUSIONS

This study developed and internally validated a DL algorithm to identify patients with a CSF leak based on brain MR imaging findings. The algorithm yielded a promising level of performance, but further refinement and external validation are required prior to its clinical adoption.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

- Cheema S, Anderson J, Angus-Leppan H, et al. **Multidisciplinary consensus guideline for the diagnosis and management of spontaneous intracranial hypotension.** *J Neurol Neurosurg Psychiatry* 2023; 94:835–43 CrossRef Medline
- Severson M, Schaurich CG, Strecker-McGraw MK. *Cerebrospinal Fluid Leak*. StatPearls Publishing 2023.
- Brinjikji W, Garza I, Whealy M, et al. **Clinical and imaging outcomes of cerebrospinal fluid-venous fistula embolization.** *J Neurointerv Surg* 2022;14:953–56 CrossRef Medline
- Callen AL, Timpone VM, Schwertner A, et al. **Algorithmic multimodality approach to diagnosis and treatment of spinal CSF leak and venous fistula in patients with spontaneous intracranial hypotension.** *AJR Am J Roentgenol* 2022;219:292–301 CrossRef Medline
- Benson JC, Madhavan AA, Cutsforth-Gregory JK, et al. **The Monroe-Kellie doctrine: A review and call for revision.** *AJNR Am J Neuroradiol* 2023;44:2–6 CrossRef
- Dobrocky T, Grunder L, Breiding PS, et al. **Assessing spinal cerebrospinal fluid leaks in spontaneous intracranial hypotension with a scoring system based on brain magnetic resonance imaging findings.** *JAMA Neurol* 2019;76:580–87 CrossRef Medline
- D'Antona L, Jaime Merchan MA, Vassiliou A, et al. **Clinical presentation, investigation findings, and treatment outcomes of spontaneous intracranial hypotension syndrome: A systematic review and meta-analysis.** *JAMA Neurol* 2021;78:329–37 CrossRef Medline
- Kim DK, Brinjikji W, Morris PP, et al. **Lateral decubitus digital subtraction myelography: Tips, tricks, and pitfalls.** *AJNR Am J Neuroradiol* 2020;41:21–28 CrossRef Medline
- Rouzrokh P, Khosravi B, Faghani S, et al. **Mitigating bias in radiology machine learning: 1. Data Handling.** *Radiology Artif Intell* 2022;4:e210290
- Singh Y, Kelm ZS, Faghani S, et al. **Deep learning approach for differentiating indeterminate adrenal masses using CT imaging.** *Abdom Radiology (NY)* 2023;48:3189–94 CrossRef Medline
- Zhang K, Khosravi B, Vahdati S, et al. **Mitigating bias in radiology machine learning: 2. Model Development.** *Radiology Artif Intell* 2022;4:e220010
- Moassefi M, Faghani S, Conte GM, et al. **A deep learning model for discriminating true progression from pseudoprogression in glioblastoma patients.** *J Neurooncol* 2022;159:447–55 CrossRef Medline
- Faghani S, Khosravi B, Moassefi M, et al. **A comparison of three different deep learning-based models to predict the MGMT promoter methylation status in glioblastoma using brain MRI.** *J Digit Imaging* 2023;36:837–46 CrossRef Medline
- Huang G, Liu Z, Van Der Maaten L, et al. **Densely connected convolutional networks.** *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017. p. 2261–69 CrossRef
- Loshchilov I, Hutter F. **Decoupled weight decay regularization.** arXiv [cs.LG]. 2017. <http://arxiv.org/abs/1711.05101>
- Gotmare A, Keskar NS, Xiong C, et al. **A closer look at deep learning heuristics: learning rate restarts, warmup and distillation.** arXiv [cs.LG]. 2018. <http://arxiv.org/abs/1810.13243>
- Ho Y, Wookey S. **The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling.** arXiv [cs.LG]. 2020. <http://arxiv.org/abs/2001.00570>
- Faghani S, Khosravi B, Zhang K, et al. **Mitigating bias in radiology machine learning: 3. Performance Metrics.** *Radiology Artif Intell* 2022;4:e220061 CrossRef Medline
- Moassefi M, Faghani S, Khosravi B, et al. **Artificial intelligence in radiology: Overview of application types, design, and challenges.** *Semin Roentgenol* 2023;58:170–77 CrossRef Medline
- Fu J, Chai J-W, Chen P-L, et al. **Quantitative measurement of spinal cerebrospinal fluid by cascade artificial intelligence models in patients with spontaneous intracranial hypotension.** *Biomedicine* 2022;10 CrossRef Medline
- Arnold PG, Kaya E, Reisert M, et al. **Support vector machine-based spontaneous intracranial hypotension detection on brain MRI.** *Clin Neuroradiol* 2022;32:225–30 CrossRef Medline
- Houk JL, Morrison S, Peskoe S, et al. **Validity of the Bern score as a surrogate marker of clinical severity in patients with spontaneous intracranial hypotension.** *AJNR Am J Neuroradiol* 2023;44:1096–100 CrossRef Medline
- Steyaert S, Pizurica M, Nagaraj D, et al. **Multimodal data fusion for cancer biomarker discovery with deep learning.** *Nat Mach Intell* 2023;5:351–62 CrossRef Medline
- Couture HD. **Deep learning-based prediction of molecular tumor biomarkers from H&E: A practical review.** *J Pers Med* 2022;12 CrossRef Medline



# Compressed Sensitivity Encoding Artificial Intelligence Accelerates Brain Metastasis Imaging by Optimizing Image Quality and Reducing Scan Time

Mengmeng Wang, Yue Ma, Linna Li, Xingchen Pan, Yafei Wen, Ying Qiu, Dandan Guo, Yi Zhu, Jianxiu Lian, and Dan Tong



## ABSTRACT

**BACKGROUND AND PURPOSE:** Accelerating the image acquisition speed of MR imaging without compromising the image quality is challenging. This study aimed to evaluate the feasibility of contrast-enhanced (CE) 3D T1WI and CE 3D-FLAIR sequences reconstructed with compressed sensitivity encoding artificial intelligence (CS-AI) for detecting brain metastases (BM) and explore the optimal acceleration factor (AF) for clinical BM imaging.

**MATERIALS AND METHODS:** Fifty-one patients with cancer with suspected BM were included. Fifty participants underwent different customized CE 3D-T1WI or CE 3D-FLAIR sequence scans. Compressed SENSE encoding acceleration 6 (CS6), a commercially available standard sequence, was used as the reference standard. Quantitative and qualitative methods were used to evaluate image quality. The SNR and contrast-to-noise ratio (CNR) were calculated, and qualitative evaluations were independently conducted by 2 neuro-radiologists. After exploring the optimal AF, sample images were obtained from 1 patient by using both optimized sequences.

**RESULTS:** Quantitatively, the CNR of the CS-AI protocol for CE 3D-T1WI and CE 3D-FLAIR sequences was superior to that of the CS protocol under the same AF ( $P < .05$ ). Compared with reference CS6, the CS-AI groups had higher CNR values (all  $P < .05$ ), with the CS-AI10 scan having the highest value. The SNR of the CS-AI group was better than that of the reference for both CE 3D-T1WI and CE 3D-FLAIR sequences (all  $P < .05$ ). Qualitatively, the CS-AI protocol produced higher image quality scores than did the CS protocol with the same AF (all  $P < .05$ ). In contrast to the reference CS6, the CS-AI group showed good image quality scores until an AF of up to 10 (all  $P < .05$ ). The CS-AI10 scan provided the optimal images, improving the delineation of normal gray-white matter boundaries and lesion areas ( $P < .05$ ). Compared with the reference, CS-AI10 showed reductions in scan time of 39.25% and 39.93% for CE 3D-T1WI and CE 3D-FLAIR sequences, respectively.

**CONCLUSIONS:** CE 3D-T1WI and CE 3D-FLAIR sequences reconstructed with CS-AI for the detection of BM may provide a more effective alternative reconstruction approach than CS. CS-AI10 is suitable for clinical applications, providing optimal image quality and a shortened scan time.

**ABBREVIATIONS:** AF = acceleration factor; BM = brain metastases; CE = contrast-enhanced; CNR = contrast-to-noise ratio; CS = compressed SENSE; AI = artificial intelligence; SENSE = sensitivity encoding

In recent years, with the prolonged survival of patients with cancer, the incidence of brain metastases (BM) has increased substantially, and BM have become the most common intracranial tumors in adults.<sup>1,2</sup> BM have an important impact on the TNM-

staging and clinical therapy of tumors. In addition, the number and size of BM are closely related to the choice of treatment method (eg, stereotactic radiosurgery versus surgical excision).<sup>3</sup> Therefore, the early detection and accurate diagnosis of BM are vital.<sup>4</sup>

With its excellent soft tissue contrast and high-resolution imaging capabilities, contrast-enhanced (CE) MR imaging has emerged as the primary imaging technique for screening and diagnosing BM. According to studies by Kaufmann et al,<sup>1</sup> CE-T1WI sequences can provide high-contrast images and information on the location, morphology, and blood supply of BM. CE-FLAIR sequences have better visibility of leptomeningeal metastases and superficial cortical metastases. Moreover, 3D MR imaging scanning is more sensitive than 2D scanning, and it offers the advantage of being able to detect small BM.<sup>5,6</sup> However, the time-consuming image

Received October 11, 2023; accepted after revision December 25.

From the Department of Radiology (M.W., Y.M., L.L., X.P., Y.W., Y.Q., D.G., D.T.), The First Hospital of Jilin University, Changchun, China; and Philips Healthcare (Y.Z., J.L.), Beijing, China.

Mengmeng Wang and Yue Ma contributed equally to the work.

Grant Support: This work was supported by grants from the Jilin Provincial Department of Finance (Correction: Nos. JLSWSRCZX2021-2).

Please address correspondence to Dan Tong, Department of Radiology, The First Hospital of Jilin University, Changchun 130000, China; e-mail: tongdan@jlu.edu.cn

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

<http://dx.doi.org/10.3174/ajnr.A8161>

acquisition of 3D sequences can limit clinical applications under certain circumstances, especially in patients who cannot remain stationary for a long time due to pain or a disturbance of consciousness.<sup>7</sup>

Accelerating the image acquisition speed of MR imaging without reducing the image quality has always been challenging. Several image acquisition and postprocessing technologies have been developed over the past few decades, with 2 of the more influential developments being parallel imaging and compressed sensing.<sup>8</sup> In particular, compressed sensitivity encoding (CS), based on the principle of compressed sensing and the commonly used sensitivity encoding (SENSE) technology, has been used to optimize MR imaging scanning sequences of the brain, liver, and nerves.<sup>9–11</sup> CS implements random undersampling in *k*-space and capitalizes on the sparsity of signals during reconstruction. By using wavelet transforms, CS distinguishes between signal and noise. This strategy allows CS to accelerate the imaging process by operating at a sampling rate obviously lower than that prescribed by the Nyquist sampling theorem, and this enables the reconstruction of complete signals from limited data, reducing the acquisition time without compromising the SNR, especially at higher acceleration factors (AFs).<sup>12</sup> Compared with traditional acquisition, CS was shown to reduce the acquisition time of 3D T1-echo-spoiled gradient-echo and 3D T2-FLAIR sequences for brain tumors by 35% (2:56 minutes versus 4:43 minutes) and 25% (3:36 minutes versus 4:33 minutes), respectively, without sacrificing image quality.<sup>7</sup> Integrating artificial intelligence (AI) into MR imaging reconstruction has recently attracted considerable attention to the further acceleration of the scan speed.<sup>13,14</sup> CS-AI is a novel reconstruction algorithm that applies an adaptive CS network approach inspired by CS theory to reconstruct images, builds on the iterative shrinkage-thresholding algorithm (ISTA)-Net framework developed by Zhang et al,<sup>15</sup> and introduces strong prior information. It has demonstrated superior performance in reconstructing images from highly undersampled *k*-space data, improving the speed and quality of the reconstruction.<sup>16</sup>

To further improve the image scanning sequences of BM, we applied CS-AI technology to prospectively scan patients to achieve a balance between shortening the scanning time and improving the image quality. Therefore, the purposes of this study were to acquire highly accelerated CE 3D-T1WI and CE 3D-FLAIR images by using CS-AI framework reconstruction and to evaluate the image quality of CS-AI with different and high AFs, compared with those of the CS technique, to explore the optimal AF for the clinical imaging of BM.

## MATERIALS AND METHODS

This prospective study was approved by our institution's review board (ID: 22K036-001), and written informed consent was obtained from all participants.

### Patient Selection

From January to November 2022, we recruited 60 patients with cancer with suspected BM for brain MR imaging examinations. Five patients who exhibited BM disappearance and 4 patients with severe imaging artifacts were excluded from the study. The criteria for determining BM are as follows: 1) new enhanced

lesions, increased size of existing lesions on follow-up CE MR imaging, or decreased size or disappeared lesions after treatment; 2) lesions that were not artifacts or normally enhanced structures; and 3) enhanced lesions occurring in the brain parenchyma.<sup>5</sup> Finally, 51 patients (25 patients each in the CE 3D-T1WI and CE 3D-FLAIR groups and 1 patient in the ultimate optimized sequences group) were enrolled.

### Image Acquisition

All MR images of the patients were obtained by using a 3T MR scanner (Ingenia Elition, Philips Healthcare) with 16-channel head coils. The patients underwent a conventional MR imaging examination, including 2D-T1WI, 2D-T2WI, 2D T2-FLAIR, and DWI sequences. Enhanced scans were initiated 3 minutes after a single dose of gadolinium contrast injection (0.1 mmol/kg Gadovist at a rate of 1.0 mL/s). A timing bolus was used to determine the precise onset of imaging post contrast administration. Twenty-five patients underwent customized CE 3D-T1WI (3D BrainVIEW T1-weighted TSE, Philips Healthcare) sequence scanning, whereas another 25 patients received customized CE 3D-FLAIR (3D BrainVIEW FLAIR) sequence scanning (Fig 1). For these sequences, we only altered the AF and the reconstruction methods (CS or CS-AI); all other scanning parameters remained unchanged. In addition, to account for the impact of contrast agent concentration decay, the order of AFs (ranging from 6 to 12) used during the data acquisition was randomized. The denoising level of the CS group was set to "system default," and the CS-AI group was set to "complete." The thin section sagittal images that could be 3D-reconstructed were obtained. The parameters of the MR scanning are summarized in Tables 1 and 2.

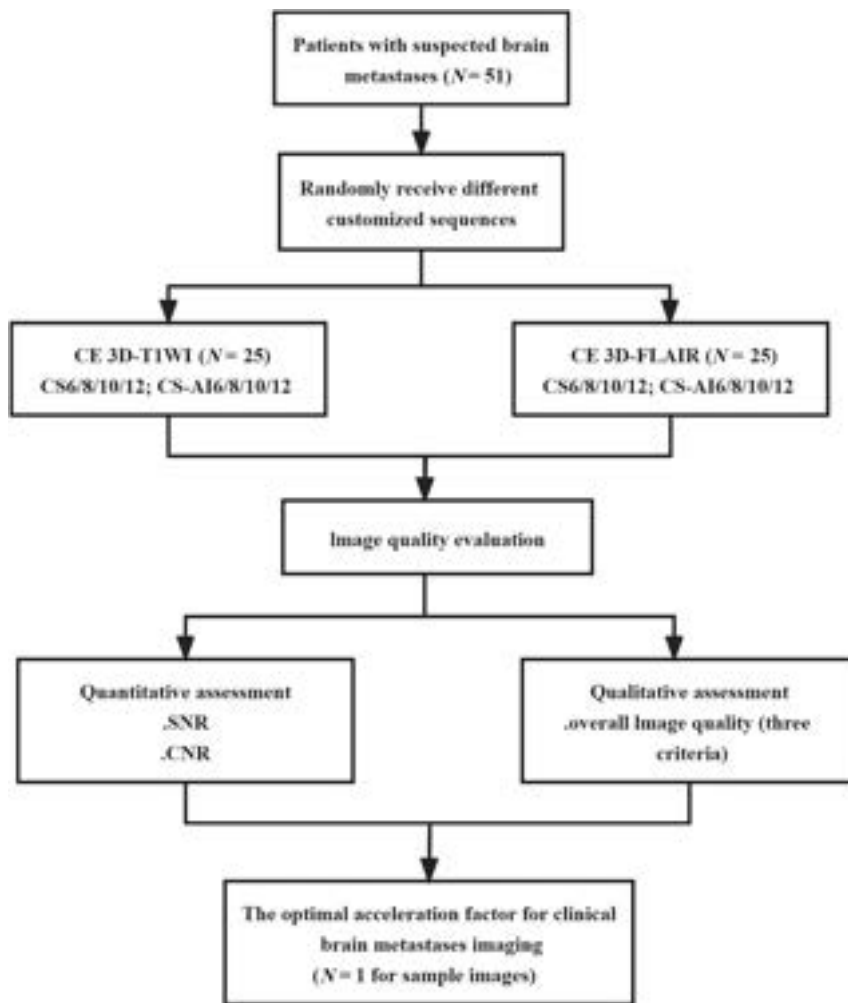
### CS-AI Algorithm

CS-AI reconstruction is an "adaptive CS network" based on deep learning and inspired by CS theory.<sup>15</sup> CS-AI replaces the iterative optimization process and the traditional wavelet transform in CS with a convolutional neural network to deal with undersampled *k*-space data. The algorithm combines a multi-scale sparse method based on learning with the compressed sensing reconstruction method to ensure data consistency.<sup>17</sup> The algorithm also integrates prior information, such as the phase constraint and image background, and uses a deep neural network to refine and modify previous assumptions. The CS-AI was trained on approximately 740,000 MR images (1.5T and 3T) of various anatomic regions and imaging contrasts.<sup>16</sup> Its performance is better than the traditional CS reconstruction algorithm and deep learning reconstruction algorithms that do not consider prior knowledge.<sup>16,18,19</sup>

### Imaging Evaluation

Quantitative and qualitative image analysis methods were applied to evaluate the image quality of the CE 3D-T1WI and CE 3D-FLAIR sequences that were reconstructed by using CS and CS-AI with different AFs (6–12). All images were transmitted and reconstructed by using a picture archiving communication system and an IntelliSpace Portal (V9; Philips Healthcare).

**Quantitative Image Analysis.** Two radiologists (Y.F. and M.W.) with more than 8 years of experience performed quantitative



**FIG 1.** Flowchart of the experimental design.

image analyses. Regions of interest were placed on the lesion and adjacent normal white matter, avoiding areas of hemorrhage, necrosis, and vessels. The size of the delineated region of interest depended on the lesion size, ranging from 4.00 to 21.00 mm<sup>2</sup>. To obtain comparable measurements, the area of interest first delineated on CS6 was copied and pasted in the corresponding regions on the images reconstructed via CS and CS-AI technology with different AFs. The standard deviation used to represent the noise and average signal intensity was recorded for each region of interest group. Based on the regions of interest, the SNR and contrast-to-noise ratio (CNR) were analyzed for objective evaluation by using Equations 1<sup>20</sup> and 2<sup>21</sup>:

$$1) \quad SNR_{lesion} = \frac{SI_{lesion}}{SD_{lesion}}$$

$$2) \quad CNR_{lesion\_white\ matter} = \frac{|SI_{lesion} - SI_{white\ matter}|}{\sqrt{SD_{lesion}^2 + SD_{white\ matter}^2}}$$

where SD stands for the standard deviation representing noise and SI represents the signal intensity.

**Qualitative Image Analysis.** All images were independently evaluated by 2 neuroradiologists (Y.M. and L.L.) with more than 15 years of experience and who were unaware of any clinical and sequence information. A 5-point Likert scale was used to grade the overall image quality: 1 = poor, nondiagnostic image quality; 2 = below average, interpretable but limited; 3 = average, interpretable; 4 = good; and 5 = excellent (including sharpness of gray-white matter boundaries, sharpness of lesion boundaries, and artifacts).<sup>22</sup> Subjective scores were independently assessed twice by 2 neuroradiologists, with the 2 readings separated by 4 weeks to avoid recall bias. The average value was used in a consistency test.

### Statistical Analysis

All statistical analyses were performed using the GraphPad 9.5.1 (GraphPad Software) and MedCalc, version 19 (<https://www.medcalc.org>), statistical software packages. The interobserver consistency of the protocol images was calculated using a weighted Cohen kappa. The kappa value rating system was as follows: a kappa value of <0.2 indicated slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, good agreement; and 0.81–1.00, excellent agreement.<sup>23</sup>

The normality of quantitative parameters and subjective scores was analyzed using the Shapiro–Wilk test. The test results indicated that the SNR, CNR, and Likert scale scores were not normally distributed. Therefore, the Wilcoxon signed-rank test was used to compare the differences between the SNR, CNR, and subjective scores of the CS-AI protocols and CS at the same AF. The Friedman test was then used to compare the quantitative values of the images reconstructed by using the CS6 and CS-AI factors. In cases of statistical significance, the Dunn pairwise post hoc test was used to perform multiple comparisons. The threshold for statistical significance was set as  $P < .05$ .

## RESULTS

### Patient Characteristics

A total of 51 patients (27 men and 24 women; mean age, 57.22 ± 8.05 years; age range, 38–73 years) were included in this study. The primary malignancies included lung cancer ( $n = 42$ ), breast cancer ( $n = 6$ ), colon cancer ( $n = 1$ ), malignant melanoma ( $n = 1$ ), and kidney cancer ( $n = 1$ ). Among them, 25 underwent accelerated CE 3D-T1WI by using the CS and CS-AI techniques, and a further 25 patients underwent CS- and CS-AI-accelerated CE

**Table 1: Imaging parameters for CS- and CS-AI-accelerated CE 3D-T1WI sequences**

Parameter	CE 3D-T1WI		
	CS6 (RS)	CS 8/10/12	CS-AI 6/8/10/12
TR (ms)	600	600	600
TE (ms)	28	28	28
FOV (mm <sup>2</sup> )	250 × 250	250 × 250	250 × 250
Voxel size (mm <sup>3</sup> )	0.99 × 1.00 × 1.10	0.99 × 1.00 × 1.10	0.99 × 1.00 × 1.10
Slices	327	327	327
Acceleration factor	6	8/10/12	6/8/10/12
Scan time (min)	3:34	2:42/2:10/1:49	3:34/2:42/2:10/1:49
Scan time reduction	-	24.29%/39.25%/49.06%	-/24.29%/39.25%/49.06%

**Note:**—dash indicates no value; RS, reference sequence; FOV, field of view; min, minute.

**Table 2: Imaging parameters for CS- and CS-AI-accelerated CE 3D-FLAIR sequences**

Parameter	CE 3D-FLAIR		
	CS6 (RS)	CS 8/10/12	CS-AI 6/8/10/12
TR (ms)	4800	4800	4800
TE (ms)	340	340	340
FOV (mm <sup>2</sup> )	250 × 250	250 × 250	250 × 250
Voxel size (mm <sup>3</sup> )	1.12 × 1.12 × 1.12	1.12 × 1.12 × 1.12	1.12 × 1.12 × 1.12
Slices	326	326	326
Acceleration factor	6	8/10/12	6/8/10/12
Scan time (min)	4:48	3:41/2:53/2:24	4:48/3:41/2:53/2:24
Scan time reduction	-	23.26%/39.93%/50.00%	-/23.26%/39.93%/50.00%

**Note:**—dash indicates no value; RS, reference sequence; FOV, field of view; min, minute.

3D-FLAIR imaging. After determining the optimal AF, 1 patient was scanned by using both optimized sequences.

### Quantitative Assessment

Figure 2 shows the quantitative analysis results of the SNR and CNR for CS- and CS-AI-accelerated CE 3D-T1WI with different AFs. Under the same AF, the SNR and CNR of the CS-AI group were superior to those of the CS group (all  $P < .05$ ), except at AF8, for which there was no statistically significant difference between the SNR values of the CS8 and CS-AI8 sequences ( $P > .05$ ). Compared with the reference sequence CS6, the CS-AI groups had higher SNR and CNR values (all  $P < .05$ ), with the CS-AI10 protocol showing the highest quantitative parameters.

Figure 3 shows the quantitative SNR and CNR analysis results for the CS and CS-AI-accelerated CE 3D-FLAIR images with different AFs. With the increase of the AF, the benefits of the CS-AI group continued up to AF10 (all  $P < .05$ ) under the same AF conditions. At AF=12, the SNR values of both CS and CS-AI decreased, and there was no statistical difference between the 2 groups ( $P > .05$ ). Compared with the reference sequence CS6, the CS-AI groups exhibited significantly higher SNR and CNR values (all  $P < .05$ ), with CS-AI10 being the most prominent.

In addition, for both the CE 3D-T1WI and CE 3D-FLAIR sequences, as the AF increased, the SNR and CNR of the images first increased and then decreased.

### Qualitative Assessment

Between the 2 observers, the image scores of CE 3D-T1WI with different AFs were in good agreement [kappa (95% confidence

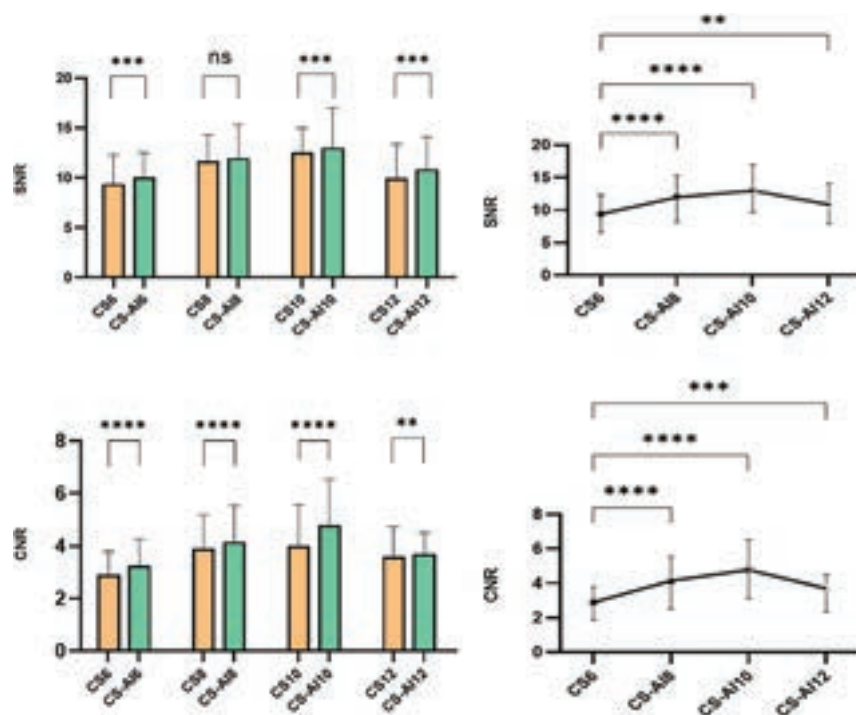
interval) = 0.74 (0.68–0.80)], and the image scores of CE 3D-FLAIR were in excellent agreement [kappa (95% confidence interval) = 0.82 (0.77–0.87)]. Figure 4 shows that for both CE 3D-T1WI and CE 3D-FLAIR sequences, the image quality of the CS-AI group was superior to that of the CS group under the same AF (all  $P < .05$ ). No statistically significant difference was observed in the overall image quality between the reference sequence CS6 and CS-AI12-accelerated CE 3D-T1WI ( $P > .05$ ). In contrast, the image quality of all other CS-AI groups was superior to that of the reference sequence CS6 (all  $P < .05$ ). Both neuroradiologists observed that CS-AI10 provided the best quality images ( $P < .05$ ), with improved sharpness at the lesion boundaries and gray-white matter connections compared with the other acceleration protocols (Fig 5C1, C2 and Fig 6C1, C2). Furthermore, in the CS group, as the AF increased, the gray-white matter junction of the CS10 and CS12 images became unsharp, boundaries of the lesions started becoming blurred, and artifact was more evident (Fig 5 and 6).

### DISCUSSION

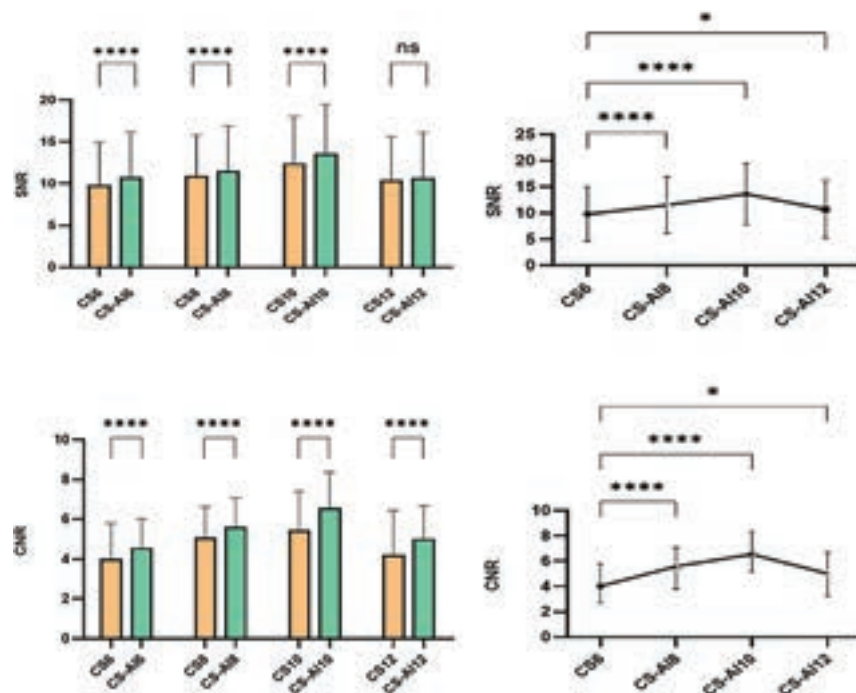
The clinical feasibility of CS-AI for the accelerated CE 3D-T1WI and CE 3D-FLAIR imaging of BM was investigated, and the optimal AF was explored by comparing CS and CS-AI technology with different AFs (6–12). Quantitative and qualitative assessments were used to analyze the acquired images to explore the value of applying CS-AI technology for patients with BM.

The quantitative analysis showed that the parameters of CE 3D-T1WI and CE-FLAIR reconstruction accelerated by CS-AI10 were superior to those of the other AF sequences. In the qualitative analysis, for the CE 3D-T1WI and CE 3D-FLAIR sequences, the image quality obtained by the CS-AI technique with different AFs was equal to or better than the overall image quality of the reference CS6, which indicated that the image quality obtained by the CS-AI protocol was at least comparable with that of the reference sequence. In addition, the 2 neuroradiologists observed that the CS-AI10-accelerated images improved the delineation of gray-white matter boundaries and lesion areas with better image quality than the other sequences. Therefore, when applying CS-AI technology to CE 3D-T1WI and CE 3D-FLAIR sequences on 3T MR systems and considering improving the image quality and shortening the scan time, AF10 is recommended. CS-AI exhibited higher CNR and better image quality at higher AFs, which may be attributed to its adaptive noise reduction capability. As the AF increases, the model's denoising effectiveness becomes more pronounced, resulting in images with enhanced contrast and improved image content, such as sharper gray-white matter connections and lesion





**FIG 2.** Quantitative analysis results of the SNR and CNR of CS- and CS-AI-accelerated CE 3D-T1WI under different AFs. Note: NS indicates not significant; \*,  $P < .05$ ; \*\*,  $P < .01$ ; \*\*\*,  $P < .001$ ; \*\*\*\*,  $P < .0001$ .

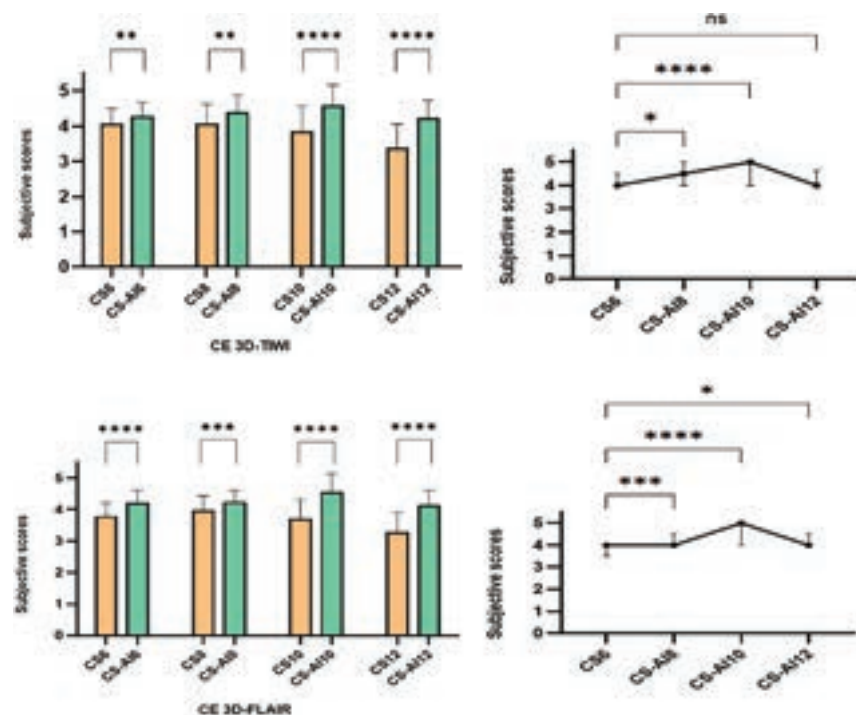


**FIG 3.** Quantitative analysis results of the SNR and CNR values of CS- and CS-AI-accelerated CE 3D-FLAIR images under different AFs. Note: NS indicates not significant; \*,  $P < .05$ ; \*\*,  $P < .01$ ; \*\*\*,  $P < .001$ ; \*\*\*\*,  $P < .0001$ .

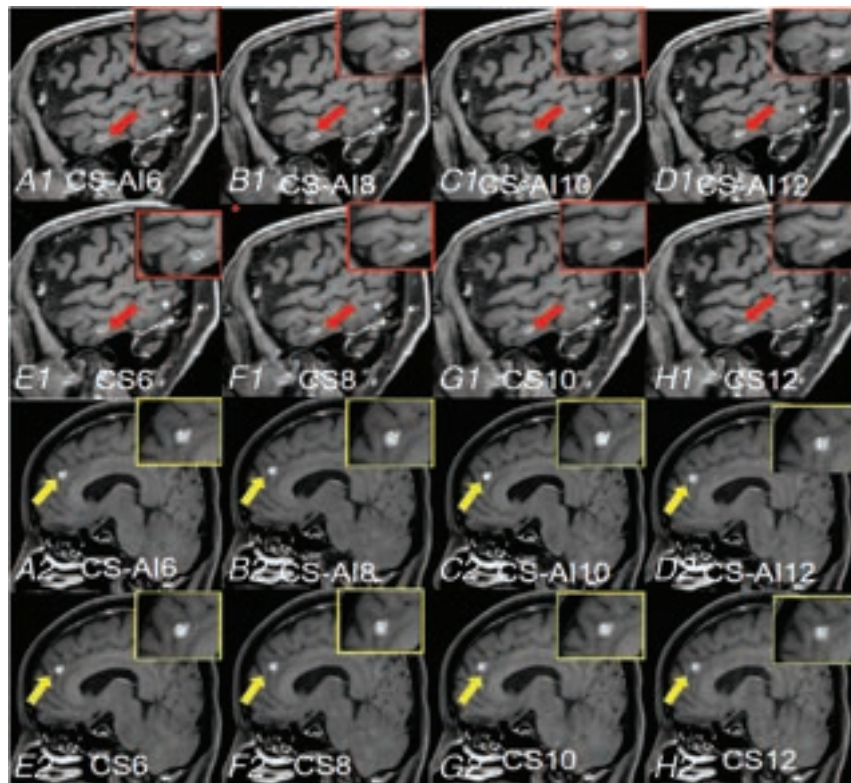
boundaries. However, there is a limit to denoising. When the level of undersampling is exceedingly high, even effective noise removal cannot fully compensate for the loss of image quality. The images tend to become blurred, and critical details are often lost, which

adversely affects the diagnostic quality of these images. These phenomena are supported by our subjective scores.

CE-T1WI and CE-FLAIR sequences play essential roles in detecting BM. Some studies have shown that CE-T1WI sequences



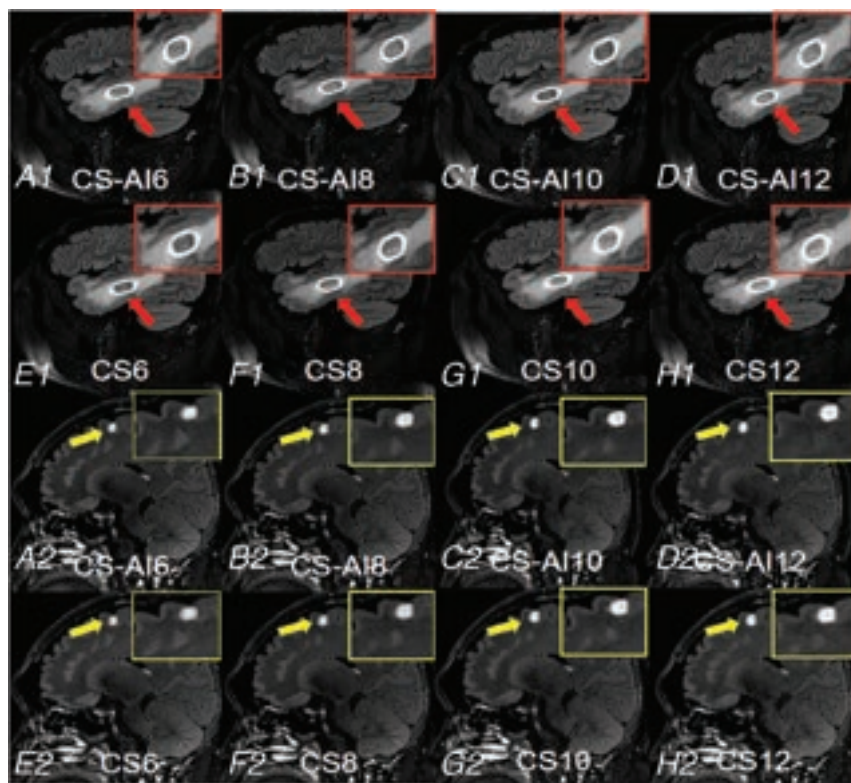
**FIG 4.** Comparison of the overall image quality score results of different accelerated sequences in patients with BM. Note: NS indicates not significant; \*,  $P < .05$ ; \*\*,  $P < .01$ ; \*\*\*,  $P < .001$ ; \*\*\*\*,  $P < .0001$ .



**FIG 5.** MR images of a 68-year-old woman with breast cancer and a 57-year-old man with lung cancer. Reconstructed CE 3D-T1-weighted sagittal images for CS-AI6 (A1, A2), CS-AI8 (B1, B2), CS-AI10 (C1, C2), CS-AI12 (D1, D2), reference CS6 (E1, E2), CS8 (F1, F2), CS10 (G1, G2), and CS12 (H1, H2). The arrows indicate enhanced BM lesions. Notably, the gray-white matter boundaries and lesion edges are clearer in the images accelerated by CS-AI technology, and the image quality is comparable with or better than that of CS6.

clearly show lesions in the brain parenchyma, whereas the CE-FLAIR sequence is more likely to detect BM on the surface of the cerebral cortex and leptomeningeal metastases and to help distinguish small lesions from cerebral vessels.<sup>24,25</sup> In addition, 3D sequences with higher spatial resolution are more advantageous for detecting small metastatic lesions when evaluating BM because 3D sequences can perform multiplanar reconstruction and reduce the partial volume effect compared with 2D sequences. However, the disadvantage of a 3D scan is that a longer scan time is required. Increasing scan times can be challenging for some patients, leading to possible image artifacts and even complete scan failures. Furthermore, longer acquisition times can increase the costs of clinical examinations by reducing patient throughput.<sup>26</sup>

Various MR imaging acceleration techniques have been developed to reduce image acquisition and reconstruction times, among which parallel imaging and compressed sensing are the most influential. Both methods improve the scanning speed by under-sampling the  $k$ -space. However, the larger the AF of parallel imaging, the



**FIG 6.** MR images of a 55-year-old male patient and a 38-year-old male patient with lung cancer. Reconstructed CE 3D-FLAIR sagittal images for CS-AI6 (A1, A2), CS-AI8 (B1, B2), CS-AI10 (C1, C2), CS-AI12 (D1, D2), reference CS6 (E1, E2), CS8 (F1, F2), CS10 (G1, G2), and CS12 (H1, H2). The arrows indicate enhanced BM. In the CS group, as the AF increases, the images become blurred, and the junctions of gray-white matter and the boundaries of the lesions become less clear.

lower the SNR of the image and the easier it is to produce fold-over artifacts. Though CS is superior to conventional SENSE-type parallel imaging in maintaining SNR and is less prone to generating artifacts, the image quality often degrades due to insufficient noise removal when the AF is higher than a certain level.<sup>27</sup> In this study, the CS group exhibited lower overall image quality scores for CS12-accelerated CE 3D-T1WI and CE 3D-FLAIR sequences compared with other AFs. Addressing these problems by applying a deep learning framework in the original data reconstruction is expected to reduce the scanning time while preserving image quality.

Deep learning is a popular machine learning approach that can automatically extract features and process high-dimensional medical image data.<sup>28</sup> Currently, this is the most widely used AI approach for medical images. One of its primary applications in medical imaging is accelerated MR imaging.<sup>29,30</sup> Recently, researchers have applied this method to sequence optimization. For instance, Park et al<sup>31</sup> demonstrated that fast MR imaging based on deep learning can be used to reduce the acquisition time of prostate MR imaging without compromising diagnostic performance. Zhao et al<sup>32</sup> used AI-assisted CS technology to accelerate T2-weighted kidney imaging, which obviously shortened the scanning time, and the image quality was equal to or better than that of traditional technology. Sheng et al<sup>33</sup> used a convolutional neural network framework to improve the reconstruction speed and image quality of single breath-hold T2-weighted sequences

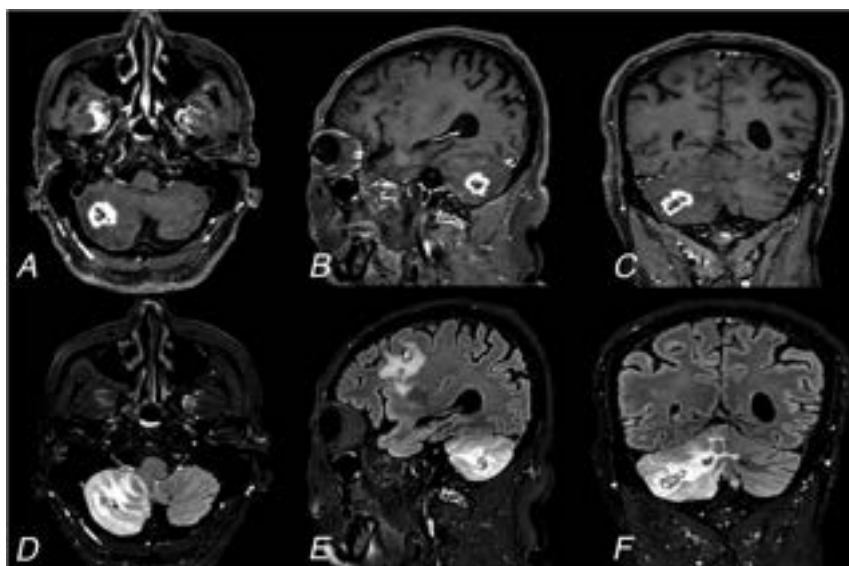
in liver scanning. This implies that image reconstruction functions in AI can further enhance the clinical applicability of time-consuming imaging sequences.

To our knowledge, this is the first study to comprehensively evaluate CS-AI for the imaging of patients with BM. Vranic et al<sup>7</sup> used CS to accelerate 3D sequence scanning in patients with brain tumors. However, that study utilized only 1 specific AF. While the findings underscore the superiority of CS over traditional scanning methods, the research did not determine the optimal AF, indicating a need for further exploration in this area. Zhang et al<sup>34</sup> evaluated the feasibility of CS with different AFs for the imaging of patients with brain tumors, but this was done only for 3D amide proton transfer-weighted sequences, and they concluded by recommending CS4. In this study, we obtained corresponding images for objective and subjective evaluations by applying CS and CS-AI to patients with BM and explored the optimal AF for clinical BM imaging. Notably, in this study, though the SNR and CNR of CS12 were better than or roughly similar to those of CS6, the image quality obtained after the reconstruction of the 2 sequences was inferior to that of CS6 (Fig 5 H1, H2; and Fig 6 H1, H2). This indicated that the quantitative parameters did not fully represent the overall image quality in terms of iterative image reconstruction and denoising. Objective and subjective evaluations must be combined to comprehensively analyze image diagnostic performance.

This study has some limitations. First, this was a single-center prospective study with a small sample size that exclusively used single vendor equipment. Due to the incompatibility of CS-AI with other manufacturers' devices, its applicability is restricted to Philips MR imaging systems. In addition, the study did not explore the use of this technology with various MR imaging coils or lower field strength systems, nor did it explore its generalizability to similar equipment in different hospital settings. Therefore, future research should consider these aspects to further enhance the understanding and application of CS-AI in various MR imaging contexts.

Second, BM were mainly confirmed during follow-ups. Patients with multiple BM frequently do not undergo surgery, so lesions are rarely pathologically confirmed. Third, CS- or CS-AI-accelerated scans with different AFs were performed in a random acquisition order to reduce time bias. This may have influenced the enhancement of the lesions to some extent; however, the 2 raters considered no significant difference in enhancement between the observed images. Finally, 42 of the 51 patients included in this study had lung cancer. Theoretically, these findings can only be applied to BM due to





**FIG 7.** A 60-year-old male patient with lung cancer was scanned with 2 optimized sequences to obtain example images of a good quality in a short time (5:03 minutes). The axial (A), sagittal (B), and coronal (C) images of accelerated CE 3D-T1WI of CS-AI10. The axial (D), sagittal (E), and coronal (F) images of accelerated CE 3D-FLAIR of CS-AI10.

lung cancer. Therefore, more heterogeneous populations are needed to confirm these results in patients with other metastatic tumors.

## CONCLUSIONS

Combining CS and deep learning can shorten scanning time while maintaining or improving image quality compared with conventional CS. Using CS-AI with AFs up to 10 to reconstruct CE 3D-T1WI and CE 3D-FLAIR sequences on a 3T MR machine could reduce scanning time and improve diagnostic performance. When CS-AI10 (Fig 7) was used for acceleration, CE 3D-T1WI and CE 3D-FLAIR sequences were 39.25% (2:10 minutes versus 3:34 minutes) and 39.93% (2:53 minutes versus 4:48 minutes) faster than CS6, respectively. The use of CS-AI technology could optimize CE 3D scans when performing clinical imaging of patients with BM to help radiologists better evaluate lesions in detail while improving scanning efficiency.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

- Kaufmann TJ, Smits M, Boxerman J, et al. Consensus recommendations for a standardized brain tumor imaging protocol for clinical trials in brain metastases. *Neuro Oncol* 2020;22:757–72 CrossRef Medline
- Li R, Guo Y, Zhao Z, et al. MRI-based two-stage deep learning model for automatic detection and segmentation of brain metastases. *Eur Radiology* 2023;33:3521–31 CrossRef Medline
- Derks SH, van der Veldt AA, Smits M. Brain metastases: the role of clinical imaging. *Br J Radiology* 2022;95:20210944 CrossRef Medline
- Villanueva-Meyer JE, Mabray MC, Cha S. Current clinical brain tumor imaging. *Neurosurgery* 2017;81:397–415 CrossRef Medline
- Kim D, Heo YJ, Jeong HW, et al. Usefulness of the delay alternating with nutation for tailored excitation pulse with T1-weighted sampling perfection with application-optimized contrasts using different flip angle evolution in the detection of cerebral metastases: Comparison with MPRAGE imaging. *AJNR Am J Neuroradiol* 2019;40:1469–75 CrossRef Medline
- Langer CJ, Mehta MP. Current management of brain metastases, with a focus on systemic options. *J Clin Oncol* 2005;23:6207–19 CrossRef Medline
- Vranic JE, Cross NM, Wang Y, et al. Compressed sensing-sensitivity encoding (CS-SENSE) accelerated brain imaging: reduced scan time without reduced image quality. *AJNR Am J Neuroradiol* 2019;40:92–98 CrossRef Medline
- Duan Y, Zhang J, Zhuo Z, et al. Accelerating brain 3D T1-weighted turbo field echo MRI using compressed sensing-sensitivity encoding (CS-SENSE). *Eur J Radiology* 2020;131:109255 CrossRef Medline
- Sasi SD, Ramanikharan AK, Bhattacharjee R, et al. Evaluating feasibility of high resolution T1-perfusion MRI with whole brain coverage using compressed SENSE: Application to glioma grading. *Eur J Radiology* 2020;129:109049 CrossRef Medline
- Boyarko AC, Dillman JR, Tkach JA, et al. Comparison of compressed SENSE and SENSE for quantitative liver MRI in children and young adults. *Abdom Radiology (NY)* 2021;46:4567–75 CrossRef Medline
- Cho SJ, Choi YJ, Chung SR, et al. High-resolution MRI using compressed sensing-sensitivity encoding (CS-SENSE) for patients with suspected neurovascular compression syndrome: comparison with the conventional SENSE parallel acquisition technique. *Clin Radiology* 2019;74:817.e9–e14 CrossRef Medline
- Hsiao A, Lustig M, Alley MT, et al. Evaluation of valvular insufficiency and shunts with parallel-imaging compressed-sensing 4D phase-contrast MR imaging with stereoscopic 3D velocity-fusion volume-rendered visualization. *Radiology* 2012;265:87–95 CrossRef Medline
- Knoll F, Zbontar J, Sriram A, et al. fastMRI: A publicly available raw k-space and DICOM dataset of knee images for accelerated MR image reconstruction using machine learning. *Radiology Artif Intell* 2020;2:e190007 CrossRef Medline
- Knoll F, Murrell T, Sriram A, et al. Advancing machine learning for MR image reconstruction with an open competition: Overview of the 2019 fastMRI challenge. *Magn Reson Med* 2020;84:3054–70 CrossRef Medline
- Pezzotti N, Yousefi S, Elmahdy MS, et al. An adaptive intelligence algorithm for undersampled knee MRI reconstruction. *IEEE Access* 2020;8:204825–38 CrossRef
- Foreman SC, Neumann J, Han J, et al. Deep learning-based acceleration of compressed sense MR imaging of the ankle. *Eur Radiology* 2022;32:8376–85 CrossRef Medline
- Fervers P, Zaeske C, Rauen P, et al. Conventional and deep-learning-based image reconstructions of undersampled K-space data of the lumbar spine using compressed sensing in MRI: A comparative study on 20 subjects. *Diagnostics (Basel)* 2023;13:418 CrossRef Medline
- Dratsch T, Siedek F, Zäske C, et al. Reconstruction of shoulder MRI using deep learning and compressed sensing: a validation study on healthy volunteers. *Eur Radiology Exp* 2023;7:66 CrossRef Medline
- Ni M, He M, Yang Y, et al. Application research of AI-assisted compressed sensing technology in MRI scanning of the knee joint: 3D-MRI perspective. *Eur Radiology* 2023 CrossRef Medline
- Dietrich O, Raya JG, Reeder SB, et al. Measurement of signal-to-noise ratios in MR images: influence of multichannel coils, parallel



- imaging, and reconstruction filters. *J Magn Reson Imaging* 2007;26:375–85 CrossRef Medline
21. Kakeda S, Korogi Y, Hiai Y, et al. **Detection of brain metastasis at 3T: comparison among SE, IR-FSE and 3D-GRE sequences.** *Eur Radiology* 2007;17:2345–51 CrossRef Medline
22. Conklin J, Longo MGF, Cauley SF, et al. **Validation of highly accelerated wave-CAIPI SWI compared with conventional SWI and T2\*-weighted gradient recalled-echo for routine clinical brain MRI at 3T.** *AJNR Am J Neuroradiol* 2019;40:2073–80 CrossRef Medline
23. Chung MS, Yim Y, Sung JK, et al. **CS-VIBE accelerates cranial nerve MR imaging for the diagnosis of facial neuritis: comparison of the diagnostic performance of post-contrast MPRAGE and CS-VIBE.** *Eur Radiology* 2022;32:223–33 CrossRef Medline
24. Jin T, Zhang H, Liu X, et al. **Enhancement degree of brain metastases: correlation analysis between enhanced T2 FLAIR and vascular permeability parameters of dynamic contrast-enhanced MRI.** *Eur Radiology* 2021;31:5595–604 CrossRef Medline
25. Ercan N, Gultekin S, Celik H, et al. **Diagnostic value of contrast-enhanced fluid-attenuated inversion recovery MR imaging of intracranial metastases.** *AJNR Am J Neuroradiol* 2004;25:761–65 Medline
26. Toledano-Massiah S, Sayadi A, de Boer R, et al. **Accuracy of the compressed sensing accelerated 3D-FLAIR sequence for the detection of MS plaques at 3T.** *AJNR Am J Neuroradiol* 2018;39:454–58 CrossRef Medline
27. Dieckmeyer M, Roy AG, Senapati J, et al. **Effect of MRI acquisition acceleration via compressed sensing and parallel imaging on brain volumetry.** *MAGMA* 2021;34:487–97 CrossRef Medline
28. Schlemper J, Caballero J, Hajnal JV, et al. **A deep cascade of convolutional neural networks for dynamic MR image reconstruction.** *IEEE Trans Med Imaging* 2018;37:491–503 CrossRef Medline
29. Gu H, Yaman B, Moeller S, et al. **Revisiting [Formula: see text]-wavelet compressed-sensing MRI in the era of deep learning.** *Proc Natl Acad Sci U S A* 2022;119:e2201062119 CrossRef Medline
30. Bash S, Wang L, Airriess C, et al. **Deep learning enables 60% accelerated volumetric brain MRI while preserving quantitative performance: A prospective, multicenter, multireader trial.** *AJNR Am J Neuroradiol* 2021;42:2130–37 CrossRef Medline
31. Park JC, Park KJ, Park MY, et al. **Fast T2-weighted imaging with deep learning-based reconstruction: Evaluation of image quality and diagnostic performance in patients undergoing radical prostatectomy.** *J Magn Reson Imaging* 2022;55:1735–44 CrossRef Medline
32. Zhao Y, Peng C, Wang S, et al. **The feasibility investigation of AI-assisted compressed sensing in kidney MR imaging: an ultra-fast T2WI imaging technology.** *BMC Med Imaging* 2022;22:119 CrossRef Medline
33. Sheng RF, Zheng LY, Jin KP, et al. **Single-breath-hold T2WI liver MRI with deep learning-based reconstruction: A clinical feasibility study in comparison to conventional multi-breath-hold T2WI liver MRI.** *Magn Reson Imaging* 2021;81:75–81 CrossRef Medline
34. Zhang N, Zhang H, Gao B, et al. **3D amide proton transfer weighted brain tumor imaging with compressed SENSE: Effects of different acceleration factors.** *Front Neurosci* 2022;16:876587 CrossRef Medline

# Arterial Spin-Labeling and DSC Perfusion Metrics Improve Agreement in Neuroradiologists' Clinical Interpretations of Posttreatment High-Grade Glioma Surveillance MR Imaging—An Institutional Experience

Ghiam Yamin, Eric Tranvinh, Bryan A. Lanzman, Elizabeth Tong, Syed S. Hashmi, Chirag B. Patel, and Michael Iv



## ABSTRACT

**BACKGROUND AND PURPOSE:** MR perfusion has shown value in the evaluation of posttreatment high-grade gliomas, but few studies have shown its impact on the consistency and confidence of neuroradiologists' interpretation in routine clinical practice. We evaluated the impact of adding MR perfusion metrics to conventional contrast-enhanced MR imaging in posttreatment high-grade glioma surveillance imaging.

**MATERIALS AND METHODS:** This retrospective study included 45 adults with high-grade gliomas who had posttreatment perfusion MR imaging. Four neuroradiologists assigned Brain Tumor Reporting and Data System scores for each examination on the basis of the interpretation of contrast-enhanced MR imaging and then after the addition of arterial spin-labeling-CBF, DSC-relative CBV, and DSC-fractional tumor burden. Interrater agreement and rater agreement with a multidisciplinary consensus group were assessed with  $\kappa$  statistics. Raters used a 5-point Likert scale to report confidence scores. The frequency of clinically meaningful score changes resulting from the addition of each perfusion metric was determined.

**RESULTS:** Interrater agreement was moderate for contrast-enhanced MR imaging alone ( $\kappa = 0.63$ ) and higher with perfusion metrics (arterial spin-labeling-CBF,  $\kappa = 0.67$ ; DSC-relative CBV,  $\kappa = 0.66$ ; DSC-fractional tumor burden,  $\kappa = 0.70$ ). Agreement between raters and consensus was highest with DSC-fractional tumor burden ( $\kappa = 0.66$ – $0.80$ ). Confidence scores were highest with DSC-fractional tumor burden. Across all raters, the addition of perfusion resulted in clinically meaningful interpretation changes in 2%–20% of patients compared with contrast-enhanced MR imaging alone.

**CONCLUSIONS:** Adding perfusion to contrast-enhanced MR imaging improved interrater agreement, rater agreement with consensus, and rater confidence in the interpretation of posttreatment high-grade glioma MR imaging, with the highest agreement and confidence scores seen with DSC-fractional tumor burden. Perfusion MR imaging also resulted in interpretation changes that could change therapeutic management in up to 20% of patients.

**ABBREVIATIONS:** ASL = arterial spin-labeling; BT-RADS = Brain Tumor Reporting and Data System; CE = contrast-enhanced; CRT = chemoradiotherapy; FTB = fractional tumor burden; GBM = glioblastoma; HGG = high-grade glioma; IRA = interrater agreement; rCBV = relative CBV

Surveillance MR imaging following surgery and chemoradiotherapy (CRT) is integral to the care of patients with high-grade gliomas (HGGs). In the Response Assessment in Neuro-Oncology criteria, the assessment of tumor status is based on

changes in contrast-enhanced (CE) T1-weighted images and T2 FLAIR signal across time, taking into consideration clinical factors such as functional status, interscan therapy, and time since completion of CRT.<sup>1,2</sup> However, interpretation of posttreatment HGG MR imaging is often challenging because of the overlapping imaging findings between tumor progression and CRT effects.

Perfusion metrics are not yet universally included in the assessment, in part due to a perceived lack of standardization and validation, but studies have shown them to be valuable in the evaluation of posttreatment HGGs. DSC perfusion-derived relative CBV (rCBV), which is the most widely used perfusion metric in brain tumor imaging, is a biomarker for tumor angiogenesis and can distinguish HGG recurrence from radiation effects.<sup>3-5</sup> Arterial spin-labeling (ASL) perfusion-derived CBF has also been shown to be higher in recurrent HGGs than in treatment-related

Received July 6, 2023; accepted after revision November 16.

From the Department of Radiology (G.Y., E. Tranvinh, B.A.L., E. Tong, S.S.H., M.I.), Division of Neuroimaging and Neurointervention, Stanford University Medical Center, Stanford, California; and Department of Neuro-Oncology (C.B.P.), The University of Texas MD Anderson Cancer Center, Houston, Texas.

Ghiam Yamin and Eric Tranvinh are co-first authors and contributed equally to this work.

Please address correspondence to Michael Iv, MD, Stanford University Medical Center, Center for Academic Medicine, Radiology +; MC: 5659, Room 323A, 453 Quarry Rd, Palo Alto, CA 94304; e-mail: miv@stanford.edu; @Michael\_Iv\_MD

Indicates open access to non-subscribers at [www.ajnr.org](http://www.ajnr.org)

<http://dx.doi.org/10.3174/ajnr.A8190>

effects and correlates well with DSC-derived rCBV.<sup>6-8</sup> DSC perfusion-derived fractional tumor burden (FTB) can delineate areas of low, intermediate, and high likelihood of tumor burden within a contrast-enhancing volume on the basis of predefined rCBV thresholds.<sup>9</sup> Fewer studies, however, have shown the positive clinical impact of adding MR perfusion to conventional CE-MR imaging in the assessment of posttreatment HGG surveillance imaging in real-time practice. Geer et al<sup>10</sup> showed that the addition of perfusion metrics to CE-MR imaging changed the management plan in 8.5% of patients and increased confidence in the management plan by 57.6%. Iv et al<sup>9</sup> showed that the addition of FTB can influence clinical decision-making among a panel of physicians involved in the care of patients with gliomas.

In this study, we investigated the impact of adding ASL-derived CBF, DSC-derived rCBV, and DSC-derived FTB to conventional CE-MR imaging on neuroradiologists' clinical interpretations of posttreatment MR imaging of HGGs. We hypothesized that the addition of perfusion metrics would improve interrater agreement (IRA) among neuroradiologists and between neuroradiologists and an experienced multidisciplinary group, improve confidence in interpretation, and yield clinically meaningful changes in interpretation.

## MATERIALS AND METHODS

### Patients

This retrospective study was approved by the Stanford University institutional review board. The records of patients with treated HGGs who had MR imaging performed between May 2019 and December 2019 were reviewed. Inclusion criteria were as follows: 1) 18 years of age or older; 2) pathologically proved *IDH* wild-type HGG (we used the integrated tumor diagnosis from the time of the original histopathologic evaluation); 3) prior treatment with surgical resection followed by CRT; and 4) posttreatment MR imaging with ASL and DSC. For each patient, the first MR images with complete ASL and DSC acquisitions and postprocessed imaging data were selected. Patients were excluded if either the perfusion imaging set was incomplete or the contrast-enhancing lesion was not located within the brain parenchyma. *IDH*-mutant gliomas were also excluded because their biology is different from that of *IDH* wild-type glioblastomas (GBM). Clinical and histopathologic information was obtained through the electronic medical records.

### MR Imaging Acquisition

MR imaging was performed on 1.5T (Signa Explorer; GE Healthcare) and 3T (Discovery MR750; GE Healthcare) scanners using our institutional glioma-specific MR imaging protocol. While this protocol has undergone slight changes with time, the sequences relevant to this study have remained constant: 3D T1 fast-spoiled gradient recalled-echo brain volume (BRAVO), 3D T2 FLAIR, ASL, DSC, axial T2, and 3D T1 BRAVO postcontrast images.

ASL was performed using pseudocontinuous labeling with the following parameters: TR/TE/label time/postlabel delay = 5500/2.5/1500/2000 ms, 3D background-suppressed fast-spin-echo stack-of-spirals readout, and 4-mm in-plane and 6-mm through-plane

resolution. ASL postprocessing was performed by an automated reconstruction script that sent CBF images to PACS.

Before acquisition of DSC perfusion images, a 0.05-mmol/kg dose of gadobenate dimeglumine was administered as a preload to help correct for leakage effects.<sup>11</sup> DSC perfusion was performed with a 4-mL/s bolus of 0.05-mmol/kg gadobenate dimeglumine and single-echo gradient-echo-planar imaging using the following parameters: TR/TE = 1800/35–40 ms, section thickness = 5 mm, no interslice gaps with 20 slices covering the brain, flip angle = 30°, matrix = 128 × 128 mm, FOV = 240 mm.

### DSC Image Processing

DSC perfusion images were postprocessed by our 3D and Quantitative Imaging Laboratory at a workstation equipped with Aycan (Version 3.16.002) and IB Neuro (Version 2.0; Imaging Biometrics), a commercially available image-processing plug-in that implements a leakage-correction algorithm and generates standardized rCBV and FTB maps.<sup>12-16</sup> Semiautomated image segmentation and processing were performed with IB Rad Tech (Version 2.0; Imaging Biometrics), a workflow engine that generates quantitative  $\Delta$  T1 and FTB maps from the IB Delta Suite (Version 2.0; Imaging Biometrics) and IB Neuro plug-in. This workflow has been described in detail previously.<sup>9,16</sup> Two standardized rCBV thresholds (1 and 1.56) were used to define 3 FTB classes: FTB<sub>low</sub>, percentage of contrast-enhancing voxels with rCBV  $\leq 1.0$ ; FTB<sub>mid</sub>, percentage of voxels with rCBV between 1.0 and 1.56; and FTB<sub>high</sub>, percentage of voxels with rCBV of  $\geq 1.56$ . A standardized rCBV value of 1.56 was used as the higher threshold for tumor prediction considering our own clinical experience and a prior report of this value indicating >88% probability of tumor.<sup>17</sup> Mean and median rCBV values of the contrast-enhancing volumes were generated for each patient. Volumetric images of the contrast-enhancing lesion superimposed on the FTB map containing colored voxels of each class (FTB<sub>low</sub> = blue; FTB<sub>mid</sub> = yellow; FTB<sub>high</sub> = red) and a histogram displaying voxels for the entire contrast-enhancing volume were also produced.

### MR Imaging Interpretation and Brain Tumor Reporting and Data System Scoring

Studies were interpreted by 4 neuroradiology faculty with 1 (E. Tong), 3 (S.S.H.), 4 (B.A.L.) and 6 (E. Tranvinh) years of practice experience following 1 year of a dedicated neuroradiology fellowship at the time of the study. The 2 raters with fewer experience years also had less experience with perfusion imaging than the other 2 raters. Each rater was provided a worklist of 45 de-identified patients in PACS, including the MR imaging of interest and 2 of the most recent prior MR images for comparison, during 1 session. Clinical information such as tumor histology and molecular status and a brief treatment course was also provided. For each patient, raters provided scores using the Brain Tumor Reporting and Data System (BT-RADS).<sup>2</sup> In BT-RADS, MR imaging is assigned a score from 0 to 4 based on changes in the contrast-enhancing lesion and T2 FLAIR signal compared with the most recent prior MR imaging, taking into consideration the time since completion of CRT and treatment with adjuvant medications such as bevacizumab and steroids: 0, baseline; 1a, improving imaging findings due to decreased tumor burden and/or treatment

effect; 1b, improving imaging findings due to bevacizumab or steroids; 2, no change; 3a, worsening imaging findings thought to represent treatment effects; 3b, worsening imaging findings representing an indeterminate mix of tumor and treatment effects; 3c, worsening imaging findings thought to represent increasing tumor burden; and 4, worsening imaging findings highly suspicious for tumor progression.

Considering minor differences between the following pairs of scores in their influence on patient management, BT-RADS scores 1a and 1b were combined into a single 1a/b category and scores 3c and 4 were combined into a single 3c/4 category to facilitate the determination of whether the addition of perfusion imaging resulted in clinically meaningful changes in interpretation. A clinically meaningful change in interpretation was defined as a change in score from  $\leq 3a$  to  $\geq 3b$  (and vice versa) and from  $\leq 3b$  to 3c/4 (and vice versa), because these changes reflected a change in the proportion and confidence of predicted tumor burden. BT-RADS scores are associated with management recommendations, and the higher the predicted tumor burden, the more likely it is that there will be a recommended change in management.<sup>2</sup> In practice, the stability or worsening of clinical symptoms may also help to guide the management. Often, worsening clinical symptoms coupled with worsening imaging findings, particularly if there is any tumor concern (eg, score of  $\geq 3b$ ), may favor an earlier change in management (eg, administration of steroids or bevacizumab or surgical resection) over just shortening the surveillance imaging time.

Each rater provided an initial BT-RADS score based only on conventional CE-MR imaging before any perfusion maps were evaluated. In the same session, each rater then provided additional scores after reviewing each of the ASL-CBF, DSC-rCBV, and DSC-FTB maps. The order in which the perfusion maps were reviewed was randomized from patient to patient by a study member not involved in image interpretation and scoring (G.Y.). This process resulted in a total of 4 scores per patient. Quantitative and qualitative imaging data were provided for DSC-rCBV and DSC-FTB as previously mentioned. Quantitative values were not available for ASL-CBF, reflecting the lack of a clinically validated threshold for differentiating tumor and treatment effect. Raters then provided a confidence score for each BT-RADS assessment. Confidence scores were based on a 5-point Likert scale (1, not confident; 2, less confident; 3, average confidence; 4, more confident; and 5, very confident).

### Reference Standard

Because histopathology was not available at each MR imaging timepoint, a reference standard BT-RADS score was determined for each patient by an experienced multidisciplinary consensus group comprising a board-certified neuroradiologist with 8 years of postfellowship experience (M.I.) and a board-eligible neuro-oncologist (C.B.P.), both of whom attend our institution's weekly neuro-oncology tumor board and are regularly involved in patient care-related decision-making. They each assessed the MR imaging, including perfusion data, relevant priors, and clinical information. If there was disagreement, follow-up MR imaging and any available pathology results during the

**Table 1: Patient demographics**

Demographics	
Age (yr)	
Mean (SD)	61 (13)
Range	31–88
Sex	
Male	25 (56%)
Female	20 (44%)
Integrated diagnosis	
GBM, <i>IDH</i> wild-type, WHO 4	44 (98%)
Astrocytoma, <i>IDH</i> wild-type, WHO 3	1 (2%)
HGG molecular features	
<i>IDH</i> wild-type	45 (100%)
MGMT promoter-unmethylated	24 (53%)
MGMT promoter-methylated	19 (42%)
Unknown MGMT promoter methylation status	2 (4%)

**Note:**—WHO indicates World Health Organization.

follow-up period were reviewed to determine a single consensus score.

### Statistical Analysis

The IRA among the 4 raters was determined using the Fleiss  $\kappa$ . Agreement between the consensus and each rater's scores was assessed using the Cohen  $\kappa$ . The strength of agreement was interpreted using the following scale:  $<0.2$ , poor;  $0.2$ – $0.39$ , slight;  $0.4$ – $0.59$ , fair;  $0.6$ – $0.79$ , moderate;  $0.8$ – $0.99$ , substantial; 1, perfect. The Kruskal-Wallis test was used to evaluate whether ASL-CBF, DSC-rCBV, and DSC-FTB when added to CE-MR imaging had an effect on the number of clinically meaningful score changes. Randomizing the order of perfusion map interpretations made observations in each group and between groups independent. Descriptive statistics were used to evaluate raters' confidence in the BT-RADS scoring. While confidence scores were assessed with an ordinal scale (Likert scale), the mean of scores was calculated to interpret trends; in other words, for clarity of interpretation, we treated the scale as an approximation to an interval scale. In addition, for each rater, the Friedman test was used to determine whether there was a difference in confidence scores among the dependent variables of CE-MR imaging, CE-MR imaging+ASL-CBF, CE-MR imaging+DSC-rCBV, and CE-MR imaging+DSC-FTB. The post hoc Dunn-Bonferroni test was performed to detect pair-wise differences among groups. For all analyses,  $P < .05$  was considered significant. Statistical analyses were performed with the DATAtab online Statistics Calculator (<http://www.datatab.net>).

## RESULTS

### Patients

Ninety-three MR imaging examinations were eligible for the study. Thirty-one MR imaging examinations were excluded because they were follow-up examinations of patients already included in the study, leaving 62 unique patient examinations. Ten examinations were excluded for lack of postprocessed DSC data, and 1 was excluded for lack of ASL. One patient was excluded because the enhancing lesion was outside the brain parenchyma. Five *IDH*-mutant tumors were excluded from the analysis. The final cohort consisted of 45 patients (25 men, 20 women; mean age, 61 [SD, 13] years; range, 31–88 years) (Table 1). Of 45 HGGs, 44 were *IDH* wild-type grade 4 GBMs, and 1



was an *IDH* wild-type grade 3 astrocytoma. The median number of days between the first histologic diagnosis and the perfusion MR imaging assessed was 296 days (range, 70–2777 days).

**Agreement in MR Imaging Interpretation among Raters**

The IRA was 0.63 (95% CI, 0.56–0.69; *P* < .001) for CE-MR imaging alone; 0.67 (95% CI, 0.60–0.74; *P* < .001) for CE-MR imaging+ASL-CBF; 0.66 (95% CI, 0.60–0.73; *P* < .001) for CE-MR imaging+DSC-rCBV; and 0.70 (95% CI, 0.63–0.77; *P* < .001) for CE-MR imaging+DSC-FTB (Table 2). Figures 1 and 2 show BT-RADS scores by raters and consensus in representative patients with posttreatment GBMs on surveillance MR imaging.

**Agreement in MR Imaging Interpretation between the Multidisciplinary Consensus Group and Raters**

Scores from the consensus group yielded 10 patients with a BT-RADS score of 1a/1b, 9 with a score of 2, one with a score of 3a,

**Table 2: Agreement in MR imaging interpretation among 4 neuroradiologists**

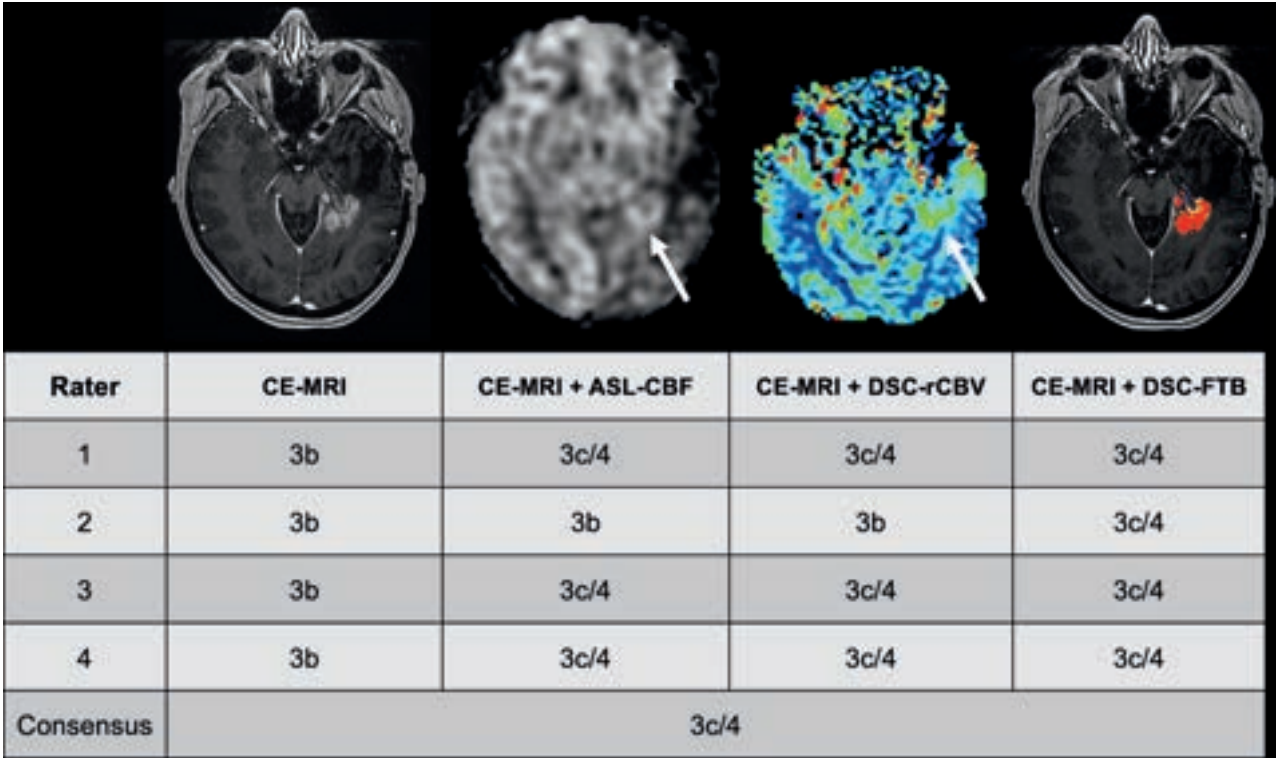
	Fleiss $\kappa$	95% CI	<i>P</i> Value
Conventional CE-MR imaging	0.63	0.56–0.69	<.001
CE-MR imaging + ASL-CBF	0.67	0.60–0.74	<.001
CE-MR imaging + DSC-rCBV	0.66	0.60–0.73	<.001
CE-MR imaging + DSC-FTB	0.70	0.63–0.77	<.001

four with a score of 3b, and 21 with a score of 3c/4. None were assigned a score of 0.

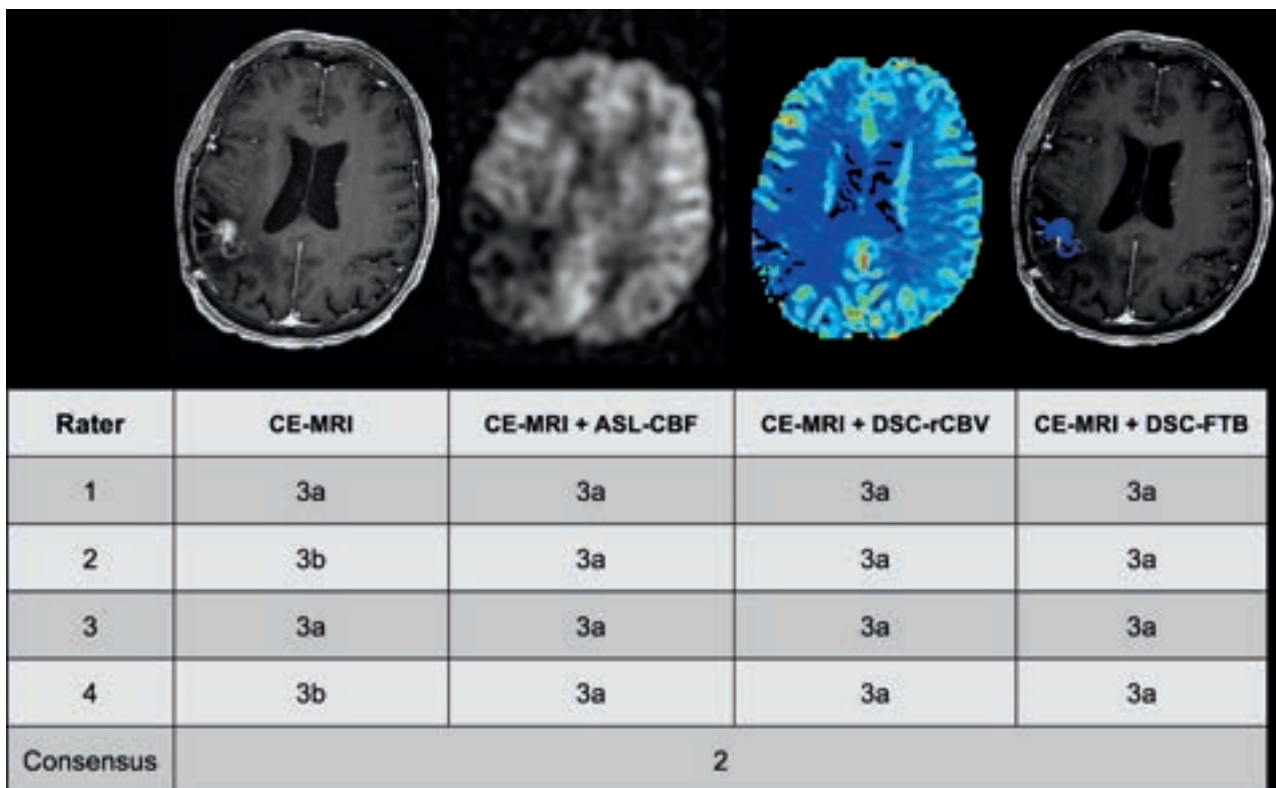
The Cohen  $\kappa$  was used to determine agreement between consensus and individual raters’ scores (Table 3). For rater 1,  $\kappa$  scores for CE-MR imaging alone, CE-MR imaging+ASL-CBF, CE-MR imaging+DSC-rCBV, and CE-MR imaging+DSC-FTB were 0.53 (95% CI, 0.31–0.75), 0.58 (95% CI, 0.36–0.80), 0.58 (95% CI, 0.36–0.80), and 0.66 (95% CI, 0.46–0.87), respectively. For rater 2,  $\kappa$  scores were 0.70 (95% CI, 0.50–0.91), 0.69 (95% CI, 0.48–0.89), 0.71 (95% CI, 0.51–0.90), and 0.80 (95% CI, 0.63–0.97), respectively. For rater 3,  $\kappa$  scores were 0.63 (95% CI, 0.43–0.84), 0.61 (95% CI, 0.39–0.82), 0.63 (95% CI, 0.42–0.84), and 0.66 (95% CI, 0.46–0.86), respectively. For rater 4,  $\kappa$  scores were 0.58 (95% CI, 0.36–0.80), 0.65 (95% CI, 0.44–0.86), 0.68 (95% CI, 0.47–0.88), and 0.73 (95% CI, 0.55–0.92), respectively. All analyses showed *P* < .001.

**Clinically Meaningful Changes in BT-RADS Scores with Perfusion Imaging**

The number of clinically meaningful score changes following the addition of perfusion metrics relative to conventional CE-MR imaging alone is shown in Table 4 and illustrated in Fig 3. For rater 1, the number of score changes occurred in 11% (5/45) of patients with CE-MR imaging+ASL-CBF, 11% (5/45) with CE-MR imaging+DSC-rCBV, and 18% (8/45) with CE-MR imaging+DSC-FTB. For rater 2, the number of score changes



**FIG 1.** Example of rater BT-RADS scores in a 57-year-old woman with previously treated GBM and worsening findings on surveillance MR imaging. The T1 postgadolinium image demonstrates an enhancing lesion in the left mesial temporal lobe. The lesion has elevated ASL-CBF and DSC-rCBV (white arrows). The DSC-FTB image shows that the enhancing voxels are in the “high” fractional tumor burden (red voxels) category. The addition of perfusion metrics to CE-MR imaging resulted in a scoring upgrade from 3b (worsening imaging findings, indeterminate mix of treatment effects and tumor) to 3c/4 (likely tumor progression) across all raters and agreed with the consensus score of 3c/4. For 3 of 4 raters, the upgrade occurred with all perfusion metrics, and for rater 2, it occurred only with DSC-FTB.



**FIG 2.** Example of rater BT-RADS scores in a 61-year-old woman with previously treated GBM and equivocally worsening findings on surveillance MR imaging. The T1 postgadolinium image demonstrates an enhancing lesion in the right temporal lobe. The lesion shows no elevated ASL-CBF or DSC-rCBV, and DSC-FTB shows that the enhancing voxels are in the “low” FTB (blue voxels) category. For two raters, the addition of perfusion metrics to CE-MR imaging resulted in a scoring downgrade from 3b (worsening imaging findings, indeterminate mix of treatment effects and tumor) to 3a (worsening imaging findings, likely treatment effects). For the other raters, perfusion metrics did not influence their assessment. The consensus score in this case was 2 (no change). The discrepancy between the consensus group and the raters was due to differences in opinion as to whether the enhancing lesion had subtly increased in size from the prior MR imaging (not shown).

**Table 3: Agreement in MR imaging interpretation between an experienced multidisciplinary consensus group and each of 4 neuroradiologists<sup>a</sup>**

	More Experienced		Less Experienced	
	Rater 1	Rater 2	Rater 3	Rater 4
Conventional CE-MR imaging	0.53 (0.31–0.75)	0.70 (0.50–0.91)	0.63 (0.43–0.84)	0.58 (0.36–0.80)
CE-MR imaging + ASL-CBF	0.58 (0.36–0.80)	0.69 (0.48–0.89)	0.61 (0.39–0.82)	0.65 (0.44–0.86)
CE-MR imaging + DSC-rCBV	0.58 (0.36–0.80)	0.71 (0.51–0.90)	0.63 (0.42–0.84)	0.68 (0.47–0.88)
CE-MR imaging + DSC-FTB	0.66 (0.46–0.87)	0.80 (0.63–0.97)	0.66 (0.46–0.86)	0.73 (0.55–0.92)

<sup>a</sup> All analyses showed  $P < .001$ . Values are Cohen  $\kappa$  with 95% confidence intervals in the parentheses.

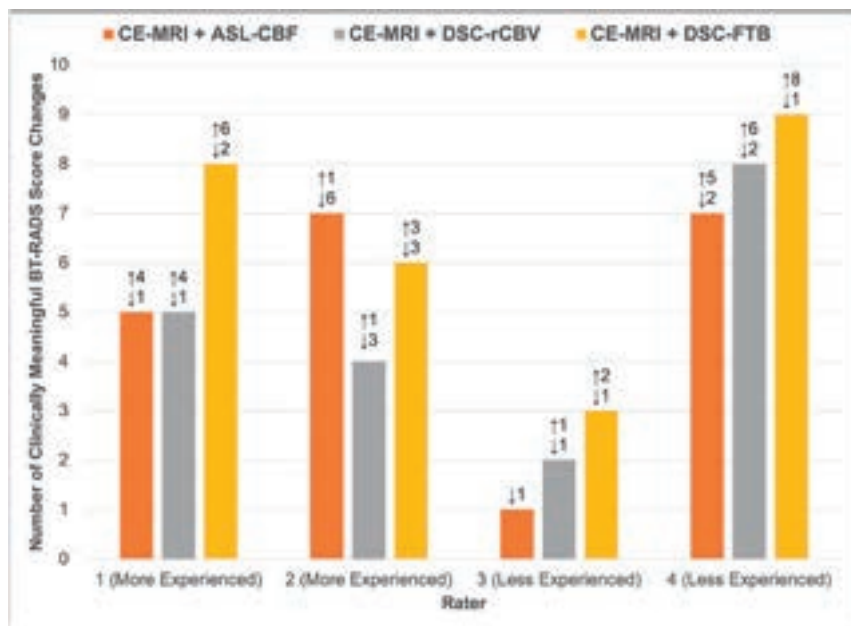
**Table 4: Frequency of clinically meaningful changes in BT-RADS scores in 45 patients following the inclusion of perfusion metrics compared with conventional CE-MR imaging alone<sup>a</sup>**

	More Experienced		Less Experienced	
	Rater 1	Rater 2	Rater 3	Rater 4
CE-MR imaging + ASL-CBF	5 (11%)	7 (16%)	1 (2%)	7 (16%)
CE-MR imaging + DSC-rCBV	5 (11%)	4 (9%)	2 (4%)	8 (18%)
CE-MR imaging + DSC-FTB	8 (18%)	6 (13%)	3 (7%)	9 (20%)

<sup>a</sup> No significance was found among ASL-CBF, DSC-rCBV, and DSC-FTB when added to CE-MR imaging with respect to the number of clinically meaningful score changes ( $P = .53$ ). Percentages are in parentheses.

occurred in 16% (7/45), 9% (4/45), and 13% (6/45) of patients, respectively. For rater 3, the number of score changes occurred in 2% (1/45), 4% (2/45), and 7% (3/45) of patients, respectively. For rater 4, the number of score changes occurred in 16% (7/45), 18% (8/45), and 20% (9/45) of patients, respectively.

The range of clinically meaningful changes across all 4 raters was 2%–20%. No significance was found among ASL-CBF, DSC-rCBV, and DSC-FTB when added to CE-MR imaging with respect to the number of clinically meaningful changes ( $P = .53$ ).



**FIG 3.** Clinically meaningful changes in BT-RADS scores following the inclusion of perfusion metrics compared with conventional CE-MR imaging alone. Clinically meaningful upgrades or downgrades were defined as score changes from  $\leq 3a \leftrightarrow 3b$  or  $3b \leftrightarrow 3c/4$  and from  $3c/4 \leftrightarrow \leq 3b$  or  $3b \leftrightarrow \leq 3a$ , respectively. The numbers and arrows above the bar graph indicate the number of score upgrades (upward facing arrow) or downgrades (downward facing arrow). The greatest number of score changes was observed with the addition of DSC-FTB.

### Confidence in MR Imaging Interpretations

Confidence scores for all raters, regardless of experience, were higher with the addition of perfusion metrics and were highest with FTB. Figure 4 shows the frequency and mean of scores. For raters 1, 2, and 3, there were no significant differences in confidence scores among CE-MR imaging, CE-MR imaging+ASL-CBF, CE-MR imaging+DSC-rCBV, and CE-MR imaging+DSC-FTB ( $P = .195$ ,  $.052$ , and  $.540$ , respectively). For rater 4, a significant difference in confidence scores was detected among the metrics assessed ( $P = .001$ ). Pair-wise comparisons revealed significant differences between CE-MR imaging and CE-MR imaging+DSC-rCBV and between CE-MR imaging and CE-MR imaging+DSC-FTB (adjusted  $P = .005$ , and  $< .001$ , respectively). No other differences were significant.

### DISCUSSION

Our study demonstrates that the addition of perfusion metrics to conventional CE-MR imaging resulted in higher agreement and confidence in neuroradiologists' clinical interpretation of post-treatment MR imaging of HGGs. Among the perfusion metrics evaluated in this study—ASL-CBF, DSC-rCBV, and DSC-FTB—the highest agreement and confidence were seen with DSC-FTB, a relatively newer described metric in brain tumor imaging. Moreover, our study found that the addition of perfusion metrics to CE-MR imaging yielded clinically meaningful changes in MR imaging interpretation that could affect management in up to 20% of patients compared with the use of CE-MR imaging alone.

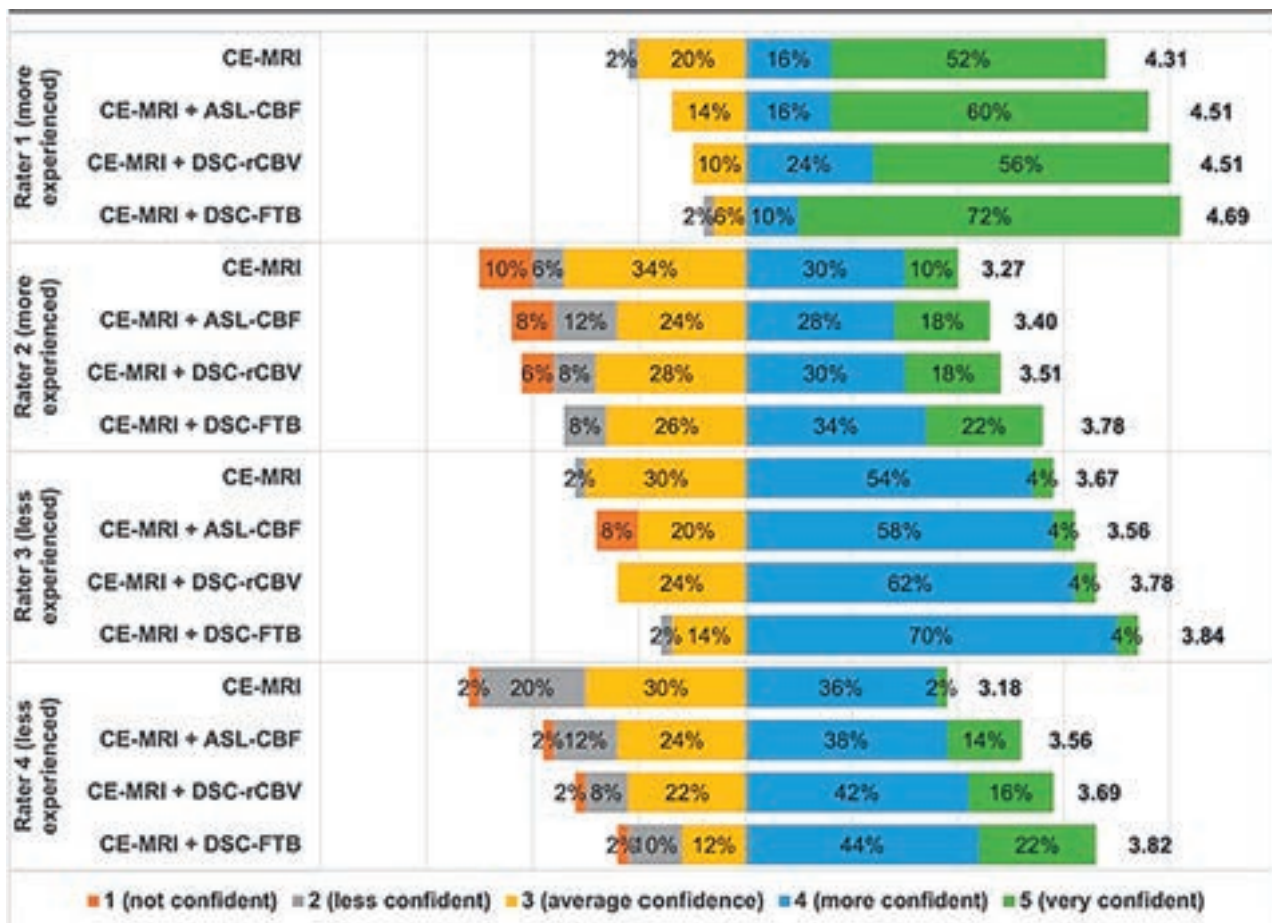
Adding perfusion metrics to conventional CE-MR imaging resulted in higher IRA in BT-RADS scores compared with CE-

MR imaging alone in this study. Prior studies have similarly shown that adding perfusion metrics to CE-MR imaging can increase the IRA. Maiter et al<sup>18</sup> showed fair agreement among raters ( $\kappa = 0.58$ ) when rCBV maps were qualitatively assessed with conventional MR imaging in posttreatment gliomas. Kerkhof et al<sup>19</sup> found good interobserver agreement ( $\kappa = 0.63$ ) on the visual assessment of rCBV maps in GBM as being high or low, though agreement on the overall tumor status was lower ( $\kappa = 0.48$ ). While the IRA improved with the addition of all 3 perfusion metrics, the highest agreement was seen with DSC-FTB, which was not evaluated in previous studies. Even agreement between a multidisciplinary consensus group and raters, despite years of experience, was consistently the highest with DSC-FTB, indicating that this metric may improve the clarity of interpretation. In DSC-FTB, the contrast-enhancing volume is classified into areas of the likelihood of tumor burden based on

predefined rCBV thresholds, which are used to color-code the images for ease of interpretation. This step removes some of the subjectivity inherent in qualitative image interpretation and accounts for the greater agreement seen with DSC-FTB compared with ASL-CBF and DSC-rCBV, particularly in lesions that have an admixture of tumor and treatment effect.

Furthermore, our results show that clinically meaningful changes in interpretation following the use of any perfusion metric, which, in practice, may result in a change in therapeutic management, occurred in up to 20% of patients compared with the use of CE-MR imaging alone. These results are similar to those in previous studies showing the potential of perfusion MR imaging to influence patient management. Geer et al<sup>10</sup> evaluated the impact of DSC and ASL on clinical management in patients with gliomas and found that the addition of perfusion imaging was associated with a change in management plan in up to 8.5% of patients. Maiter et al<sup>18</sup> showed that  $rCBV_{mean}$  ratios in GBM changed the interpretation in 6.3% of reports among readers. Yang et al<sup>20</sup> demonstrated that the addition of DSC PWI-derived  $rCBV_{max}$  (and DWI) improved the diagnostic performance for distinguishing recurrent GBM from nonrecurrence in BT-RADS category 3 lesions. Iv et al<sup>9</sup> showed that FTB can help distinguish recurrent GBM from treatment effects and guide clinical decision-making in posttreatment GBMs. In our study, while no significance was found among the various perfusion metrics with respect to the number of clinically meaningful score changes, the greatest number of changes was seen with DSC-FTB in 3 of the 4 raters. The absence of statistical significance may, in part, be related to our small sample size, but these findings have important implications when taken on an individual patient level.





**FIG 4.** Rater confidence in MR imaging interpretation. Raters graded their confidence in interpretation and assignment of BT-RADS scores for conventional CE-MR imaging, CE-MR imaging + ASL-CBF, CE-MR imaging + DSC-rCBV, and CE-MR imaging + DSC-FTB using a 5-point Likert scale. The number to the right of each color bar represents the mean score. In general, confidence was higher with the addition of any perfusion metric but was highest with DSC-FTB in all raters.

Finally, we found that perfusion MR imaging resulted in higher confidence in interpretation, which has also been demonstrated in other studies. Maiter et al<sup>18</sup> showed an 11.5% increased confidence in the overall opinion of GBM status (progression, pseudoprogression, mixed, or indeterminate) with rCBV<sub>mean</sub> ratios. Geer et al<sup>10</sup> showed that the addition of DSC or ASL to CE-MR imaging increased treatment confidence in the management plan in 57.6% of patient episodes. In this study, the addition of DSC-rCBV and DSC-FTB to CE-MR imaging increased the confidence in interpretation for one of the less experienced raters. Despite no significance found in confidence scores among CE-MR imaging, CE-MR imaging+ASL-CBF, CE-MR imaging+DSC-rCBV, and CE-MR imaging+DSC-FTB for the other 3 raters, scores were generally higher with the addition of perfusion metrics across all raters and were highest with DSC-FTB despite varying experience years.

Although several studies have shown the value of perfusion imaging in the identification of tumor progression and treatment effect, perfusion imaging is still not widely adopted for this patient population at many centers. This study shows the value of adding any perfusion metric, whether derived from ASL or DSC, to conventional CE-MR imaging for routine posttreatment HGG surveillance imaging, not only when there is a concern for tumor

progression or pseudoprogression, and supports its inclusion in a standard glioma MR imaging protocol. Whereas prior studies have primarily looked at ASL-CBF and DSC-rCBV, this study also evaluated the recently described DSC-FTB metric, which incorporates quantitative rCBV thresholds, and shows that it provides better interrater agreement and higher confidence scores in routine interpretation, in large part because FTB maps are typically easier to interpret. Thus, our study demonstrates the clinical translation and integration of quantitative perfusion markers in neuro-oncology imaging and their value in day-to-day clinical practice.

Some important limitations in this study must be considered. One limitation is the lack of histopathologic confirmation to establish a “ground truth” for most cases. In the clinical setting, however, histopathology is not available at each imaging timepoint and, most important, it is not always required for treatment decisions to be made. We, instead, used an experienced consensus group as our reference standard, as used in similar clinical imaging studies<sup>18</sup> and in real-time practice. Another limitation of this study is the small sample size. While DSC-FTB resulted in a greater number of clinically meaningful changes in BT-RADS scores than ASL-CBF and DSC-rCBV, a larger study must be undertaken to determine whether this change holds true. While most patients undergoing surveillance MR imaging for posttreatment HGG can



receive gadolinium-based contrast, some cannot. A future study can evaluate how the addition of ASL-CBF to noncontrast MR imaging influences the interpretation of posttreatment HGG in patients who cannot receive contrast. Finally, given the retrospective nature of the study, there was some variation in MR imaging techniques within the same patient and between patients. However, this reflects real-life clinical practice in which ensuring homogeneity in image acquisition across patients and timepoints is challenging, if not impossible.

## CONCLUSIONS

Adding ASL-CBF, DSC-rCBV, and DSC-FTB to conventional CE-MR imaging improves the IRA in neuroradiologists' clinical interpretations of posttreatment MR imaging of HGGs. Of the perfusion metrics evaluated, DSC-FTB resulted in the highest agreement in BT-RADS scores among all raters and between an experienced multidisciplinary consensus group and individual raters regardless of practice experience years. The addition of perfusion metrics also resulted in higher confidence in interpretation, with the highest scores seen with DSC-FTB, and yielded potential clinically meaningful changes in MR imaging interpretation in up to 20% of patients. Our results highlight the value of adding perfusion imaging to a standard glioma MR imaging protocol for routine interpretation of posttreatment HGGs. Of the perfusion metrics analyzed, DSC-FTB may provide the most benefit regarding multireader agreement and confidence in interpretation, which may be helpful in those with less experience in neuro-oncology imaging.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

- Wen PY, Macdonald DR, Reardon DA, et al. Updated response assessment criteria for high-grade gliomas: Response Assessment in Neuro-Oncology Working Group. *J Clin Oncol* 2010;28:1963–72 CrossRef Medline
- Weinberg BD, Gore A, Shu HG, et al. Management-based structured reporting of posttreatment glioma response with the Brain Tumor Reporting and Data System. *J Am Coll Radiology* 2018;15:767–71 CrossRef Medline
- Barajas RF Jr, Chang JS, Segal MR, et al. Differentiation of recurrent glioblastoma multiforme from radiation necrosis after external beam radiation therapy with dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging. *Radiology* 2009;253:486–96 CrossRef Medline
- Hu LS, Baxter LC, Smith KA, et al. Relative cerebral blood volume values to differentiate high-grade glioma recurrence from post-treatment radiation effect: direct correlation between image-guided tissue histopathology and localized dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging measurements. *AJNR Am J Neuroradiol* 2009;30:552–58 CrossRef Medline
- Prager AJ, Martinez N, Beal K, et al. Diffusion and perfusion MRI to differentiate treatment-related changes including pseudoprogression from recurrent tumors in high-grade gliomas with histopathologic evidence. *AJNR Am J Neuroradiol* 2015;36:877–85 CrossRef Medline
- Ozsunar Y, Mullins ME, Kwong K, et al. Glioma recurrence versus radiation necrosis? A pilot comparison of arterial spin-labeled, dynamic susceptibility contrast enhanced MRI, and FDG-PET imaging. *Acad Radiol* 2010;17:282–90 CrossRef Medline
- Liu Y, Chen G, Tang H, et al. Systematic review and meta-analysis of arterial spin-labeling imaging to distinguish between glioma recurrence and post-treatment radiation effect. *Ann Palliat Med* 2021;10:12488–97 CrossRef Medline
- Nguyen TB, Zakhari N, Velasco Sandoval S, et al. Diagnostic accuracy of arterial spin-labeling, dynamic contrast-enhanced, and DSC perfusion imaging in the diagnosis of recurrent high-grade gliomas: a prospective study. *AJNR Am J Neuroradiol* 2023;44:134–42 CrossRef Medline
- Iv M, Liu X, Lavezo J, et al. Perfusion MRI-based fractional tumor burden differentiates between tumor and treatment effect in recurrent glioblastomas and informs clinical decision-making. *AJNR Am J Neuroradiol* 2019;40:1649–57 CrossRef Medline
- Geer CP, Simonds J, Anverly A, et al. Does MR perfusion imaging impact management decisions for patients with brain tumors? A prospective study. *AJNR Am J Neuroradiol* 2012;33:556–62 CrossRef Medline
- Hu LS, Baxter LC, Pinnaduwa DS, et al. Optimized preload leakage-correction methods to improve the diagnostic accuracy of dynamic susceptibility-weighted contrast-enhanced perfusion MR imaging in posttreatment gliomas. *AJNR Am J Neuroradiol* 2010;31:40–48 CrossRef Medline
- Schminda KM, Prah MA, Rand SD, et al. Multisite concordance of DSC-MRI analysis for brain tumors: results of a National Cancer Institute Quantitative Imaging Network collaborative project. *AJNR Am J Neuroradiol* 2018;39:1008–16 CrossRef Medline
- Schminda KM, Prah MA, Hu LS, et al. Moving toward a consensus DSC-MRI protocol: validation of a low-flip angle single-dose option as a reference standard for brain tumors. *AJNR Am J Neuroradiol* 2019;40:626–33 CrossRef Medline
- Prah MA, Al-Gizawi MM, Mueller WM, et al. Spatial discrimination of glioblastoma and treatment effect with histologically-validated perfusion and diffusion magnetic resonance imaging metrics. *J Neurooncol* 2018;136:13–21 CrossRef Medline
- Hu LS, Kelm Z, Korfiatis P, et al. Impact of software modeling on the accuracy of perfusion MRI in glioma. *AJNR Am J Neuroradiol* 2015;36:2242–49 CrossRef Medline
- Hu LS, Eschbacher JM, Heiserman JE, et al. Reevaluating the imaging definition of tumor progression: perfusion MRI quantifies recurrent glioblastoma tumor fraction, pseudoprogression, and radiation necrosis to predict survival. *Neuro Oncol* 2012;14:919–30 CrossRef Medline
- Connelly JM, Prah MA, Santos-Pinheiro F, et al. Magnetic resonance imaging mapping of brain tumor burden: clinical implications for neurosurgical management: case report. *Neurosurg Open* 2021;2:okab029 CrossRef Medline
- Maiter A, Butteriss D, English P, et al. Assessing the diagnostic accuracy and interobserver agreement of MRI perfusion in differentiating disease progression and pseudoprogression following treatment for glioblastoma in a tertiary UK centre. *Clin Radiol* 2022;77:e568–75 CrossRef Medline
- Kerkhof M, Hagenbeek RE, van der Kallen BF, et al. Interobserver variability in the radiological assessment of magnetic resonance imaging (MRI) including perfusion MRI in glioblastoma multiforme. *Eur J Neurol* 2016;23:1528–33 CrossRef Medline
- Yang Y, Yang Y, Wu X, et al. Adding DSC PWI and DWI to BT-RADS can help identify postoperative recurrence in patients with high-grade gliomas. *J Neurooncol* 2020;146:363–71 CrossRef Medline

# Improved Detection of Target Metabolites in Brain Tumors with Intermediate TE, High SNR, and High Bandwidth Spin-Echo Sequence at 5T

Wenbo Sun, Dan Xu, YanXing Yang, Linfei Wen, Hanjiang Yu, Yaowen Xing, Xiaopeng Song, Huan Li, and Haibo Xu

## ABSTRACT

**BACKGROUND AND PURPOSE:** Due to high chemical shift displacement, challenges emerge at ultra-high fields when measuring metabolites using  $^1\text{H}$ -MRS. Our goal was to investigate how well the high SNR and high bandwidth spin-echo (HISE) technique perform at 5T for detecting target metabolites in brain tumors.

**MATERIALS AND METHODS:** Twenty-six subjects suspected of having brain tumors were enrolled. HISE and point-resolved spectroscopy (PRESS) single-voxel spectroscopy scans were collected with a 5T clinical scanner with an intermediate TE (TE = 144 ms). The main metabolites, including total NAA, Cr, and total Cho, were accessed and compared between HISE and PRESS using a paired Student *t* test, with full width at half maximum and SNR as covariates. The detection rate of specific metabolites, including lactate, alanine, and lipid, and subjective spectral quality were accessed and compared between HISE and PRESS.

**RESULTS:** Twenty-three pathologically confirmed brain tumors were included. Only the full width at half maximum for total NAA was significantly lower with HISE than with PRESS ( $P < .05$ ). HISE showed a significantly higher SNR for total NAA, Cr, and total Cho compared with PRESS ( $P < .05$ ). Lactate was detected in 21 of the 23 cases using HISE, but in only 4 cases using PRESS. HISE detected alanine in 8 of 9 meningiomas, whereas PRESS detected alanine in just 3 meningiomas. PRESS found lipid in more cases than HISE, while HISE outperformed PRESS in terms of subjective spectral quality.

**CONCLUSIONS:** HISE outperformed the clinical standard PRESS technique in detecting target metabolites of brain tumors at 5T, particularly lactate and alanine.

**ABBREVIATIONS:** Ala = alanine; CSDE = chemical shift displacement error; FWHM = full width at half maximum; 2-HG = 2-hydroxyglutarate; HISE = high bandwidth spin-echo; Lac = lactate; Lip = lipid; PRESS = point-resolved spectroscopy; RF = radiofrequency; STEAM = stimulated echo acquisition mode; SVS = single-voxel spectroscopy; *t* = total

Brain tumors encompass >120 different types, with some common primary tumors being highly malignant.<sup>1</sup> The criterion standard for clinical diagnosis of brain tumors involves pathologic testing conducted after invasive surgical resection or needle biopsy.<sup>2</sup> Nevertheless, the development of noninvasive diagnostic tools is necessary.

MR imaging is a crucial noninvasive imaging tool for brain tumors.<sup>3</sup>  $^1\text{H}$ -MRS is an MR imaging technique that enables

quantitation of different metabolite profiles in vivo.<sup>4</sup> The  $^1\text{H}$ -MRS spectrum shows 3 primary metabolite peaks for brain tumors, which have been linked to intact glioneuronal structures (NAA), energy homeostasis (Cr), and tumor membrane turnover and proliferation (Cho), respectively.<sup>5</sup> Meanwhile, depending on the metabolite patterns of tumors, lactate (Lac), alanine (Ala), and mobile lipid (Lip) may also be observed. A metabolic shift toward glycolysis has been found in tumor cells, which is known as the Warburg effect.<sup>6</sup> Lac is an end product of glycolysis, while Ala is a reduced partner of pyruvate formed from glycolysis.<sup>7</sup> Both Lac and Ala have been associated with a range of brain tumors.<sup>7</sup> Lip signal is linked to necrosis and apoptosis.<sup>7</sup> However, the routine application of  $^1\text{H}$ -MRS in detecting the above metabolites is limited by the clinical field strength, often 1.5T, which provides low SNR and spectra resolution.<sup>8</sup> Typically, it is difficult to resolve Lac and Ala peaks individually at 1.5T.<sup>8</sup>

Higher magnetic field strengths provide an opportunity for improving  $^1\text{H}$ -MRS data quality.<sup>9</sup> Previous studies have demonstrated the advantages of using 3T over 1.5T in diagnosing brain

Received August 10, 2023; accepted after revision November 6.

From the Departments of Radiology (W.S., H.L., H.X.) and Nuclear Medicine (D.X.), Zhongnan Hospital of Wuhan University, Wuhan, Hubei, P.R. China; and United-Imaging Healthcare (Y.Y., L.W., H.Y., Y.X., X.S.), Shanghai, China.

Wenbo Sun and Dan Xu share first authorship.

This work was supported by the National NSFC International (regional) Cooperation and Exchange Project x 82111530204; the Key R&D Program of Hubei Province x 2020BCB030; and the Improvement Project for Theranostic Ability on Difficulty Miscellaneous Disease (tumor) of the Zhongnan Hospital of Wuhan University.

Please address correspondence to Haibo Xu, MD, PhD, Department of Radiology, Zhongnan Hospital of Wuhan University, 169 East Lake Rd, Wuhan 430071, Hubei, P. R. China; e-mail: xuhaibo1120@hotmail.com

<http://dx.doi.org/10.3174/ajnr.A8150>

tumors with  $^1\text{H}$ -MRS, in terms of higher SNR, lower spectra overlap, and shorter acquisition times.<sup>10,11</sup> Moreover, the application of ultra-high-field (7T) MR imaging systems further enhances SNR and provides superior spatial and spectral resolution compared with 3T.<sup>12,13</sup> Nearby metabolites, such as 2-hydroxyglutarate (2HG), glutamate/glutamine, and  $\gamma$ -aminobutyric acid, could be separated and quantified at 7T.<sup>13</sup> Despite these advantages, higher magnetic field strengths pose some challenges related to field inhomogeneities and a larger chemical shift displacement error (CSDE); therefore, it is necessary to optimize the acquisition scheme of  $^1\text{H}$ -MRS when working with higher magnetic field strengths.<sup>14,15</sup>

Currently, in clinical practice, the standard acquisition sequences include stimulated echo acquisition mode (STEAM) and point-resolved spectroscopy (PRESS).<sup>14</sup> In STEAM, 3 section-selective radiofrequency (RF) pulses, each with a 90° flip angle, generate a stimulated echo, while the PRESS technique uses a 90° excitation pulse combined with 2 refocusing pulses to produce a spin-echo.<sup>16</sup> PRESS is preferred over STEAM due to its ability to double the signal output.<sup>14</sup> However, both STEAM and PRESS are susceptible to the increased CSDE at higher field strengths.<sup>14</sup> With increased CSDE, PRESS and STEAM would both have anomalous J-modulation, and additional signal cancellation would occur in metabolites with weakly coupled resonances, like Lac.<sup>14</sup> The high SNR and high bandwidth spin-echo (HISE) technique, also called semi-adiabatic localization by adiabatic selective refocusing (sLASER) by some MR imaging manufacturers, has been recently applied to overcome this challenge.<sup>17</sup> In HISE, the 2 refocusing pulses are replaced by 2 pairs of adiabatic full-passage pulses, allowing a higher RF pulse bandwidth, less sensitivity to field inhomogeneities, and superior section-selection profiles, decreasing the CSDE to a minimum.<sup>17</sup> Therefore, HISE is recommended for high and ultra-high-field applications.<sup>14</sup> Several investigations at 7T have already applied similar HISE approaches and proved the benefits of HISE in detecting unique compounds and resolving adjacent metabolites.<sup>13,17-20</sup> Nevertheless, clinical use of 7T MR imaging systems is not yet common, even though the FDA has approved the technology.<sup>21</sup> As a result, the clinical potential of the combination of ultra-highfield and the HISE technique has not been fully explored.

Recently, a 5T clinical scanner has also shown its capabilities in neuroimaging. In displaying the intracranial distal small branches with TOF MRA, it surpasses 3T and is comparable with 7T.<sup>22</sup> Additionally, 5T enables better visualization of intracranial vessels than 3T for vessel wall imaging and SWI.<sup>23</sup> For quantitative T1 $\rho$  mapping of brain tissues, 5T showed more benefits in the SNR than 3T at a high spatial resolution.<sup>24</sup> It has also been demonstrated that 5T had a lower RF inhomogeneity and specific absorption rate than 7T<sup>23</sup> and was capable of scanning abdominal organs, including the pancreas,<sup>25,26</sup> kidney,<sup>27</sup> spleen,<sup>28</sup> and liver,<sup>28</sup> whereas 7T has not been approved for such applications, due to the RF inhomogeneity in large-body cross-sections and safety considerations concerning RF power deposition.<sup>29</sup> Imaging at 5T may offer a good balance between the benefits and drawbacks of ultra-high-field MR imaging and thus may see wider applicable clinical scenarios than 7T in the future.

So far, the performance of  $^1\text{H}$ -MRS in diagnosing brain tumors has never been investigated at 5T. Because 5T belongs to

**Table 1: Acquisition parameters for the HISE and PRESS SVS sequences**

Parameters	HISE	PRESS
TR	2500 ms	2500 ms
TE	144 ms	144 ms
Voxel size	15 × 15 × 15 mm <sup>3</sup>	15 × 15 × 15 mm <sup>3</sup>
Bandwidth	1000 kHz	1000 kHz
Averages	100	100
Spectral sampling	1024	1024
Flip angle	90°	90°
Phase-cycling schemes	2 blocks	4 blocks
Water-suppressed bandwidth	90 kHz	90 kHz
Acquisition time	4 min 19 sec	4 min 14 sec

**Note:**—min indicates minute; sec, second.

the ultra-high field, we hypothesize that HISE will show obvious advantages at 5T. The objective of this study was to validate the application of the HISE technique at 5T in the clinic and compare its performance in detecting target metabolites of brain tumors with that of the clinical standard PRESS technique.

## MATERIALS AND METHODS

### Subjects

This prospective study was approved by the ethics committee of our hospital (ethics number 2021110). From March 2023 to July 2023, a total of 26 patients suspected of having brain tumors were enrolled in this study. Informed consent was obtained from all participants. The inclusion criteria were as follows: between 18 and 80 years of age, suspected of having brain tumors, no contraindications to MR imaging (including metal implants and claustrophobia), and no history of brain surgery, radiation therapy, or chemotherapy. The exclusion criteria comprised patients who did not undergo brain surgery after MR imaging ( $n = 1$ ), patients who were pathologically confirmed not to have a brain tumor ( $n = 1$ ), and incomplete image acquisition ( $n = 1$ ).

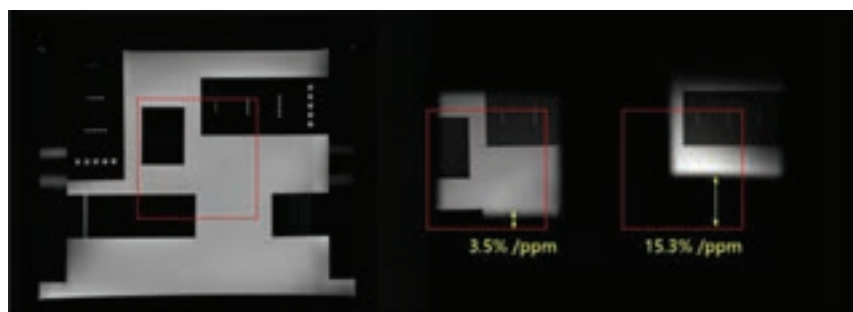
### Image Acquisition and Analysis

All MR images were obtained using a 5T whole-body MR scanner (uMR Jupiter; United Imaging Healthcare) with a 2-channel transmit and 48-channel receive (2Tx/48Rx; United Imaging Healthcare) head coil. For both HISE and PRESS, the CSDE data were first acquired in a phantom containing 100% water. For phantom and in vivo measurements, HISE and PRESS were scanned on the basis of the parameters specified in Table 1. Before the single-voxel spectroscopy (SVS) scans, T2-weighted FSE sequences were acquired in 3 anatomic directions (sagittal, coronal, and axial) to cover the entire brain and aid in the localization of tumor lesions. The acquisition parameters for the T2-weighted FSE sequences were as follows: FOV = 230 × 200 mm<sup>2</sup>, TR = 4000 ms, TE = 94 ms, section thickness = 5 mm, section number = 19, acceleration = 2.0, gap = 30%, flip angle = 90°, and resolution in-plane = 0.65 × 0.65 mm<sup>2</sup>. Subsequently, HISE and PRESS SVS scans were obtained in a random order with an intermediate TE of 144 ms. An intermediate TE was used for the detection of Lac signals at 7T.<sup>17</sup> Before scanning, a B<sub>0</sub> shimming was performed. During scans, a dynamic frequency calibration module to correct the frequency shift during measurement and a motion-monitoring module were added.

Prominent metabolites, including total NAA (tNAA, composed of NAA and N-Acetyl-Aspartyl-Glutamate (NAAG)), Cr, and total Cho (tCho, composed of glycerophosphocholine and phosphorylcholine) were measured at chemical shift positions of 2.03, 3.03, and 3.21 ppm, respectively. The full width at half maximum (FWHM) and SNR of tNAA, Cr, and tCho were calculated using custom-designed software. The FWHM was used to evaluate the quality of shimming and was measured from the spectral width at the half amplitude of the metabolite signal. The SNR was calculated using the highest baseline subtracted metabolite signal intensity divided by the SD of the noise on the spectral baseline estimated from a region free from metabolite signals. To detect specific metabolites, we used an in-house MATLAB script (MathWorks) for spectra editing. Lac, Ala, and Lip signals were detected at 1.34, 1.45, and 1.3 ppm, respectively. A comparative analysis of the detection rate in Lac, Ala, and Lip signals between HISE and PRESS was performed. Moreover, the overall spectral quality of HISE and PRESS for each subject was rated by 2 radiologists with 5 and 10 years of experience, using a 5-point Likert scale. The Likert scale was used with the following grades: 1 = not diagnostic; 2 = spectra are markedly distorted with poor diagnostic value; 3 = spectra are minimally distorted with reduced diagnostic value; 4 = no distortion with good diagnostic value; and 5 = spectra with excellent diagnostic quality by presenting more peaks of metabolites. The 2 radiologists reached a consensus over the scores.

### Pathologic Analysis

After surgery, tumor specimens were fixed in 4% paraformaldehyde, embedded in paraffin, and sectioned into 4- $\mu$ m-thick histologic sections for pathologic diagnosis. Both routine H&E staining and immunohistochemistry analysis were performed.



**FIG 1.** A comparison of CSDE with HISE and PRESS in a phantom containing 100% water. The red square indicates the FOV in the center of the phantom, while the middle image is the result of HISE, and the right image is the result of PRESS.

Additionally, DNA sequencing was performed to confirm the presence of gene mutations for gliomas, including the *IDH* mutation and 1p/19q codeletion. The diagnosis was made by a pathologist with 10 years of experience following the guidelines outlined in the 2016 World Health Organization Classification of Tumors of the Central Nervous System.

### Statistical Analysis

To compare the FWHM and SNR of tNAA, Cr, and tCho between 2 scans, we used a paired Student *t* test. The detection rate of Lac, Ala, and Lip signals among all cases was calculated for both sequences. For the evaluation of spectral quality, the Likert scale was compared between HISE and PRESS with frequency tables and a paired Student *t* test. All statistical analyses were performed using SPSS Version 24.0 software (IBM). A *P* value < .05 suggested a statistically significant difference.

### RESULTS

A total of 23 patients (mean age, 55.39 [SD, 11.17] years, male, *n* = 8) were finally included in this study, with tumors consisting of acoustic neuroma (*n* = 2), non-Hodgkin lymphoma (*n* = 1), meningiomas (*n* = 9), gliomas (*n* = 9), and metastasis (*n* = 2).

#### CSDE for HISE and PRESS with 5T

Due to the high bandwidth property of the adiabatic pulse, the phantom experiment showed a obviously reduced CSDE with 5T (3.5%/ppm for the adiabatic refocusing pulse) compared with a normal PRESS sequence (15.3%/ppm), as shown in Fig 1.

#### FWHM and SNR of Prominent Target Metabolites

As shown in Table 2, the FWHM for tNAA was significantly lower with HISE than with PRESS (*P* < .05), while the SNR for tNAA, Cr, and tCho was significantly higher with HISE compared with PRESS (*P* < .05).

#### Detection Rate of Specific Target Metabolites

As shown in Table 3, among all brain tumors, HISE detected Lac in 21 cases, whereas PRESS detected Lac in just in 4 cases. HISE detected Ala in 8 of 9 meningiomas, while PRESS detected Ala in only 3 meningiomas. PRESS detected Lip signals in 14 cases, while HISE detected Lip signals in 11 cases.

**Table 2: A comparison of average values (mean [SD]) of FWHM and SNR of prominent metabolites for the HISE and PRESS SVS scans among all cases<sup>a</sup>**

Metabolite	FWHM			SNR		
	HISE	PRESS	<i>P</i> Value	HISE	PRESS	<i>P</i> Value
tNAA	10.92 (SD, 2.59)	14.12 (SD, 3.62)	.001 <sup>b</sup>	14.61 (SD, 7.66)	9.53 (SD, 5.48)	.001 <sup>b</sup>
Cr	10.87 (SD, 3.31)	11.75 (SD, 4.72)	.455	20.02 (SD, 14.06)	9.44 (SD, 6.12)	.000 <sup>c</sup>
tCho	12.07 (SD, 2.74)	12.74 (SD, 3.45)	.215	63.23 (SD, 43.32)	29.72 (SD, 19.19)	.000 <sup>c</sup>

<sup>a</sup> The *P* value shown is for the paired *t* test performed between HISE and PRESS.

<sup>b</sup> *P* < .01.

<sup>c</sup> *P* < .001.



**Table 3: The detection of specific metabolites (Lac, Ala, and Lip) at TE = 144 ms for the HISE and PRESS SVS scans, respectively<sup>a</sup>**

No.	Sex	Age (yr)	Pathologic Diagnosis	HISE			PRESS		
				Lac	Ala	Lip	Lac	Ala	Lip
Case 1	F	71	Acoustic neuroma	+	—	+	—	—	+
Case 2	M	47	Acoustic neuroma	+	—	—	—	—	—
Case 3	M	52	Metastases (lung origin)	+	?	+	—	—	+
Case 4	M	58	Metastases (gastrointestinal origin)	+	—	+	—	—	+
Case 5	M	61	Non-Hodgkin lymphoma	+	—	+	—	—	+
Case 6	M	66	Meningioma, WHO I	+	+	—	—	—	—
Case 7	F	52	Meningioma, WHO I	+	+	—	—	—	—
Case 8	F	67	Meningioma, WHO I	+	?	—	+	—	—
Case 9	F	43	Meningioma, WHO II	+	+	—	—	+	—
Case 10	M	52	Meningioma, WHO I	+	?	+	—	—	+
Case 11	F	34	Meningioma, WHO I	+	+	—	—	?	—
Case 12	F	56	Meningioma, WHO I	—	—	+	—	—	+
Case 13	F	50	Meningioma, WHO I	+	+	—	—	—	—
Case 14	F	58	Meningioma, WHO I	—	+	+	—	?	+
Case 15	F	42	Oligodendroglioma, WHO III	+	—	—	—	—	—
Case 16	F	49	Oligodendroglioma, WHO II	+	—	—	—	—	—
Case 17	F	35	GBM, WHO IV	+	—	—	—	—	+
Case 18	M	45	GBM, WHO IV	+	?	+	—	—	+
Case 19	F	64	GBM, WHO IV	+	?	+	+	—	+
Case 20	F	67	GBM, WHO IV	+	—	—	—	—	+
Case 21	F	70	GBM, WHO IV	+	?	+	—	—	+
Case 22	M	66	GBM, WHO IV	+	—	—	+	—	+
Case 23	F	69	GBM, WHO IV	+	—	+	+	—	+

**Note:**—F indicates female; M, male; WHO, World Health Organization; GBM, glioblastoma.

<sup>a</sup>The plus sign represents the detection of strong signal peaks; the question mark represents the presence of faint or inconspicuous signal peaks; and the minus sign represents the almost undetectable signal peaks.

**Table 4: A comparison of results of spectral quality rating between HISE and PRESS SVS**

Rate	HISE SVS	PRESS-SVS	P Value
1	0	0	NA
2	0	2	NA
3	2	8	NA
4	5	8	NA
5	16	5	NA
Mean value	4.61 (SD, 0.66)	3.70 (SD, 0.93)	.001 <sup>a</sup>

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

<sup>a</sup> $P < .01$ .

### Subjective Spectral Quality Evaluations

The HISE showed significantly higher scores in spectral quality than PRESS ( $P < .001$ ), as demonstrated in Table 4. Representative spectra for typical patients (meningioma, glioma, metastases, and acoustic neuroma) are demonstrated in Figs 2–5, respectively.

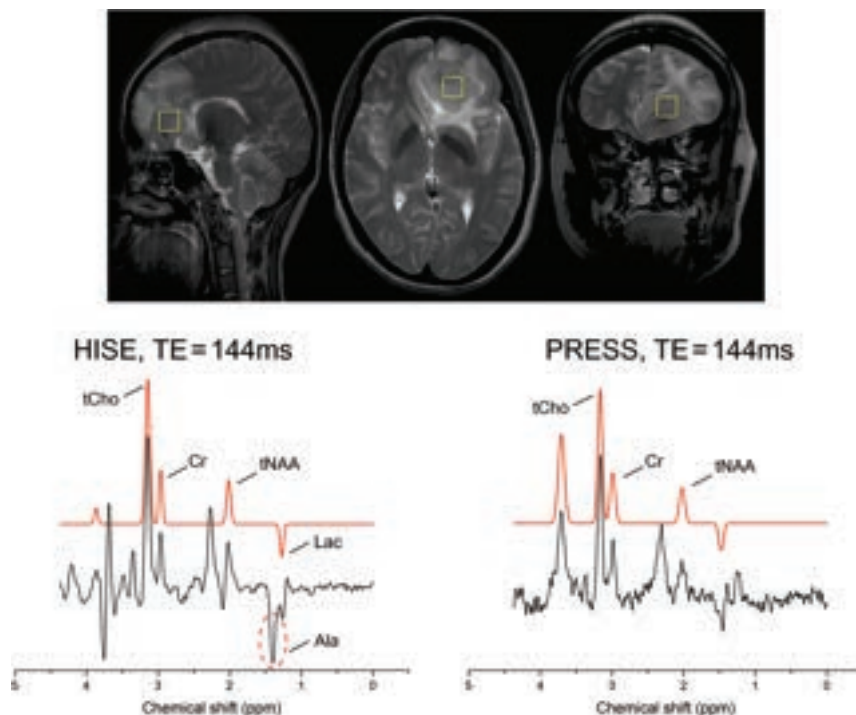
## DISCUSSION

This study aimed to investigate the performance of HISE in diagnosing brain tumors with a 5T ultra-high-field whole-body clinical scanner by comparing HISE with the clinical standard PRESS. We found that HISE demonstrated significantly higher SNR in tNAA, Cr and tCho compared with PRESS. Furthermore, we observed a higher detection rate of Lac signal in all tumors, as well as a higher detection rate of Ala in meningiomas with HISE. In addition, PRESS showed a higher detection rate for Lip signals. According to the 5-point Likert scale, a significant, higher spectral quality with HISE than with PRESS was found. Altogether, these findings suggested that the application of HISE holds promise in clinical

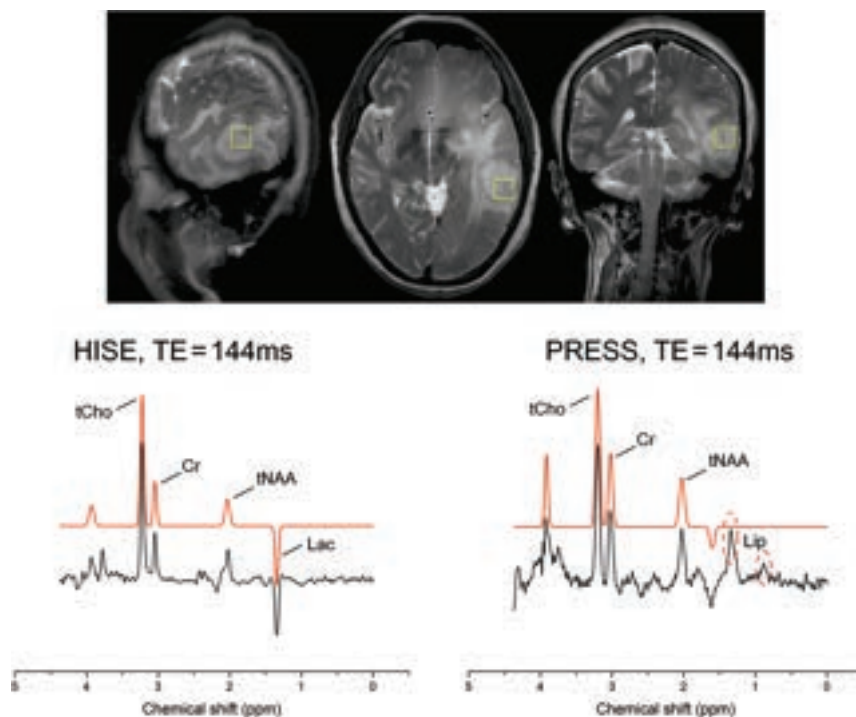
settings for diagnosing brain tumors with a recently developed 5T clinical scanner.

Several studies have already compared HISE or similar techniques with PRESS at 3T.<sup>30,31</sup> A prior study compared HISE with PRESS and discovered that HISE has a lower CSDE than PRESS (6% versus 24%) and a stronger signal of mIns and Cr in its spectra than PRESS.<sup>30</sup> Another study compared similar techniques, Mescher-Garwood (MEGA) semi-localised by adiabatic selective refocusing (MEGA-sLASER) and MEGA-PRESS at Siemens 3T system, and found obvious SNR increases (mean, 10.9 [SD, 5.2] versus 5.0 [SD, 3.0]) using MEGA-sLASER for Lac detection in distinct brain areas.<sup>31</sup> Meanwhile, only a few studies have looked into PRESS at 7T, which could be due to the drawback of PRESS at high fields, a very high CSDE due to a constricted refocusing pulse bandwidth.<sup>32</sup> An earlier investigation in 12 patients with gliomas found that a modified PRESS sequence could detect 2HG and separate it from glutamate/glutamine and  $\gamma$ -aminobutyric acid signals.<sup>33</sup> However, it showed a large CSDE (20%) of PRESS at 7T.<sup>33</sup>

In comparison with the scant literature on PRESS at 7T, many studies have focused on the use of HISE in neuroimaging at 7T.<sup>13,18–20,34</sup> A study on 5 healthy volunteers with HISE found an average SNR (48 [SD, 6]) for NAA and reliable readings for Cho, NAA, Cr, mIns, and glutamate/glutamine.<sup>18</sup> Another study used HISE in 7 gliomas and found an average SNR (24.4 [SD, 13.6]) for the obtained spectra in tumor locations, and 8 metabolites, including lactate, NAA, Cho, and Cr, could be identified.<sup>20</sup> HISE was also shown to be capable of detecting 2HG with an average SNR (77 [SD, 26]) in a study involving 9 patients with gliomas<sup>13</sup> and detected 2HG concentration as low as 0.5 mM in a study involving 4 *IDH1* mutant glioms.<sup>34</sup> In addition, a study on healthy volunteers using an HISE sequence at 7T revealed high



**FIG 2.** A 43-year-old woman with a mass located in the left frontal lobe, suggesting a meningioma. It was an atypical meningioma (World Health Organization grade II) with invasion into the brain parenchyma. The HISE technique observed relatively strong Ala signal, while the PRESS technique could not differentiate between the Ala and Lac peaks. Both HISE and PRESS did not detect Lip signals. The yellow box in anatomical images represented the region of the volume of the SVS scan.

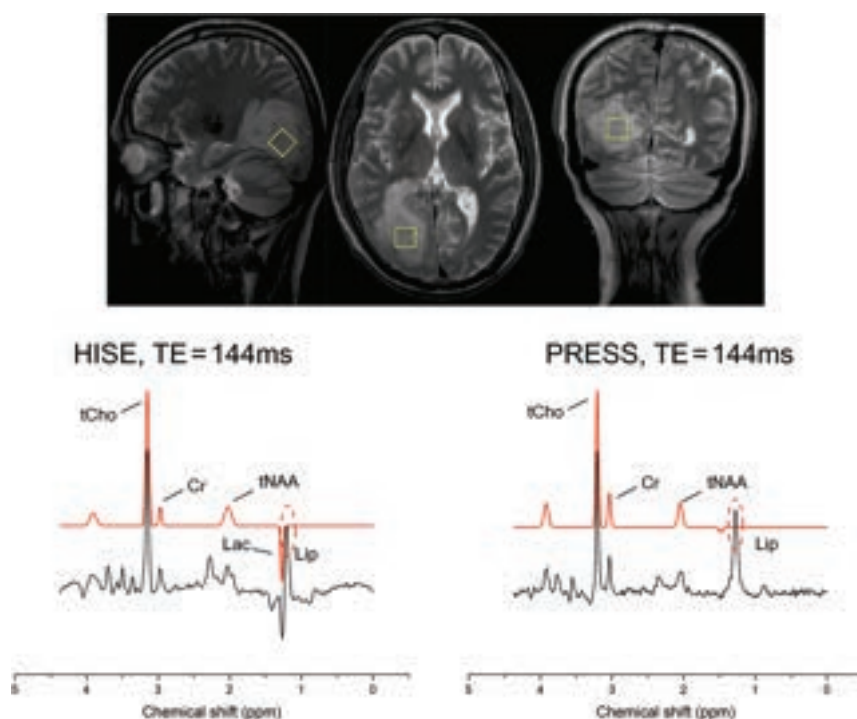


**FIG 3.** A 67-year-old woman with a mass in the left temporal lobe with surrounding edema. On the basis of multimodal MR imaging, a high-grade glioma was suspected. Clinical correlation was recommended. The glioma was classified as a glioblastoma (World Health Organization grade IV, *IDH1* wild-type). In the HISE, a Lac signal was detected. In PRESS, due to chemical shift displacement effects, a Lip signal originating from the scalp was detected, and the Lac signal was covered by the Lip signal. The yellow box in anatomical images represented the region of the volume of the SVS scan.

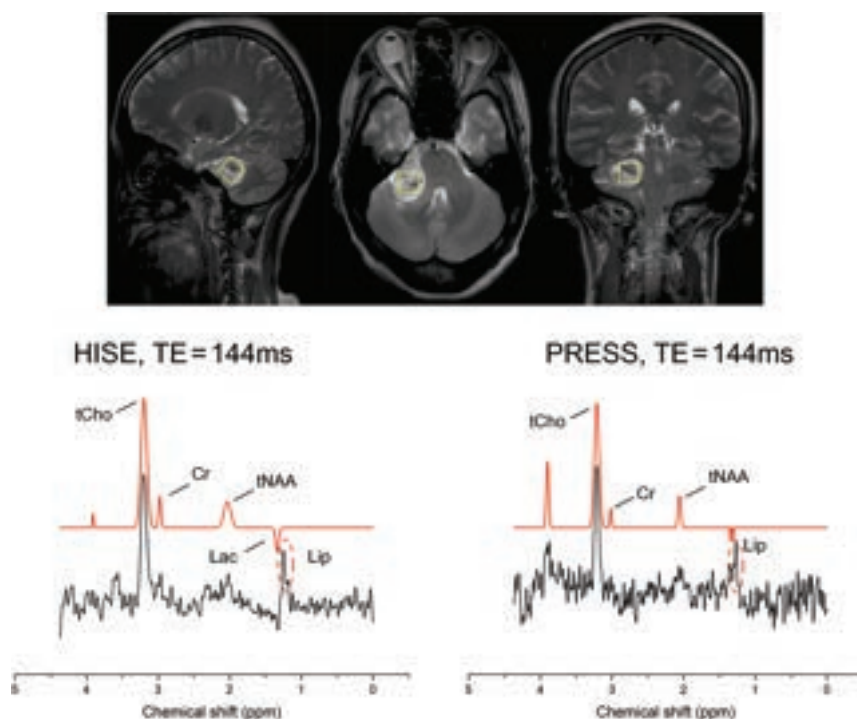
SNR (mean, 119 [SD, 39]) and well-fitting results for brain glutamate.<sup>19</sup> These findings at 7T support the experts' consensus that ultra-high fields generally favor HISE over PRESS; however, these results have never been repeated at 5T, and there is little evidence for the routine use of HISE for brain tumors at 5T. Hence, a comparison study of HISE and PRESS at 5T is still essential.

At ultra-high field, inhomogeneities ( $B_0$  and  $B_1$ ) are prevalent and may have an influence on the quality of the  $^1\text{H}$ -MRS spectrum. FWHM is an important indicator of the  $B_0$  shimming.<sup>35</sup> In our study, we found a similar or better  $B_0$  shimming with HISE than PRESS at 5T, despite the lesions being located near the base of the skull or scalp, which could be influenced by susceptibility artifacts. Moreover, we found significantly higher SNR in tNAA, Cr, and tCho when using HISE compared with PRESS at 5T. These results indicate that at 5T, HISE may still show advantages over PRESS in terms of SNR. It might be due to the following reasons:<sup>32</sup> When the magnetic field strength increases, a larger RF field is needed to flip the spin. Due to the limitation of the RF field amplitude of the system and to minimize CSDE, the return flip angle of the PRESS sequence is usually  $<180^\circ$ . However, by using an adiabatic refocusing pulse in the HISE sequence, the return flip angle can reach  $180^\circ$  under the same condition, resulting in a lower sensitivity to field inhomogeneities and a higher SNR.

Besides SNR, we observed a higher detection rate of Lac and Ala using HISE compared with PRESS at 5T. Lac accumulation serves as an important marker of bulk tumors.<sup>36</sup> According to the hypothesis of the Warburg effect, increased Lac levels may be caused by increased glycolysis and hypoxia in tumors.<sup>6</sup> However, the limited detection rate of Lac in PRESS can be because Lac is a metabolite with frequency-separated J-coupled multiplets.<sup>14</sup> As a result, it may experience signal loss in PRESS due to the unequal exposure of regions around the voxel edge periphery to two  $180^\circ$  refocusing pulses.<sup>14</sup> HISE increased the bandwidth of the RF pulses, thus reducing the signal cancellation in Lac.<sup>17</sup>



**FIG 4.** A 52-year-old man with a right occipital lobe space-occupying lesion, which was classified as a brain metastasis (originating from lung cancer). In the HISE, both Lac and Lip signals were detected. In PRESS, only a Lip signal was detected, and the Lac signal was covered by the Lip signal. The *yellow box* in anatomical images represented the region of the volume of the SVS scan.



**FIG 5.** A 71-year-old woman with a space-occupying lesion in the right cerebellopontine angle, confirmed to be an acoustic neuroma after surgery. In the HISE, a small Lac signal was detected. In PRESS, the Lac signal was also covered by the Lip signal. The *yellow box* in anatomical images represented the region of the volume of the SVS scan.

Additionally, we detected more Ala cases with HISE in gliomas and meningiomas. Ala could enter the metabolic stream to provide energy and precursors for rapidly proliferating tumor cells in gliomas.<sup>37</sup> Furthermore, because Ala is a hallmark of meningioma,<sup>8</sup> the high detection rate of Ala in meningiomas could be attributed to the enhanced spectral resolution and less CSDE of HISE at 5T, as shown in Fig 2. Last, PRESS showed more cases with Lip signals. This could be owing to the difficulty of PRESS in completely suppressing scalp Lip signals, as demonstrated in Fig 3. However, compared with PRESS, HISE has superior section-selection characteristics, thus reducing Lip contaminations from outside the measurement volume.<sup>30</sup>

Some limitations of this study should be acknowledged. First, the size of our research cohort was relatively small. A larger sample size would provide more statistical power and enhance the generalizability of the findings. Second, we did not conduct reliable and reproducible comparisons of these 2 sequences between 3T and 5T. Although it is common sense that a higher field strength would have a higher CSDE, further investigations comparing the 2 sequences between 3T and 5T are warranted. Third, we chose only an intermediate TE (144 ms), which is recommended at 7T for the detection of negative in-phase Lac.<sup>17</sup> A shorter TE could be achieved by PRESS than by HISE, which could reduce the signal loss due to T2 relaxation. A further comparison study using shorter TEs toward other metabolites, 2HG, which is crucial for the diagnosis of isocitrate dehydrogenase mutant gliomas, is planned at 5T. Last, we evaluated only the SVS scans. Although SVS has been proved to be robust in clinical settings, particularly for most lesions located at anatomic regions with B<sub>0</sub> field inhomogeneities in our cohort, it would be beneficial to investigate the HISE technique based on a 2D or 3D chemical shift imaging approach. This process would provide a more comprehensive understanding of the performance of the HISE technique at 5T.



## CONCLUSIONS

In a recently developed whole-body 5T clinical scanner, the HISE technique is preferable to PRESS for the clinical diagnosis of brain tumors. This result is attributed to its higher SNR and detection rate of target metabolites in brain tumors, especially Lac and Ala.

## ACKNOWLEDGMENTS

We would like to express our gratitude to Prof. Qu Xiaobo and Dr. Tu Zhangren from Xiamen University for their assistance in testing the data for 2HG detection using the LCModel.









Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

1. Mack SC, Hubert CG, Miller TE, et al. **An epigenetic gateway to brain tumor cell identity.** *Nat Neurosci* 2016;19:10–19 CrossRef Medline
2. Di Bonaventura R, Montano N, Giordano M, et al. **Reassessing the role of brain tumor biopsy in the era of advanced surgical, molecular, and imaging techniques—a single-center experience with long-term follow-up.** *J Pers Med* 2021;11:909 CrossRef Medline
3. Haubold J, Demircioglu A, Gratz M, et al. **Non-invasive tumor decoding and phenotyping of cerebral gliomas utilizing multiparametric 18F-FET PET-MRI and MR fingerprinting.** *Eur J Nucl Med Mol Imaging* 2020;47:1435–45 CrossRef Medline
4. Bednarik P, Goranovic D, Svatkova A, et al. **<sup>1</sup>H magnetic resonance spectroscopic imaging of deuterated glucose and of neurotransmitter metabolism at 7 T in the human brain.** *Nat Biomed Eng* 2023;7:1001–13 CrossRef Medline
5. Wang Q, Zhang H, Zhang J, et al. **The diagnostic performance of magnetic resonance spectroscopy in differentiating high-from low-grade gliomas: a systematic review and meta-analysis.** *Eur Radiol* 2016;26:2670–84 CrossRef Medline
6. Gatenby RA, Gillies RJ. **Why do cancers have high aerobic glycolysis?** *Nat Rev Cancer* 2004;4:891–99 CrossRef Medline
7. Griffin JL, Kauppinen RA. **A metabolomics perspective of human brain tumours.** *FEBS J* 2007;274:1132–39 CrossRef Medline
8. Krishnamoorthy T, Radhakrishnan VV, Thomas B, et al. **Alanine peak in central neurocytomas on proton MR spectroscopy.** *Neuroradiology* 2007;49:551–54 CrossRef Medline
9. Maudsley AA, Andronesi OC, Barker PB, et al. **Advanced magnetic resonance spectroscopic neuroimaging: experts' consensus recommendations.** *NMR Biomed* 2021;34:e4309 CrossRef Medline
10. Kim JH, Chang KH, Na DG, et al. **Comparison of 1.5T and 3T 1H MR spectroscopy for human brain tumors.** *Korean J Radiol* 2006;7:156–61 CrossRef Medline
11. Sjøbakk TE, Lundgren S, Kristoffersen A, et al. **Clinical <sup>1</sup>H magnetic resonance spectroscopy of brain metastases at 1.5T and 3T.** *Acta Radiol* 2006;47:501–08 CrossRef Medline
12. McCarthy L, Verma G, Hangel G, et al. **Application of 7T MRS to high-grade gliomas.** *AJNR Am J Neuroradiol* 2022;43:1378–95 CrossRef Medline
13. Berrington A, Voets NL, Larkin SJ, et al. **A comparison of 2-hydroxyglutarate detection at 3 and 7 T with long-TE semi-LASER.** *NMR Biomed* 2018;31 CrossRef Medline
14. Wilson M, Andronesi O, Barker PB, et al. **Methodological consensus on clinical proton MRS of the brain: review and recommendations.** *Magn Reson Med* 2019;82:527–50 CrossRef Medline
15. Henning A. **Proton and multinuclear magnetic resonance spectroscopy in the human brain at ultra-high field strength: a review.** *Neuroimage* 2018;168:181–98 CrossRef Medline
16. Bingölbalı A, Fallone BG, Yahya A. **Comparison of optimized long echo time STEAM and PRESS proton MR spectroscopy of lipid olefinic protons at 3 Tesla.** *J Magn Reson Imaging* 2015;41:481–86 CrossRef Medline
17. Fernandes CC, Lanz B, Chen C, et al. **Measurement of brain lactate during visual stimulation using a long TE semi-LASER sequence at 7 T.** *NMR Biomed* 2020;33:e4223 CrossRef Medline
18. Penner J, Bartha R. **Semi-LASER <sup>1</sup>H MR spectroscopy at 7 Tesla in human brain: metabolite quantification incorporating subject-specific macromolecule removal.** *Magn Reson Med* 2015;74:4–12 CrossRef Medline
19. Wong D, Schranz AL, Bartha R. **Optimized in vivo brain glutamate measurement using long-echo-time semi-LASER at 7 T.** *NMR Biomed* 2018;31:e4002 CrossRef Medline
20. Prener M, Opheim G, Shams Z, et al. **Single-voxel MR spectroscopy of gliomas with s-LASER at 7T.** *Diagnostics (Basel)* 2023;13:1805 CrossRef Medline
21. Jones SE, Lee J, Law M. **Neuroimaging at 3T vs 7T: is it really worth it?** *Magn Reson Imaging Clin N Am* 2021;29:1–12 CrossRef Medline
22. Shi Z, Zhao X, Zhu S, et al. **Time-of-flight intracranial MRA at 3 T versus 5 T versus 7 T: visualization of distal small cerebral arteries.** *Radiology* 2022;305:E72 CrossRef Medline
23. Wei Z, Chen Q, Han S, et al. **5T magnetic resonance imaging: radio frequency hardware and initial brain imaging.** *Quant Imaging Med Surg* 2023;13:3222–40 CrossRef Medline
24. Liu Y, Wang W, Zheng Y, et al. **Magnetic resonance T1ρ quantification of human brain at 5.0 T: a pilot study.** *Front Phys* 2022;10:1016932 CrossRef Medline
25. Jiang Z, Sun W, Xu D, et al. **Stability and repeatability of diffusion-weighted imaging (DWI) of normal pancreas on 5.0 Tesla magnetic resonance imaging (MRI).** *Sci Rep* 2023;13:11954 CrossRef Medline
26. Zheng L, Yang C, Liang L, et al. **T2-weighted MRI and reduced-FOV diffusion-weighted imaging of the human pancreas at 5 T: a comparison study with 3 T.** *Med Phys* 2023;50:344–53 CrossRef Medline
27. Zheng L, Yang C, Sheng R, et al. **Renal imaging at 5 T versus 3 T: a comparison study.** *Insights Imaging* 2022;13:155 CrossRef Medline
28. Zhang Y, Yang C, Liang L, et al. **Preliminary experience of 5.0 T higher field abdominal diffusion-weighted MRI: agreement of apparent diffusion coefficient with 3.0 T imaging.** *J Magn Reson Imaging* 2022;56:1009–17 CrossRef Medline
29. Kraff O, Quick HH. **7T: physics, safety, and potential clinical applications.** *J Magn Reson Imaging* 2017;46:1573–89 CrossRef Medline
30. Scheenen TW, Klomp DW, Wijnen JP, et al. **Short echo time 1H-MRSI of the human brain at 3T with minimal chemical shift displacement errors using adiabatic refocusing pulses.** *Magn Reson Med* 2008;59:1–6 CrossRef Medline
31. Dacko M, Lange T. **Improved detection of lactate and β-hydroxybutyrate using MEGA-sLASER at 3 T.** *NMR Biomed* 2019;32:e4100 CrossRef Medline
32. Fuchs A, Luttje M, Boesiger P, et al. **SPECIAL semi-LASER with lipid artifact compensation for 1H MRS at 7 T.** *Magn Reson Med* 2013;69:603–12 CrossRef Medline
33. Ganji SK, An Z, Tiwari V, et al. **In vivo detection of 2-hydroxyglutarate in brain tumors by optimized point-resolved spectroscopy (PRESS) at 7T.** *Magn Reson Med* 2017;77:936–44 CrossRef Medline
34. Shams Z, van der Kemp WJ, Emir U, et al. **Comparison of 2-hydroxyglutarate detection with sLASER and MEGA-sLASER at 7T.** *Front Neurosci* 2021;12:718423 CrossRef Medline
35. Enoki T, Jomoto W, Yamano T, et al. **Influences of tumor volume and FWHM of the water peak and T2\* value of water on the detection rate of the choline peaks in proton MR spectroscopy of breast cancer at 3.0 T-MRI [in Japanese].** *Nihon Hoshasen Gijutsu Gakkai Zasshi* 2021;77:351–57 CrossRef Medline
36. Li X, Yang Y, Zhang B, et al. **Lactate metabolism in human health and disease.** *Signal Transduct Target Ther* 2022;7:305 CrossRef Medline
37. Ijare O, Baskin D, Pichumani K. **CBMT-01: alanine fuels energy metabolism of glioblastoma cells.** *Neuro Oncol* 2019;21:vi32–33 CrossRef Medline



# High-Grade Astrocytoma with Piloid Features: A Dual Institutional Review of Imaging Findings of a Novel Entity

 Neetu Soni,  Amit Agarwal,  Pranav Ajmera,  Parv Mehta,  Vivek Gupta, Mukta Vibhute, Maria Gubbiotti,  Ian T. Mark,  Steven A. Messina, Suyash Mohan, and  Girish Bathla



## ABSTRACT

**SUMMARY:** High-grade astrocytoma with piloid features (HGAP) is a recently identified brain tumor characterized by a distinct DNA methylation profile. Predominantly located in the posterior fossa of adults, HGAP is notably prevalent in individuals with neurofibromatosis type 1. We present an image-centric review of HGAP and explore the association between HGAP and neurofibromatosis type 1. Data were collected from 8 HGAP patients treated at two tertiary care institutions between January 2020 and October 2023. Demographic details, clinical records, management, and tumor molecular profiles were analyzed. Tumor characteristics, including location and imaging features on MR imaging, were reviewed. Clinical or imaging features suggestive of neurofibromatosis 1 or the presence of *NF1* gene alteration were documented. The mean age at presentation was 45.5 years (male/female = 5:3). Tumors were midline, localized in the posterior fossa ( $n = 4$ ), diencephalic/thalamic ( $n = 2$ ), and spinal cord ( $n = 2$ ). HGAP lesions were T1 hypointense, T2-hyperintense, mostly without diffusion restriction, predominantly peripheral irregular enhancement with central necrosis ( $n = 3$ ) followed by mixed heterogeneous enhancement ( $n = 2$ ). Two *NF1* mutation carriers showed signs of neurofibromatosis type 1 before HGAP diagnosis, with one diagnosed during HGAP evaluation, strengthening the HGAP-*NF1* link, particularly in patients with posterior fossa masses. All tumors were *IDH1* wild-type, often with *ATRX*, *CDKN2A/B*, and *NF1* gene alteration. Six patients underwent surgical resection followed by adjuvant chemoradiation. Six patients were alive, and two died during the last follow-up. Histone H3 mutations were not detected in our cohort, such as the common H3K27M typically seen in diffuse midline gliomas, linked to aggressive clinical behavior and poor prognosis. HGAP lesions may involve the brain or spine and tend to be midline or paramedian in location. Underlying neurofibromatosis type 1 diagnosis or imaging findings are important diagnostic cues.

**ABBREVIATIONS:** *ATRX* = Alpha thalassemia/mental retardation syndrome X-linked; DCE = dynamic contrast-enhanced; DMG = diffuse midline glioma; GTR = gross-total resection; HGAP = high-grade astrocytoma with piloid features; MAPK = mitogen-activated protein kinase; MC AAP = methylation-class anaplastic astrocytoma with piloid features; *NF1* = neurofibromatosis type 1; PA = pilocytic astrocytoma; PFS = progression-free survival; TPF3 = tumor protein p53

CNS tumor classification has incorporated several molecular markers and genetic mutations of prognostic value, as seen with the latest 5th edition of World Health Organization CNS tumor classification.<sup>1,2</sup> For characterizing tumors with unusual morphologic features and overlapping characteristics on conventional histology, DNA methylation profiling is helpful for accurate classification and can alter the diagnosis in up to 12% of patients.<sup>3</sup> Reinhardt et al<sup>4</sup> used DNA methylation profiling on


existing tumor data to identify a new subtype, initially termed “methylation-class anaplastic astrocytoma with piloid features (MC AAP),” later incorporated into the 5th edition as “High-Grade Astrocytoma with Piloid Features (HGAP).”<sup>5</sup>

HGAP is a rare tumor, primarily found in the posterior fossa (74%), but it can occur in supratentorial and spinal locations.<sup>4</sup> It can emerge de novo with neurofibromatosis type 1 (*NF1*) and has also been reported to progress from a prior lower-grade glioma, often pilocytic astrocytoma (PA). Cimino et al<sup>6</sup> identified 3 distinct epigenetic groups (g) of HGAPs based on DNA methylation: *gNF1* ( $n = 18$ ), *g1* ( $n = 72$ ), and *g2* ( $n = 54$ ), with median ages of 43.5, 47, and 32 years. Notably, *gNF1* was strongly associated with a clinical diagnosis of *NF1* (33.3%,  $P < .001$ ), posterior fossa localization, neurofibromin 1 (*NF1*) hypermethylation, and reduced progression-free survival (PFS) ( $P < .058$ ). The tumor histology varies and often displays frequent mitotic activity, elongated glial tumor cell processes, known as “piloid” features, and Rosenthal fibers or eosinophilic granular bodies. These

Received October 31, 2023; accepted after revision December 21.

From the Mayo Clinic (N.S., A.A., V.G.), Jacksonville, Florida; Mayo Clinic (P.A., P.M., I.T.M., S.A.M., G.B.), Rochester, Minnesota; College of Medicine (M.V.), St. George's University, Grenada, West Indies; MD Anderson Cancer Center (M.G.), University of Texas, Houston, Texas; and Perelman School of Medicine (S.M.), University of Pennsylvania, Philadelphia, Pennsylvania.

Please address correspondence to Neetu Soni, MD, DNB, FRCC, Radiology, Mayo Clinic, 4500 San Pablo Rd S, Jacksonville, FL 32224; e-mail: drneetusoni98@gmail.com; @NeetuSo27437480

 Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8166>

tumors feature a unique epigenetic profile: mitogen-activated protein kinase (MAPK) pathway activation (often *NF1*, *FGFR1*, or *BRAF* alterations), cyclin-dependent kinase inhibitor 2 (*CDKN2A/B*) deletions (>70%), adenosine triphosphate-dependent helicase (*ATRX*) loss (50%), chromosome 19q loss (>50%), and occasional *BRAF* duplications. Some patients with HGAP may also exhibit a methylated *MGMT* promoter. HGAP may resemble glioblastoma and diffuse midline glioma (DMG) on imaging, posing a diagnostic challenge. The prognosis is intermediate between *IDH*-mutant gliomas and *IDH* wild-type glioblastomas with a 5-year survival rate of around 50%.<sup>4</sup> Considering the rarity and uniqueness of HGAP, there is a lack of specific imaging biomarkers, which makes DNA methylation profiling the sole diagnostic method. Limited radiologic data are available, primarily derived from pathologic studies aimed at re-assigning CNS tumors to HGAP on the basis of DNA methylation.<sup>4,7,8</sup> Herein, we present a dual-institution review of imaging findings in HGAP, supplementing molecular testing for precise diagnosis.

## MATERIALS AND METHODS

The study was approved by the local institutional review boards at both institutions (Mayo Clinic Rochester, Minnesota Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania) with waived patient consent due to its retrospective nature. Institutional pathology records were searched between January 2020 and October 2023 for patients with an integrated HGAP diagnosis via histologic and genomic features (molecular DNA methylation classification and next-generation sequencing). We excluded patients in whom pathologists had determined that the DNA methylation confidence scores were below the agreed threshold (<0.9) for an HGAP diagnosis (as per the Bethesda classifier/Heidelberg classifier tool), implying a final reported non-HGAP integrated diagnosis.<sup>3</sup> A total of 8 unique cases of HGAP were identified, 6 from the first and 2 from the second institution. We gathered clinical, demographic, imaging data and histopathologic, immunohistochemical, surgical, and postoperative details. All patients underwent a 3T whole-body MR imaging (Magnetom Skyra or Magnetom Prisma and Tim Trio; Siemens). The anatomic imaging protocol included axial 3D T1-weighted MPRAGE, axial T2 FLAIR, DWI, 3D SWI, and sagittal T1-sampling perfection with application-optimized contrasts by using different flip angle evolution (SPACE sequence; Siemens) imaging using standard parameters. The postcontrast T1-weighted images were acquired with the same parameters as the precontrast acquisition after administration of three-quarters of the standard dose (0.1 mmol/kg) of gadoterate meglumine (Dotarem; Guerbet) IV contrast agent, administered twice for dynamic contrast-enhanced (DCE) and DSC scans (a total of 1.5 full standard dose) using a power injector (Medrad, Idianola, PA). The DCE scans were performed after the first dose and included a fast 3D spoiled gradient-echo sequence (TR/TE = 5.09 ms/1.57 ms; flip angle = 23°; section thickness = 3.5 mm; FOV = 22 × 22 cm<sup>2</sup>; matrix size = 256 × 256; 28 slices per measurement, with 30 sequential measurements). The DSC scans included a T2\*-weighted gradient-echo EPI sequence and were performed following the second dosage of contrast injection (TR/TE = 2000/45 ms; FOV = 22 × 22 cm<sup>2</sup>; matrix size = 128 × 128;

in-plane resolution = 1.72 × 1.72 mm<sup>2</sup>; section thickness = 3 mm; bandwidth = 1346 Hz/pixel; flip angle = 90°; EPI factor = 128; echo spacing = 0.83; acquisition time = 3 minutes 10 seconds). Forty-five sequential measurements were taken per section. The injection rate for both scans was 5 mL/s, followed by a 20-mL saline flush at the same rate.

Four patients (P2, P6, P7, P8) had dynamic DSC, and 2 (P7, P8) had additional DCE-perfusion MR imaging. One patient with a spinal HGAP (P3) had whole-body FDG-PET/CT. Two neuroradiologists (N.S., G.B.) reviewed neuroimaging studies in consensus for location and MR imaging characteristics, including cysts, hemorrhage, and enhancement patterns (heterogeneous, necrotic with rim or patchy enhancement) alongside comprehensive pathologic and immunohistochemical analyses. Additionally, treatment and outcome/survival data were also collected. PFS was defined as the duration from the initial surgery to radiologic/clinical recurrence. Overall survival, the patient survival duration from the initial diagnosis, was available for 2 patients.

## RESULTS

### Patient Characteristics

The mean and median age at diagnosis was 45.5 (SD, 19.39 years; range, 19–71 years; male/female = 5:3) and 43 years, respectively. Common presenting symptoms included headache, weakness, and backache, with isolated cases of vocal cord paralysis (P5) and hearing loss (P6). Three patients (P3, P5, P6) showed clinical stigmata of NF1 (café-au lait spots and cutaneous/plexiform neurofibromas), with a positive family history of NF1 in 1 (P5) (Table).

### Imaging Characteristics

One-half of the lesions (4/8) were located in the posterior fossa (pontomedullary, cerebellar hemisphere, pons-cerebellopontine angle, and pontomesencephalic) followed by supratentorial (2/8) (midline diencephalic/thalamic region) and intramedullary (2/8) localization. Most lesions abutted the pial (4/8) or ependymal (3/8) surface. Spinal lesions in P1 (T9–T10) and P3 (C5–T1) were characterized by long-segment (>3 vertebral segments) intramedullary expansile tumors with a dorsal predilection. Spinal HGAPs exhibited exophytic components near the pial surface and were found to adhere to the pia during surgery (Fig 1A, -C). One of the posterior fossa pontine lesions (P6) had a predominant exophytic cerebellopontine component extending into the internal auditory canal and prepontine cistern (Fig 2B). HGAP lesions were predominantly T1 hypointense, T2/FLAIR/STIR hyperintense, and mostly without diffusion restriction, except that 1 lesion (P5) showed patchy diffusion restriction. Three HGAPs (P4, P7, P8) had intralesional hemorrhage (Online Supplemental Data). Most lesions exhibited enhancement, though the enhancement varied from minimal patchy ill-defined (P5, Online Supplemental Data) to plaquelike (P1, Fig 1A) to heterogeneous enhancement involving the tumor to a variable extent, without a dominant single pattern. Three diencephalic/thalamic lesions (P2, P7, P8) displayed peripheral, irregular enhancement with central necrosis (Figs 1B and 2C, -D). One intracerebellar HGAP (P4) exhibited a multicystic lesion with peripheral enhancement and areas of hemorrhage (Fig 1D and Online Supplemental Data).

**Demographic and clinical features of the patient population with HGAP**

Patient No.	Sex	Age at Presentation (yr)	Location	Clinical Symptoms	Clinical/Imaging NF1 Features
1	M	44	Intramedullary T9-T10	Midback pain	None
2	M	42	Midline pontomesencephalic and thalamic tumor, intraventricular tumor seeding; drop mets on follow-up MR imaging spine (C5-C6)	Confusion, fatigue, and nausea, lower backache	None
3	M	74	Intramedullary C5-T1	Generalized weakness, difficulty walking, lower limb tingling and numbness	Scattered cutaneous/subcutaneous neurofibromas; multiple neurofibromas along cervicodorsal spine MR imaging; postsurgical resection of sciatic nerve plexiform NF1
4	F	38	Left cerebellum	Gradually progressive headache and dizziness	None
5	M	44	Pontomedullary	Nausea, dyspnea, vocal cord paralysis	Suboccipital neurofibroma
6	F	32	Right pontine lesion extending into the right cerebellopontine angle, internal auditory canal, prepontine cistern	Bilateral hearing loss	Hyperpigmented macules and papules on chest; multiple plexiform neurofibromas
7	M	19	Midline diencephalic/thalamic region	Severe headache	None
8	F	71	Midline diencephalic/thalamic region	Syncope, loss of consciousness, urinary incontinence	None

**Note:**—M indicates male; F, female; drop mets, leptomeningeal mets in spine.

MR perfusion in 4 patients (P2, P6, P7, P8) showed elevated perfusion parameters on DSC (relative CBV) and DCE ( $K^{\text{trans}}$ , volume transfer coefficient;  $V_e$ , extravascular extracellular volume fraction;  $V_p$ , blood plasma volume fraction; Online Supplemental Data). One (P2) pontomesencephalic HGAP showed intraventricular seeding, progressing on follow-up MRIs with more extensive ependymal involvement (Fig 1B). This patient also exhibited dorsal intramedullary T2-hyperintense signal at the C5–C6 level with focal leptomeningeal enhancement developing 5 months after the initial MR imaging, consistent with leptomeningeal spread, a finding rarely reported with HGAP (Online Supplemental Data).<sup>9</sup> One patient with a spinal HGAP (P3) showed increased tracer uptake on FDG-PET/CT, indicating the high-grade nature of the lesion (Online Supplemental Data). Figures 1 and 2, and the Online Supplemental Data depict HGAP imaging specifics.

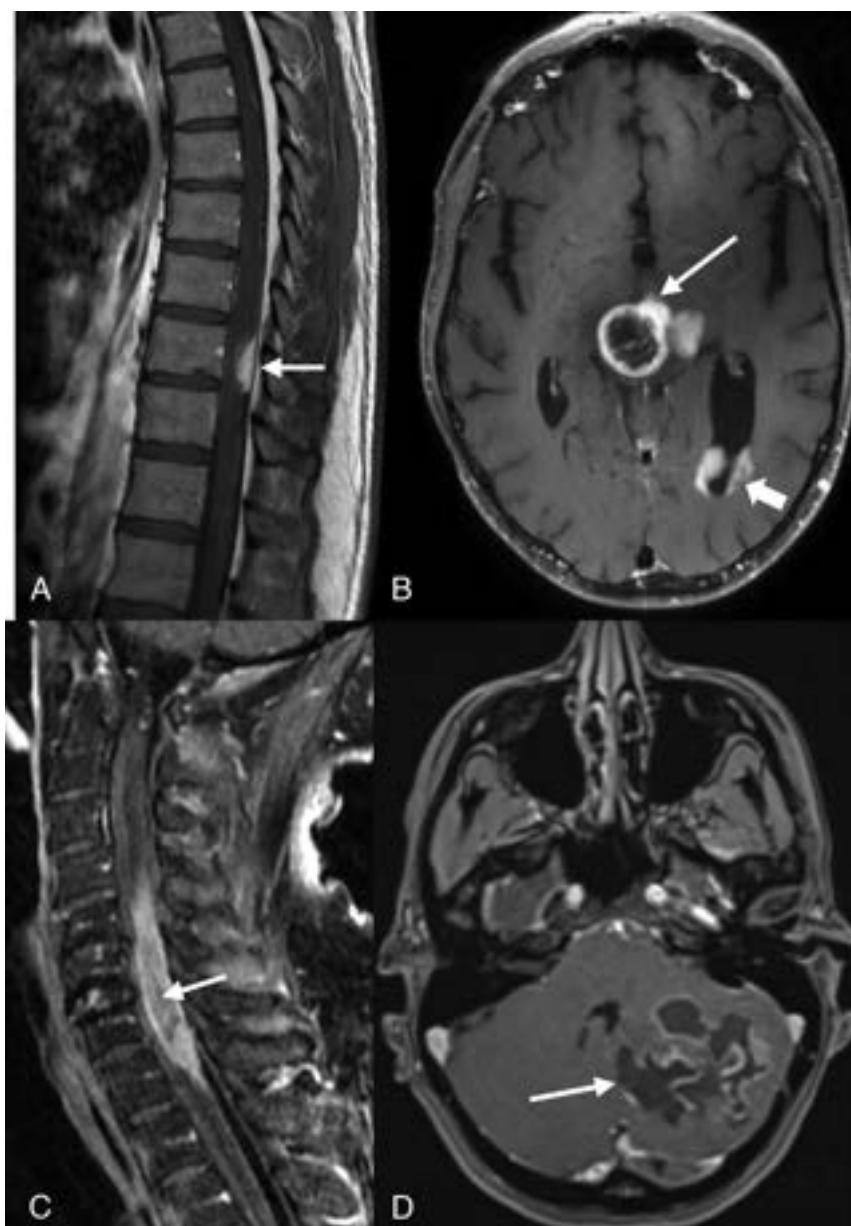
### Operative Details

Gross-total resection (GTR) was performed in 3 cases (P1, P3, P4); and subtotal resection, in 3 cases (P6, P7, P8) due to tumor vascularity. These patients received adjuvant chemoradiation (temozolomide, proton beam radiation therapy, and intensity beam radiation therapy) except one (P7), who died 1 month after surgery. Two patients did not undergo surgery due to multifocality (P2) and pontomedullary location (P5) and had only adjuvant chemoradiation (Online Supplemental Data).

### Histopathologic and Molecular Characteristics

Histopathology showed a moderately cellular glial tumor with variable mitotic activity and nuclear atypia. Some tumors showed

glomeruloid vasculature. Rosenthal fibers or eosinophilic granular bodies were seen in a minority. Microhemorrhage was present in 3 HGAPs (P6, P7, P8). Cystic changes were present in 3 HGAPs (P4, P7, P8). Next-generation sequencing data were obtained from various institutions with different gene region coverage. The most common genetic alterations were *CDKN2A/B* homozygous deletion (7/8, 87.5%), *ATRX* mutation (87.5%), *MGMT* promoter methylation (6/8, 75%), and *NF1* gene alteration (5/8) (62%). Six HGAP cases had co-occurring *CDKN2A/B* homozygous deletion and *ATRX* mutations. Other observed mutations included TP53 (3/6, 50%), RB1 loss (2/3), BRCA2 loss (1/3), and SETD2 (1/3). All tumors were *IDH1/2* wild-type. Histone H3 mutations, such as the common H3K27M typically seen in DMGs, linked to aggressive clinical behavior and poor prognosis, were not detected in our cohort.<sup>10</sup> A representative histopathologic and immunohistochemical profile is shown in Fig 3 and the Online Supplemental Data. *NF1* mutations were found in 62.5% (P1, P2, P3, P5, P6), of which 2 (P3, P6) had prior clinical (cutaneous neurofibromas) and imaging (plexiform neurofibroma) NF1 stigmata. These patients later developed intraspinal and posterior fossa HGAP after 6 months and 6 years of initial NF1 diagnosis, respectively. One patient (P5) was diagnosed with cutaneous neurofibromas during HGAP work-up along with an NF1 family history. The overall pathologic *NF1* mutation rate is 92.1% (70/76) in individuals with NF1 identified by next-generation sequencing.<sup>11</sup> CSF analysis for malignant cells was negative in all cases. Intraspinal HGAP (P3) was a highly cellular glioma with nuclear anaplasia, infiltrative growth, and high mitotic activity, with additional p53 overexpression that lacked microvascular proliferation or tumor necrosis.



**FIG 1.** Postcontrast-T1-weighted sagittal (A) image of a 44-year-old man with midback pain shows a well-defined enhancing lesion involving the dorsal aspect of thoracic cord at the T9–T10 level (arrow). Postcontrast T1-weighted axial image in a 42-year-old man (B) shows tumor with peripheral irregular enhancement at the level of thalamus (arrow) with infratentorial pontomesencephalic and cerebellar involvement (not shown) along with diffuse intraventricular tumor seeding (arrowhead). Postcontrast T1-weighted sagittal (C) image of a 74-year-old man with weakness and difficulty walking shows an intramedullary heterogeneously enhancing tumor involving nearly the entire cord C5–T1 (arrow) along with cord expansion. Postcontrast T1-weighted axial (D) image in a 38-year-old man demonstrates a large multicystic lesion with peripheral enhancement in the left cerebellum (arrow).

### Follow-Up

Mean and median follow-ups were at 17.4 and 23 months, respectively. While the treatments varied, the PFS was available in P1 (3 months) and P4 (26 months). One patient (P3) had a residual tumor after surgery without response. One patient (P1) underwent a re-resection after 15 months of initial GTR for recurrence. Six patients were alive at the time of the last follow-up, and 2 (P2, P7) died with 1–12 months of overall survival (Online Supplemental Data).

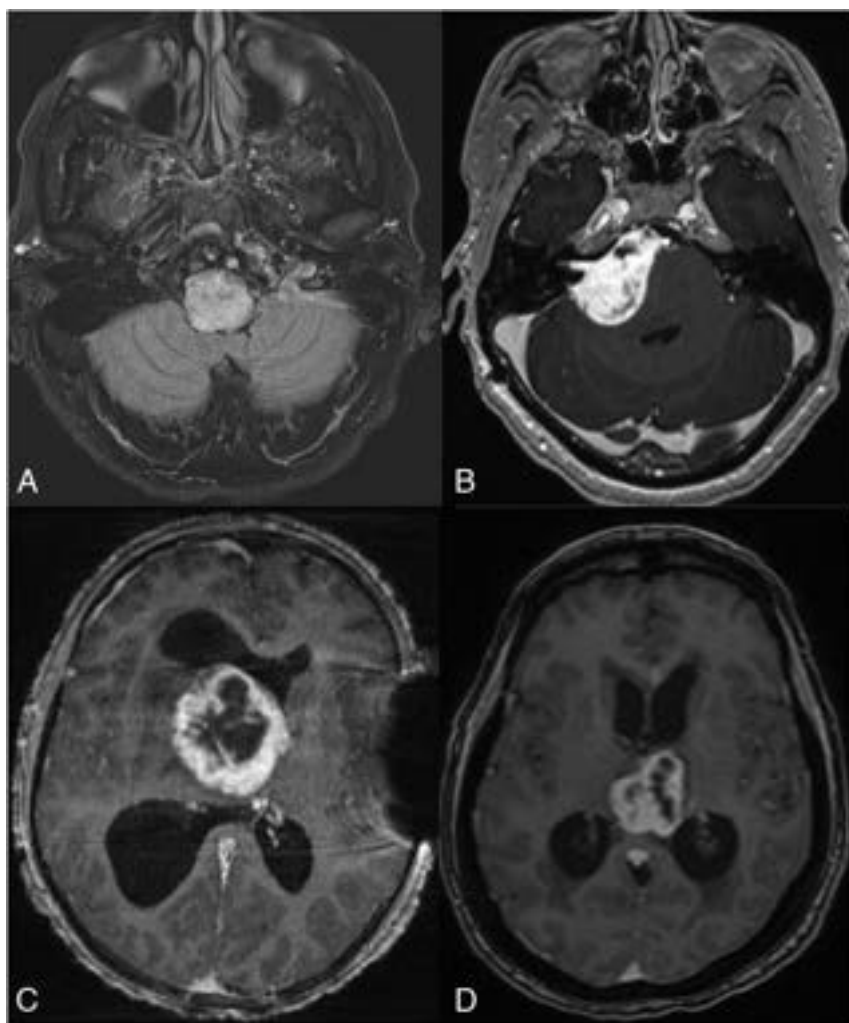
### DISCUSSION

HGAP is a new entity within circumscribed astrocytic gliomas, often resembling glioblastoma rather than PA with survival comparable with that of grade IV *IDH*-mutant astrocytoma.<sup>5</sup> DNA methylation profiling is performed to exclude HGAP, especially in atypical cerebellar gliomas (with *ATRX* loss or *CDKN2A/B* homozygous deletion) and HGAP-resembling tumors (like high-grade gliomas in NF1).<sup>6</sup> Reinhardt et al<sup>8</sup> found that molecular analysis reclassified approximately one-third of initially diagnosed glioblastoma cases as HGAP.<sup>12</sup> In a study by Lucas et al,<sup>13</sup> advanced molecular diagnosis reclassified 14 NF1-associated gliomas as HGAP, further supporting the NF1 and HGAP link. Although the exact HGAP incidence in NF1 remains uncertain, a confirmed NF1 diagnosis should alert clinicians to the potential HGAP risk.

The *NF1* gene is a tumor-suppressor gene mutated in individuals with NF1. *NF1* encodes the neurofibromin 1 protein, which negatively regulates the Ras/MAPK pathways via Ras inhibition. *NF1* mutations are prevalent in treatment-resistant gliomas and serve as supportive diagnostic biomarkers for HGAP and rosette-forming glioneuronal tumors.<sup>14,15</sup> Patients with NF1 are at a higher risk of developing low-grade gliomas. NF1 also predisposes to high-grade gliomas, with their prevalence being 10–50 times higher than in the general population. Romo et al<sup>16</sup> reported a nearly 300-fold higher prevalence of non-optic pathway gliomas in patients with NF1 (3.2%) compared with the general population (0.01%), with a median survival of 24 months. Nonoptic gliomas in adults with NF1 often have an aggressive clinical course, further underscoring the need to understand the pathobiology of NF1-associated gliomas better.<sup>16</sup> NF1-related gliomas resembling HGAP had worse

outcomes than the NF1-associated PA but fared better than sporadic *IDH* wild-type glioblastoma.<sup>13</sup> A substantial GB-HGAP overlap necessitates early HGAP detection because these HGAP cases often have a better prognosis and access to targeted therapies.<sup>13</sup> *NF1*-mutated tumors may respond to mammalian target of rapamycin (mTOR) and MAPK inhibitors, such as FDA-approved everolimus, temsirolimus, and emerging complex inhibitors.<sup>15</sup> This possibility raises the question of whether brain/





**FIG 2.** Postcontrast FLAIR axial MR image in a 44-year-old man (A) with dyspnea and vocal cord paralysis shows an expansile hyperintense lesion in the pontomedullary region with patchy ill-defined enhancement (not shown). Postcontrast T1-weighted axial MR images (not shown). B, A 32-year-old woman presenting with hearing loss demonstrates a large, right pontine tumor extending into the right cerebellopontine angle, internal auditory canal, and prepontine cistern. Postcontrast T1-weighted axial MR images (C and D) demonstrate a midline, diencephalic/thalamic region tumor with rim enhancement in a 19-year-old man (C) and a 71-year-old woman (D).

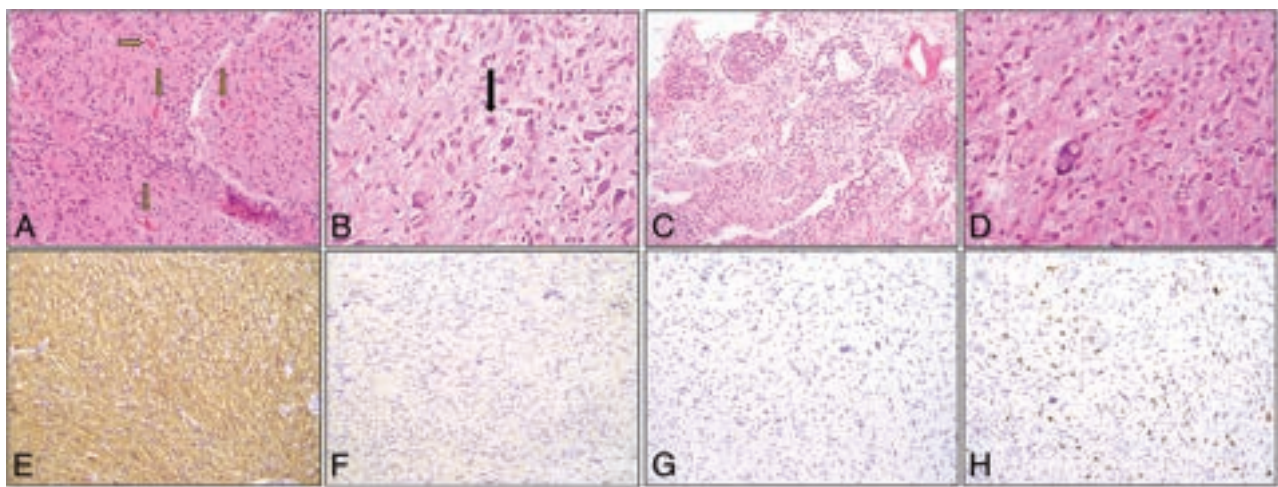
spine MR imaging screening is advisable for those with NF1, but the lack of large-scale studies on HGAP prevalence in NF1 prevents a conclusive answer.

One-half of HGAP lesions in our cohort involved the posterior fossa, followed by equal (25%) spinal cord and midline diencephalic/thalamic region involvement. Reinhardt et al<sup>4</sup> and Bender et al<sup>7</sup> identified the posterior fossa as the primary location in 74% and 66%, respectively. Spinal cord involvement varied between 7% (5/83)<sup>4</sup> and 33% (2/6).<sup>7</sup> We found 5 HGAPs (P1, P3, P5, P6, P8) in the midline, aligning with the observations of Romo et al<sup>16</sup> indicating increased midline involvement in NF1-associated tumors. Among these 5 cases, GTR was performed in 2 cases (P1, P3), with one (P1) achieving a PFS of 3 months, matching the findings of Romo et al of limited GTR success in NF1-associated tumors. While tumor-imaging descriptions for HGAP in the literature are sparse, they generally exhibit variable T1-weighted hypointensity, T2-weighted hyperintensity, and

heterogeneous enhancement. Diffusion restriction was generally absent. In our cohort, HGAP lesions closely imitated glioblastoma with variable heterogeneous enhancement and increased perfusion metrics. Intramedullary HGAPs were expansile with a dorsal predilection, and one (P3) demonstrated FDG-avidity. This imaging profile is similar to the findings of Bender et al<sup>7</sup> in their 6-case series, including radiotracer avidity on O-(2-[<sup>18</sup>F] fluoroethyl)-L-tyrosine PET/CT in 2 cases. Focal areas of signal intensity are common in patients with NF1 and may sometimes be challenging to differentiate from tumors. Perfusion-weighted MR imaging offers the potential to detect tumor-related neoangiogenesis, aiding in the differentiation to determine whether focal areas of signal intensity represent a low-grade tumor.<sup>17</sup> One of our patients with HGAP (P2) demonstrated ependymal and intraspinal leptomeningeal spread, indicating aggressiveness and a poor prognosis. Leptomeningeal or dural spread of HGAP is rarely reported, except for a recent study that identified leptomeningeal spread in a confirmed case of HGAP with a 7-year interval to metastasis and at the time of initial diagnosis in a possible HGAP. Adult high-grade astrocytic tumor types, including HGAP, are capable of leptomeningeal or dural spread.<sup>9</sup>

*MGMT* promoter methylation status is crucial for therapy and prognosis. Unmethylated status predicts a poor response to alkylating chemotherapy and worse prognosis.<sup>5,18</sup> In our cohort,

all except 1 patient received alkylating chemotherapy. We detected *MGMT* promoter methylation in 62.5% (5/8) of patients, higher than that reported by Bender et al<sup>7</sup> (33%, 2/6) and Reinhardt et al<sup>4</sup> (45%, 38/83). Four of our 5 cases of NF1 gene mutation also had co-occurring *ATRX* and *CDKN2A/B* mutations. Patients with *IDH* wild-type astrocytic gliomas with *CDKN2A/B* deletions and *ATRX* mutations generally have a better clinical outcome than patients with *IDH* wild-type glioblastoma, but the significance in NF1 is unclear.<sup>4</sup> Both *CDKN2A/B* and *ATRX* mutations were common as reported by Lucas et al,<sup>13</sup> who also noted *MAP* kinase pathway mutations (*NF1*, *FGFR-1*, *BRAF* genes). *ATRX* mutations were found in 87.5% (7/8) of our cases, slightly higher than previously reported in 45%–60% of cases.<sup>4,7</sup> Additionally, *CDKN2A/B* mutation was identified in 87.5% of our cases (7/8), previously reported in 80%–100% of cases.<sup>4,7</sup> *ATRX*, a chromatin remodeling protein, preserves genomic stability and often co-occurs with *IDH* mutations,



**FIG 3.** Morphologic and immunohistochemical profile of HGAP. A, H&E, 200 $\times$  stained section shows a moderately cellular tumor with abundant Rosenthal fibers and eosinophilic globular bodies (arrows). B, H&E (400 $\times$ ) stained section highlights cells with hairlike processes imparting a piloid appearance (arrow denotes a mitotic figure). C, H&E-stained section (100 $\times$ ) demonstrates glomeruloid vasculature often associated with HGAP. D, H&E-stained section (400 $\times$ ) shows bizarre atypia seen in scattered cells. E, Immunohistochemical (200 $\times$ ) stains. Glial fibrillary acidic protein shows diffuse positivity. F, *IDH1*-R132H with absence of staining. G, *ATRX* stain demonstrates loss. H, Ki-67.

potentially improving survival in low-grade gliomas, especially when lacking 1p/19q codeletions.<sup>19</sup> While *ATRX* alterations are characteristic of *IDH*-mutant astrocytoma, they are also frequently found in the *IDH* wild-type HGAP in conjunction with homozygous deletion of *CDKN2A/B*. This loss of *CDKN2A/B* generally indicates a higher grade and poorer prognosis, especially compared with PA.<sup>4,5</sup> HGAP is diagnosed solely by methylation profiling; however, the loss of *ATRX* on immunohistochemistry is helpful, especially with concurrent homozygous loss of *CDKN2A/B* to suggest this entity. In cases without *ATRX* mutation, *TERT* can be mutated (these are mutually exclusive events).<sup>6</sup>

In 2 cases (P7 and P8) *FGFR1* fusion was detected, which transforms primary astrocytes into highly proliferating midline gliomas.<sup>20,21</sup> *TP53* mutations were identified in 50% (3/6) of the patients in our study, a rate higher than the 5.4% reported in cases of HGAP by Cimino et al.<sup>6</sup> *TP53*, a tumor-suppressor gene, is involved in DNA damage response, cell cycle arrest, and apoptosis, with its mutations enhancing cell survival and chemoresistance.<sup>5,22</sup> *BRAF* V600E mutations were evaluated in 50% of cases, and all findings were negative. *BRAF* fusions are typically linked to World Health Organization grade 1 PA pediatric posterior fossa cases, showing a favorable outcome compared with supratentorial PAs in adults.<sup>4</sup> One patient with HGAP (P4) exhibited a rare genetic profile of *NTRK2* fusion, loss of P16 expression, a *SETD2* mutation, lack of *CDKN2A/B* deletion, and *MGMT* promoter methylation. Considering these characteristics, this case was contemplated for potential inclusion in a future trial involving *NTRK*-targeted therapy (entrectinib) in the event of disease progression. The significance of these mutations remains uncertain; however, they merit further investigation in future studies. *SETD2* mutations have been reported in high-grade hemispheric gliomas in older children<sup>23</sup> and adults with cerebellar glioblastomas.<sup>24</sup> Loss-of-function mutations in *SETD2* result in a deficiency of trimethylated histone H3K36, which appears specific to high-grade tumors. The nonsense mutation in *SETD2* in addition to a *KIAA1549-BRAF* fusion, contributes to the aggressive clinical course of this tumor.<sup>23,24</sup>

HGAP can be challenging to distinguish from PA, DMGs, and glioblastoma on imaging. PAs, typically found in children, are low-grade circumscribed gliomas with good prognosis (10-year survival rate >95%). These are strongly associated with *NF1* and *BRAF-KIAA1549* fusion/duplication. Most sporadic PAs arise from the cerebellum while affecting the optic pathway in patients with *NF1*. However, in adults, PAs, mostly supratentorial, exhibit a more aggressive clinical course, with a 50% 5-year survival rate and rare *BRAF* V600E alterations.<sup>25</sup> Radiologically, they show variable appearances, ranging from large cystic lesions with enhancing nodules to solid lesions with intense enhancement (~95%), calcification (20%), and hemorrhage.<sup>26</sup> The transformation of PA into a high-grade glioma is uncommon, especially in cases involving *BRAF* fusion, while about 20% of HGAPs with *BRAF* fusion suggest a possible link to earlier PAs or shared biologic traits.<sup>8,20</sup> Diffuse midline glioma (H3K27-altered) often occurs in pediatric patients without significant sex variation. It is characterized as an expansile, diffusely infiltrative lesion primarily in the thalamus, brainstem, and spinal cord, presenting as slightly T1 hypointense and T2 hyperintense on MR imaging with lesser enhancement, higher diffusion restriction, and perfusion. H3K27-altered DMGs have a poor prognosis, despite their histopathologic grade, mandating H3K27M detection and genetic analysis for definitive identification.<sup>27,28</sup> HGAP, prevalent in adults, associated with *NF1*, often displays high T2 signal and variable peripheral enhancement, and lacks diffusion restriction.<sup>7</sup> Both tumors share a dismal prognosis and are managed with surgery, radiation, and chemotherapy.<sup>7,29</sup> Both HGAP and glioblastoma, *IDH* wild-type high-grade CNS tumors, share similar imaging features and poor prognosis, requiring DNA methylation profiling for accurate differentiation. Glioblastoma, the most frequent primary brain tumor, typically manifests as a heterogeneous mass in the cerebral hemisphere, displaying irregular peripheral enhancement, central necrosis, diffusion restriction, extensive peritumoral edema, and glioblastoma histologic criteria (microvascular proliferation and necrosis).<sup>7</sup>

Our study did not reveal significant imaging differences between HGAPs with and without *NF1* mutations, except that spinal HGAPs were linked to *NF1*. Patient outcomes may be influenced by a complex interplay of factors, including *NF1* mutations, genetic modifiers, extent of tumor involvement, resectability, and treatment response. Neither our current study nor existing literature indicate any notable prognostic differences between HGAPs with and without *NF1* mutations; however, individual cases may differ due to unique genetic and molecular traits. Our study is constrained by a small patient sample, retrospective design, and varying next-generation sequencing data availability across institutions with differing gene region coverage. Nevertheless, this imaging-focused review adds valuable insight into the clinical, genetic diversity, and imaging characteristics of HGAP, particularly in *NF1*-associated HGAPs.<sup>4,8</sup> Targeted therapy may be pivotal due to frequent *MAP* kinase pathway gene alterations.<sup>6</sup>

## CONCLUSIONS

Radiologists should consider HGAP as a reasonable differential diagnosis in patients with high-grade/aggressive-appearing glial neoplasms in patients with *NF1*, especially with midline or paramedian localization. Nevertheless, the small sample size of our study due to HGAP rarity underscores the need for future larger studies or meta-analyses to assess the strength of the *NF1*-HGAP association.

**Disclosure forms** provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

- Osborn AG, Louis DN, Poussaint TY, et al. **The 2021 World Health Organization Classification of Tumors of the Central Nervous System: what neuroradiologists need to know.** *AJNR Am J Neuroradiol* 2022;43:928–37 CrossRef Medline
- Rigsby RK, Brahmabhatt P, Desai AB, et al. **Newly recognized CNS tumors in the 2021 World Health Organization Classification: imaging overview with histopathologic and genetic correlation.** *AJNR Am J Neuroradiol* 2023;44:367–80 CrossRef Medline
- Capper D, Jones DTW, Sill M, et al. **DNA methylation-based classification of central nervous system tumours.** *Nature* 2018;555:469–74 CrossRef Medline
- Reinhardt A, Stichel D, Schrimpf D, et al. **Anaplastic astrocytoma with piloid features, a novel molecular class of IDH wildtype glioma with recurrent *MAPK* pathway, *CDKN2A/B* and *ATRX* alterations.** *Acta Neuropathol* 2018;136:273–91 CrossRef Medline
- Louis DN, Perry A, Wesseling P, et al. **The 2021 WHO Classification of Tumors of the Central Nervous System: a summary.** *Neuro Oncol* 2021;23:1231–51 CrossRef Medline
- Cimino PJ, Ketchum C, Turakulov R, et al. **Expanded analysis of high-grade astrocytoma with piloid features identifies an epigenetically and clinically distinct subtype associated with neurofibromatosis type 1.** *Acta Neuropathol* 2023;145:71–82 CrossRef Medline
- Bender K, Perez E, Chirica M, et al. **High-grade astrocytoma with piloid features (HGAP): the Charité experience with a new central nervous system tumor entity.** *J Neurooncol* 2021;153:109–20 CrossRef Medline
- Reinhardt A, Stichel D, Schrimpf D, et al. **Tumors diagnosed as cerebellar glioblastoma comprise distinct molecular entities.** *Acta Neuropathol Commun* 2019;7:163 CrossRef Medline
- Kleinschmidt-DeMasters BK, Ormond DR. **Leptomeningeal metastases and dural spread in adult high-grade astrocytomas.** *J Neuropathol Exp Neurol* 2023;82:194–201 CrossRef Medline
- Miguel Llordes G, Medina Pérez VM, Curto Simón B, et al. **Epidemiology, diagnostic strategies, and therapeutic advances in diffuse midline glioma.** *J Clin Med* 2023;12:5261 CrossRef Medline
- Maruoka R, Takenouchi T, Torii C, et al. **The use of next-generation sequencing in molecular diagnosis of neurofibromatosis type 1: a validation study.** *Genet Test Mol Biomarkers* 2014;18:722–35 CrossRef Medline
- Yamashita K, Hiwatashi A, Togao O, et al. **MR imaging-based analysis of glioblastoma multiforme: estimation of IDH1 mutation status.** *AJNR Am J Neuroradiol* 2016;37:58–65 CrossRef Medline
- Lucas CH, Sloan EA, Gupta R, et al. **Multiplatform molecular analyses refine classification of gliomas arising in patients with neurofibromatosis type 1.** *Acta Neuropathol* 2022;144:747–65 CrossRef Medline
- Lobbous M, Bernstock JD, Coffee E, et al. **An update on neurofibromatosis type 1-associated gliomas.** *Cancers (Basel)* 2020;12:12 CrossRef Medline
- Tao J, Sun D, Dong L, et al. **Advancement in research and therapy of NF1 mutant malignant tumors.** *Cancer Cell Int* 2020;20:492 CrossRef Medline
- Romo CG, Piotrowski AF, Campian JL, et al. **Clinical, histological, and molecular features of gliomas in adults with neurofibromatosis type 1.** *Neuro Oncol* 2023;25:1474–86 CrossRef Medline
- Russo C, Russo C, Cascone D, et al. **Non-oncological neuroradiological manifestations in NF1 and their clinical implications.** *Cancers (Basel)* 2021;13:1831 CrossRef Medline
- Galbraith K, Snuderl M. **DNA methylation as a diagnostic tool.** *Acta Neuropathol Commun* 2022;10:71 CrossRef Medline
- Haase S, Garcia-Fabiani MB, Carney S, et al. **Mutant *ATRX*: uncovering a new therapeutic target for glioma.** *Expert Opin Ther Targets* 2018;22:599–613 CrossRef Medline
- Collins VP, Jones DT, Giannini C. **Pilocytic astrocytoma: pathology, molecular mechanisms and markers.** *Acta Neuropathol* 2015;129:775–88 CrossRef Medline
- Jones DT, Hutter B, Jäger N, et al; International Cancer Genome Consortium PedBrain Tumor Project. **Recurrent somatic alterations of *FGFR1* and *NTRK2* in pilocytic astrocytoma.** *Nat Genet* 2013;45:927–32 CrossRef Medline
- Marker DF, Agnihotri S, Amankulor N, et al. **The dominant TP53 hotspot mutation in IDH-mutant astrocytoma, R273C, has distinctive pathologic features and sex-specific prognostic implications.** *Neurooncol Adv* 2022;4:vda182 CrossRef Medline
- Fontebasso AM, Schwartzentruber J, Khuong-Quang DA, et al. **Mutations in *SETD2* and genes affecting histone H3K36 methylation target hemispheric high-grade gliomas.** *Acta Neuropathol* 2013;125:659–69 CrossRef Medline
- Nomura M, Mukasa A, Nagae G, et al. **Distinct molecular profile of diffuse cerebellar gliomas.** *Acta Neuropathol* 2017;134:941–56 CrossRef Medline
- Theeler BJ, Ellezam B, Sadighi ZS, et al. **Adult pilocytic astrocytomas: clinical features and molecular analysis.** *Neuro Oncol* 2014;16:841–47 CrossRef Medline
- Chourmouzi D, Papadopoulou E, Konstantinidis M, et al. **Manifestations of pilocytic astrocytoma: a pictorial review.** *Insights Imaging* 2014;5:387–402 CrossRef Medline
- Zhao JP, Liu XJ, Lin HZ, et al. **MRI comparative study of diffuse midline glioma, H3 K27-altered and glioma in the midline without H3 K27-altered.** *BMC Neurol* 2022;22:498 CrossRef Medline
- Vallero SG, Bertero L, Morana G, et al. **Pediatric diffuse midline glioma H3K27-altered: A complex clinical and biological landscape behind a neatly defined tumor type.** *Front Oncol* 2022;12:1082062 CrossRef Medline
- Aboian MS, Solomon DA, Felton E, et al. **Imaging characteristics of pediatric diffuse midline gliomas with histone H3 K27M mutation.** *AJNR Am J Neuroradiol* 2017;38:795–800 CrossRef Medline



# Comparison of Volumetric and 2D Measurements and Longitudinal Trajectories in the Response Assessment of *BRAF* V600E-Mutant Pediatric Gliomas in the Pacific Pediatric Neuro-Oncology Consortium Clinical Trial

Divya Ramakrishnan, Sarah C. Brünigk, Marc von Reppert, Fatima Memon, Nazanin Maleki, Sanjay Aneja, Anahita Fathi Kazerooni, Ali Nabavizadeh, MingDe Lin, Khaled Bousabarah, Annette Molinaro, Theodore Nicolaides, Michael Prados, Sabine Mueller, and Mariam S. Aboian

## ABSTRACT

**BACKGROUND AND PURPOSE:** Response on imaging is widely used to evaluate treatment efficacy in clinical trials of pediatric gliomas. While conventional criteria rely on 2D measurements, volumetric analysis may provide a more comprehensive response assessment. There is sparse research on the role of volumetrics in pediatric gliomas. Our purpose was to compare 2D and volumetric analysis with the assessment of neuroradiologists using the Brain Tumor Reporting and Data System (BT-RADS) in *BRAF* V600E-mutant pediatric gliomas.

**MATERIALS AND METHODS:** Manual volumetric segmentations of whole and solid tumors were compared with 2D measurements in 31 participants (292 follow-up studies) in the Pacific Pediatric Neuro-Oncology Consortium 002 trial (NCT01748149). Two neuroradiologists evaluated responses using BT-RADS. Receiver operating characteristic analysis compared classification performance of 2D and volumetrics for partial response. Agreement between volumetric and 2D mathematically modeled longitudinal trajectories for 25 participants was determined using the model-estimated time to best response.

**RESULTS:** Of 31 participants, 20 had partial responses according to BT-RADS criteria. Receiver operating characteristic curves for the classification of partial responders at the time of first detection (median = 2 months) yielded an area under the curve of 0.84 (95% CI, 0.69–0.99) for 2D area, 0.91 (95% CI, 0.80–1.00) for whole-volume, and 0.92 (95% CI, 0.82–1.00) for solid volume change. There was no significant difference in the area under the curve between 2D and solid ( $P = .34$ ) or whole volume ( $P = .39$ ). There was no significant correlation in model-estimated time to best response ( $\rho = 0.39$ ,  $P > .05$ ) between 2D and whole-volume trajectories. Eight of the 25 participants had a difference of  $\geq 90$  days in transition from partial response to stable disease between their 2D and whole-volume modeled trajectories.

**CONCLUSIONS:** Although there was no overall difference between volumetrics and 2D in classifying partial response assessment using BT-RADS, further prospective studies will be critical to elucidate how the observed differences in tumor 2D and volumetric trajectories affect clinical decision-making and outcomes in some individuals.

**ABBREVIATIONS:** AUC = area under the curve; BT-RADS = Brain Tumor Reporting and Data System; MinR = minor response; nRMSE = normalized root mean squared error; PD = progressive disease; PNOC = Pacific Pediatric Neuro-Oncology Consortium; PR = partial response; RAPNO = Response Assessment in Pediatric Neuro-Oncology; ROC = receiver operating characteristic; SD = stable disease

**G**liomas make up 53% of all primary brain and CNS tumors in children. While the most common glioma subtype is

pilocytic astrocytoma, malignant gliomas have a higher incidence in children with increasing age. Pediatric gliomas make

Received September 25, 2023; accepted after revision December 3.

From the Department of Radiology and Biomedical Imaging (D.R., M.v.R., F.M., N.M., M.L., M.S.A.), Department of Therapeutic Radiology (S.A.), and Center for Outcomes Research and Evaluation (S.A.), Yale School of Medicine, New Haven, Connecticut; Department of Biosystems Science and Engineering (S.C.B.), ETH Zürich, Basel, Switzerland; Swiss Institute for Bioinformatics (S.C.B.), Lausanne, Switzerland; Department of Neuroradiology (M.v.R.), Leipzig University Hospital, Leipzig, Germany; Center for Biomedical Image Computing and Analytics (A.F.K.), University of Pennsylvania, Philadelphia, Pennsylvania; Center for Data-Driven Discovery in Biomedicine (A.N.), Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; Visage Imaging (M.L.), San Diego, California; Visage Imaging (K.B.), Berlin, Germany; Departments of Neurological Surgery (A.M.) and Neurology, Neurosurgery, and Pediatrics (M.P., S.M.), University of California San Francisco, San Francisco, California; Caris Life Sciences (T.N.), Irving, Texas; and Children's University Hospital Zürich (S.M.), Zürich, Switzerland.

D.R. was supported by the RSNA Research & Education Foundation's Research Medical Student Grant and the Gershon, Gardner, Hirsch, Weiss Yale One-Year Sponsored Fellowship award. S.C.B. was supported by the Botnar Research Center for Child Health Postdoctoral Excellence Program (#PEP-2021-1008). M.v.R. received

funding from the German Academic Scholarship Foundation. M.S.A. received funding from the American Society of Neuroradiology Fellow Award 2018.

This publication was made possible by KL2 TR001862 (M.S.A.) from the National Center for Advancing Translational Science, components of the National Institutes of Health, and National Institutes of Health Roadmap for Medical Research.

Its contents are solely the responsibility of the authors and do not necessarily represent the official view of the National Institutes of Health.

D. Ramakrishnan and S.C. Brünigk are co-first authors.

The Pacific Pediatric Neuro-Oncology Consortium (PNOC) PNOC-002 was supported by funds from Genentech/Roche, the PNOC Foundation, and the Pediatric Brain Tumor Foundation.

Please address correspondence to Divya Ramakrishnan, BS, Boardman Building, 330 Cedar St, Rm CB-30, New Haven, CT 06510; e-mail: divya.ramakrishnan@yale.edu; @DivyaRamakris19

Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8189>



up a heterogeneous class of tumors with prognosis depending on tumor location, histology/grade, and age at diagnosis. The 10-year progression-free survival of patients with low-grade gliomas is 96%, while those with glioblastomas have a progression-free survival below 20%.<sup>1</sup> While surgery, chemotherapy, and radiation therapy are currently the standard treatments for pediatric glioma, therapies targeted to specific genetic mutations can also be used to prevent disease progression.<sup>2</sup> Although the *BRAF* V600E mutation confers a poor prognosis in pediatric gliomas, studies have shown that treatment with a *BRAF* inhibitor, such as vemurafenib or dabrafenib, can lead to disease stabilization or response. Despite these promising results, disease re-progression has been observed after discontinuation of the *BRAF* inhibitor, and pediatric high-grade gliomas do not have a response that is as sustainable as low-grade gliomas.<sup>3</sup> Pediatric glioma response assessment is largely based on imaging changes. However, the complex appearance of these tumors on imaging makes it challenging to accurately assess treatment response or failure.<sup>4</sup>

The Response Assessment in Pediatric Neuro-Oncology (RAPNO) criteria were developed to address the unique challenges of assessing pediatric gliomas on imaging, including their diffuse nature, cystic components, irregular borders, and variable enhancement pattern. RAPNO criteria are based on 2D measurements of tumor on either T1 contrast-enhancing or T2/FLAIR sequences chosen on the basis of tumor grade and location.<sup>5</sup> Given the challenge of applying 2D criteria to heterogeneous tumors like pediatric gliomas,<sup>4</sup> volumetrics (ie, an assessment of the 3D tumor volume) may better characterize a response.<sup>6</sup> While the RAPNO criteria provide some recommendations on volumetric assessment, they are not considered standard practice.<sup>5</sup> Moreover, the volume-extrapolated RAPNO thresholds used in volumetric assessment are derived from the mathematic extrapolation of 2D thresholds to a perfect sphere, assuming uniform tumor growth.<sup>7</sup> Given the heterogeneous and irregular growth pattern of pediatric gliomas, further research is needed to validate the volumetric RAPNO criteria in practice.<sup>4</sup> However, previous research has largely focused on the role of volumetrics in adult brain tumors.<sup>8–11</sup> Additionally, elucidation of the dynamics of tumor response is critical for accurate clinical management and evaluation of novel therapeutics.

We performed a retrospective analysis of images from the Pacific Pediatric Neuro-Oncology Consortium (PNOC) 002 clinical trial (NCT01748149) of vemurafenib in *BRAF* V600E-mutant pediatric gliomas.<sup>12</sup> The primary purpose of our study was to compare volumetric and 2D-based response assessment with a reference standard, the Brain Tumor Reporting and Data System (BT-RADS), which is a structured and standardized framework for neuroradiologists' visual assessment of response based on several factors, including tumor enhancement, nonenhancing tumor components, and tumor mass effect.<sup>13</sup> We hypothesized that volumetric assessment would correlate more closely with BT-RADS than 2D measurements. Because clinical decision-making is often based on trends observed across time rather than single timepoints, we also used mathematic modeling to understand the differences in volumetric and 2D longitudinal posttreatment trajectories.

**Table 1: Participant demographics**

Demographic Characteristic	No. (%)
Median age (range) (yr)	9 (3–19)
Sex	
Male	19 (61)
Female	12 (39)
Tumor location <sup>a</sup>	
Supratentorial	22 (71)
Suprasellar/optic pathway	5 (16)
Posterior fossa	4 (13)
Tumor pathology	
Pilocytic astrocytoma	13 (42)
Ganglioglioma	7 (23)
Astrocytoma (other)	6 (19)
Pleomorphic xanthoastrocytoma	2
Anaplastic astrocytoma (grade 3)	1
Fibrillary astrocytoma	1
Diffuse astrocytoma (grade 2)	1
Low-grade astrocytoma (NOS)	1
Malignant glioma	4 (13)
NOS	2
Glioneuronal tumor with desmoplasia	1
Diffuse midline glioma ( <i>H3K27M</i> -mutant)	1
Glioblastoma (NOS)	1 (3)

**Note:**—NOS indicates not otherwise specified.

<sup>a</sup> When >1 lesion is present, location indicates most of the tumor burden.

## MATERIALS AND METHODS

### Trial Data and Measurements

We conducted a retrospective analysis of 31 participants (292 total follow-up scans) enrolled in the PNOC-002 trial, a phase I/II study of vemurafenib in *BRAF* V600E-mutant pediatric gliomas (Table 1).<sup>12</sup> Before trial enrollment, all participants had recurrence of tumor after surgery and at least one other therapy (chemo- or radiation therapy) and were not on any other treatment other than vemurafenib during the trial period. A board-certified neuroradiologist (M.S.A.) unblinded to follow-up scans performed 2D measurements on the pretreatment sequence (T2/FLAIR, T1 contrast-enhancing, or T1 pre-contrast) in a plane that best captured solid tumor, with cystic components included only if inseparable from solid components. The same sequence and plane were used to perform 2D measurements in all subsequent follow-up images to ensure consistency. T2/FLAIR sequences were used for most participants. However, in cases in which 2D measurements could not reliably and consistently capture tumor burden (eg, tumors with irregular or diffuse borders on T2/FLAIR) or the solid tumor portion was difficult to clearly visualize on T2/FLAIR, alternative sequences (ie, T1 or T1 contrast-enhancing) were used according to the neuroradiologist's clinical judgment. The T1 contrast-enhancing sequence was only chosen in cases in which there was consistent tumor enhancement without transient or variable patterns of change. The Online Supplemental Data show the 2D measurements performed for each participant on the pretreatment image. According to RAPNO guidelines,<sup>5</sup> the minimal bidimensional length of a lesion was 10 mm. Lesions that did not meet these criteria were not included in 2D or volumetric measurements.

3D segmentations of whole tumor (ie, including solid and cystic components) on FLAIR and cystic tumor on T2 were performed by a fourth-year medical student (D.R.) and corrected by

**Table 2: Response-assessment criteria**

Criteria	PR	MinR	SD	PD
BT-RADS		1a (PR)	1b, 2, 3a, 3b (SD)	3c, 4 (PD)
RAPNO	≥50% Decrease	≥25% but <50% Decrease	<25% Decrease or <25% increase	≥25% increase or appearance of new lesion
Volume-extrapolated RAPNO	≥65% Decrease	≥35% but <65% Decrease	<35% Decrease or <40% increase	≥40% increase or appearance of new lesion

the same board-certified neuroradiologist who performed the 2D measurements (M.S.A.). Segmentations were performed on a PACS-integrated annotation platform whereby a volumetric tool automatically extrapolated manual segmentations performed at the outermost tumor borders.<sup>14</sup> These extrapolated segmentations were then manually adjusted slice by slice. Solid tumor volume was calculated by subtracting the cystic volume on T2 from the whole volume on FLAIR.

### Comparison of 2D and Volumetrics with the Visual Assessment of Neuroradiologists

We compared 2D and volumetric RAPNO assessments with a reference standard, which we defined as the consensus score between 2 neuroradiologists' visual assessments of the tumor response using the BT-RADS criteria.<sup>13</sup> This comparison was performed for all participants at the end of treatment or last available on-treatment scan (for participants still on therapy at the time of analysis). BT-RADS is used in several clinical neuroradiology practices and represents the clinical interpretation of imaging findings with scores corresponding to definitive recommendations in clinical monitoring frequency and management.<sup>13</sup> Standard and volume-extrapolated RAPNO criteria (Table 2) were applied to percentage change from pretreatment in the 2D area (sum of the product of maximal bidimensional diameters of all lesions) and whole volume, respectively. We used the pretreatment scan, which was also postsurgical, as the baseline.<sup>5</sup> BT-RADS scores from a pretreatment baseline were assigned to each follow-up scan on the basis of mutual consensus between 2 board-certified neuroradiologists (M.S.A. and F.M.). We classified a BT-RADS score of 1a as partial response (PR), 1b to 3b as stable disease (SD), and 3c to 4 as progressive disease (PD) (Table 2). Paired log-rank tests were used to compare the time to the first response (confirmed on a consecutive scan at least 4 weeks later) between BT-RADS and each of the following criteria: 2D area change using the RAPNO minor response (MinR) threshold and whole and solid volume change using the volume-extrapolated RAPNO MinR threshold.

Empirical receiver operating characteristic (ROC) curves were plotted to compare the performance of 2D, solid-volume, and whole-volume change in classifying participants as having BT-RADS PR at 2 time points: 1) the end of treatment or last available on-treatment follow-up scan, and 2) the time of the BT-RADS first-confirmed response or first follow-up scan if no response. The DeLong test from the pROC<sup>15</sup> package in R (<https://www.rdocumentation.org/packages/pROC/versions/1.18.5>) was used to compare whether the area under the curve (AUC) values were statistically different between measurement methods at both time points. A 95% confidence interval for the AUC was computed using 2000 stratified bootstrap replicates. All statistical analyses were performed in R studio, Version 2022.12.0 + 353 (<http://rstudio.org/download/desktop>).

### Modeling 2D and Volumetric Longitudinal Treatment Response

We describe tumor growth mechanistically as the action and counteraction between treatment response and the onset of therapy resistance to parameterize the obtained tumor growth curves.<sup>16,17</sup> This model uses 3 patient-specific parameters characterizing the tumor growth rate ( $\lambda$ ), treatment-induced tumor shrinkage ( $\gamma_0$ ), and therapy resistance ( $\epsilon$ ). The tumor volume  $V$  is approximated as an exponential function in time at rate  $\lambda$ . In contrast, systemic therapy with vemurafenib is modeled to lead to an exponential reduction in tumor volume at rate  $\gamma$ :

$$1) \quad \frac{dV}{dt} = \lambda V - \gamma(t)V.$$

Given the onset of therapy resistance, we, in turn, describe the treatment efficacy as time-dependent with an exponential decay at rate  $\epsilon$ :

$$2) \quad \frac{d\gamma}{dt} = -\epsilon\gamma.$$

The analytical solution to this system of ordinary differential equations is

$$3) \quad V(t) = V_0 e^{\lambda t} + \frac{\gamma_0}{\epsilon} \left( e^{-\epsilon(t-t_0)-1} \right),$$

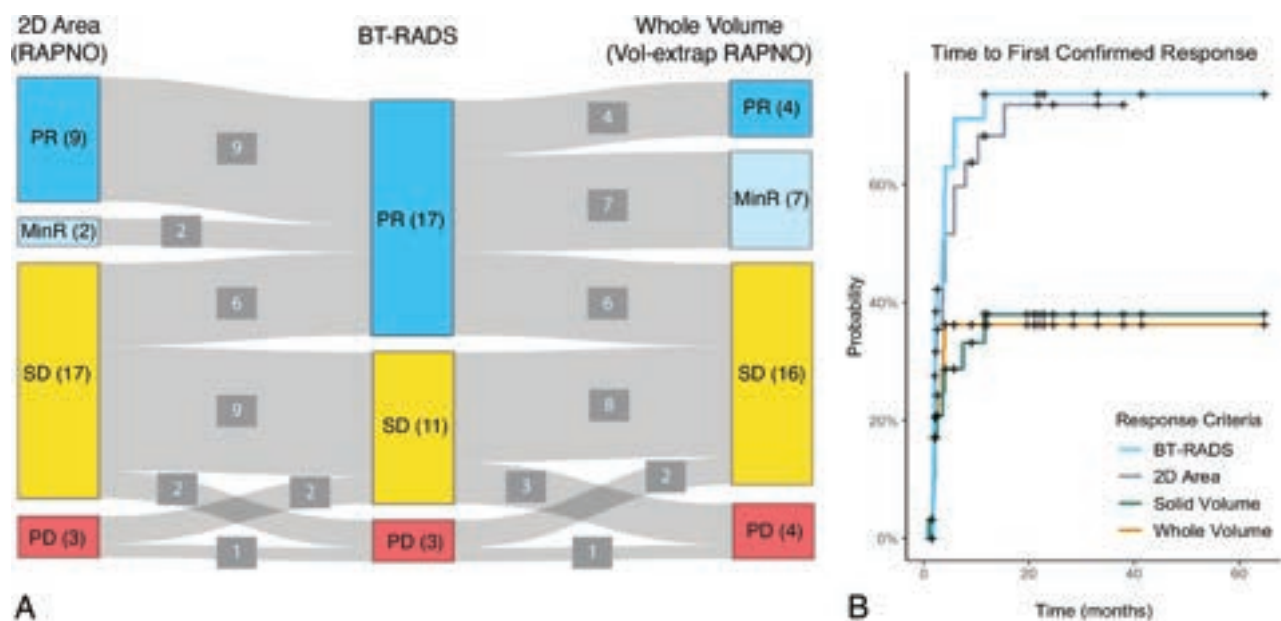
where  $\gamma_0$  indicates the initial condition at therapy initiation at time  $t_0$ .  $V_0$  is the tumor volume (or 2D area) assessed at the first scan. Patient data in either 2D or whole-volume tumor growth curves were fitted to Equation 3 using the python lmfit library (<https://lmfit.github.io/lmfit-py/>) in a bootstrap approach in which each data point was resampled  $n = 500$  times from a normal distribution around the observed value (mean) with a 10% standard deviation as an estimate of random error stemming from contouring uncertainty. Given the choice of 3 patient-specific model parameters, only growth trajectories comprising at least 3 data points (ie, 2 follow-up scans) were included in this analysis.

### Assessing Agreement between 2D and Volumetric Response Curves

From the model-obtained parametrizations, the estimated time to best response, ie, the minimal tumor volume ( $t_{Vmin}$ ), was used for further analysis as a notion of the overall trend in the data:

$$4) \quad t_{Vmin} = \frac{1}{\epsilon} \ln \left( \frac{\gamma_0}{\lambda} \right) + t_0.$$

We assigned data points preceding  $V_{min}$  (including a 10% uncertainty window) as PR, those around (within a 10% uncertainty window)  $V_{min}$  as SD, and those beyond  $V_{min}$  as PD. We



**FIG 1.** Comparison of response criteria. A, Sankey plot comparing the response classification for 31 participants at the end of treatment or last available follow-up scan based on 2D area percentage change with RAPNO thresholds, BT-RADS, and whole-volume percentage change with volume-extrapolated RAPNO thresholds. B, Time-to-response curves for 31 participants using BT-RADS, 2D area percentage change with the RAPNO MinR threshold, and solid/whole-volume percentage change with a volume-extrapolated RAPNO MinR threshold. Response had to be confirmed on a consecutive scan at least 4 weeks later.

quantified the correlation between  $t_{vmin}$  obtained from 2D and volumetric fits using the Spearman rank correlation coefficient  $\rho$ . The time to transition between response categories (ie, PR to SD or SD to PD) was also compared for all patients in whom such transitions were observed. RAPNO criteria require confirmation of response assessment on 2 consecutive scans,<sup>5</sup> and the median time interval between 2 consecutive scans in this data set was 90 days. Therefore, we defined “transition agreement” as a difference in the time to transition between response categories of <90 days. We used the Wilcoxon rank-sum test to compare the following quantitative features between participants with agreement and disagreement in 2D-versus-whole-tumor trends: FLAIR sequence slice thickness, spacing, pretreatment whole and cystic volume, and the relative model fit performance of 2D-versus-whole-tumor curves quantified by normalized root mean squared error (nRMSE). The latter was interpreted as a proxy for uncertainty in measurement/segmentation. All statistical tests were performed using the *scipy* stats module in python (<https://docs.scipy.org/doc/scipy/tutorial/stats.html>).

## RESULTS

### Comparing 2D and Volumetrics Using RAPNO and BT-RADS Criteria

Of 31 participants, 26 completed treatment while 5 were still on treatment at the time of analysis. The end of treatment ( $n=26$ ) and last available on-treatment scans ( $n=5$ ) had a median follow-up time from the pretreatment baseline scan of 21 months (range, 2–69 months). BT-RADS criteria classified 17 participants as having PR, 11 as SD, and 3 as PD. Of the 17 participants with BT-RADS PR, 11 (65%) were classified as having either PR or MinR by both 2D area and whole volume using standard and

volume-extrapolated RAPNO thresholds, respectively. The remaining 6 of 17 BT-RADS PR participants were classified as having SD by both methods. While most participants were classified as having PR by BT-RADS ( $n=17$ ), most were classified as having SD by both 2D area ( $n=17$ ) and whole volume ( $n=16$ ) (Fig 1A).

Time-to-response curves were plotted using different response criteria (Fig 1B). Of 31 participants, 20 (65%) had a response (median time to response = 3.6 months) with BT-RADS criteria, 19 (61%) (median time to response = 4.0 months) with 2D area using the MinR RAPNO threshold, and 10 (32%) with both whole and solid volume using the volume-extrapolated MinR RAPNO threshold. Paired log-rank tests revealed a significant difference in response time between BT-RADS and solid ( $P=.02$ ) and whole ( $P=.04$ ) volume. However, there was no significant difference between BT-RADS and 2D area ( $P=.40$ ).

### BT-RADS Classification Performance of 2D versus Volumetric Change

The ROC curve for classification of BT-RADS PR versus non-PR (SD/PD) at the end of treatment ( $n=26$ ) or last available on-treatment scans ( $n=5$ ) had an AUC of 0.92 (95% CI, 0.80–1.00) for 2D area change and 0.99 (95% CI, 0.97–1.00) for both whole- and solid-volume change (Fig 2A). DeLong tests revealed no significant difference in the AUC between 2D area and solid ( $P=.17$ ) or whole ( $P=.17$ ) volume. Of 31 participants, 20 experienced a confirmed BT-RADS PR during treatment, while the remaining 11 did not have confirmed BT-RADS PR. The ROC curve for differentiation of BT-RADS PR and non-PR at the time of first confirmed BT-RADS PR or first available follow-up (median follow-up time of scans = 2 months) yielded an AUC of 0.84 (95% CI, 0.69–0.99) for 2D area, 0.91 (95% CI, 0.80–1.00) for



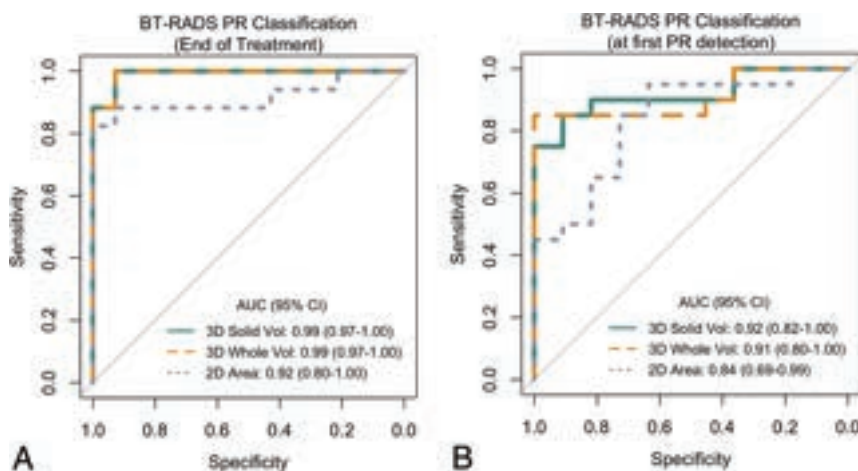
whole volume, and 0.92 (95% CI, 0.82–1.00) for solid volume (Fig 2B). DeLong tests again revealed no statistically significant difference in the AUC between the 2D area and solid (P = .34) or whole volume (P = .39) (Fig 2B).

### Longitudinal Response Assessment of 2D and Volumetric Change

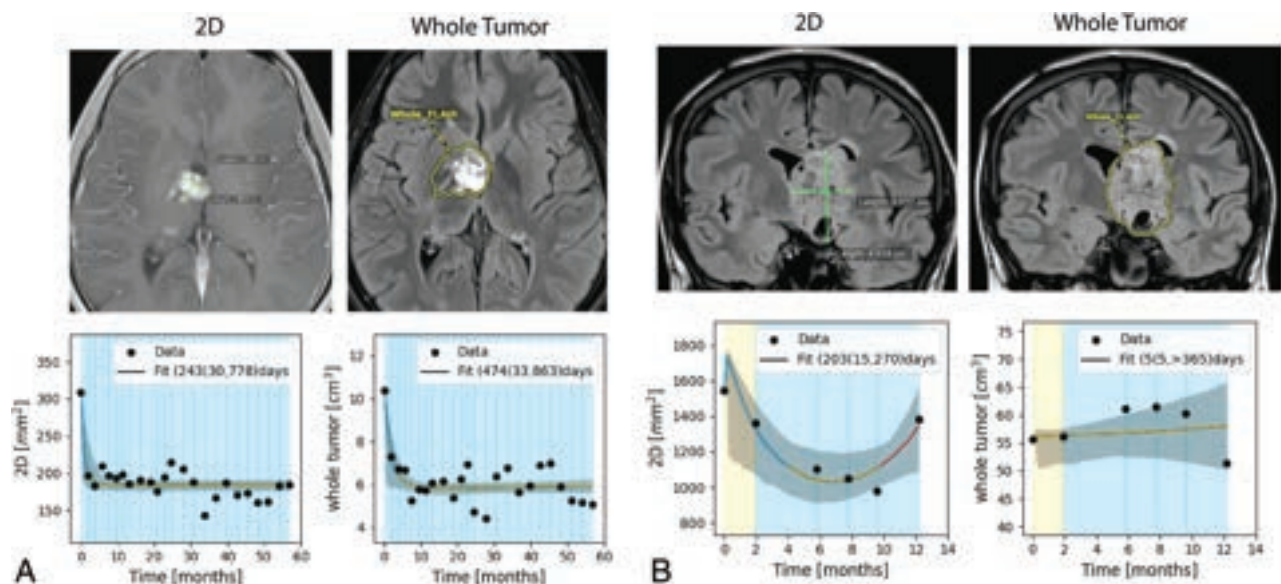
Longitudinal model fits were performed for 25 of the 31 participants. Model fits could not be performed for the remaining 6 participants because 5 of them had only 2 timepoints and 1 (participant 16) had a tumor growth pattern that did not comply with model-intended dynamics (Online Supplemental Data). Figure 3 shows 2D-versus-whole-volume growth trajectories

from 2 participants with relevant model fits, fit confidence intervals, and model-based label assignments (line color). Plots for all other participants are provided in the Online Supplemental Data. Figure 4 summarizes the model-estimated time to best response for 2D-versus-whole volume. There was no significant correlation between the values of 2D and whole-volume time to best response ( $\rho = 0.39$ ,  $P = .054$ ), with notable deviations in these trajectories indicated by deviation from the diagonal line in Fig 4. Given the many partially responding participants in this cohort and the associated limitations in fitting tumor regrowth, there were large uncertainties in the assessment of time to best response from model fits, demonstrated by the wide confidence intervals in Fig 4.

The Online Supplemental Data summarize model agreement and imaging protocol parameters for all participants with model fits. The difference in model-assigned time to transition from PR to SD between 2D and whole-volume trajectories was  $\geq 90$  days in 8 of 25 participants. Of the 8 participants, 2D showed a later PR-to-SD transition in 4 participants and an earlier transition in 1 participant. Three participants showed tumor remission on 2D but a stable trajectory in whole-tumor volume. Sixteen of 25 participants showed signs of tumor regrowth in either 2D or volumetric trajectories and were considered for scoring the agreement in the SD-PD transition. Eight of these participants did not agree in the transition from SD to PD based on a time to

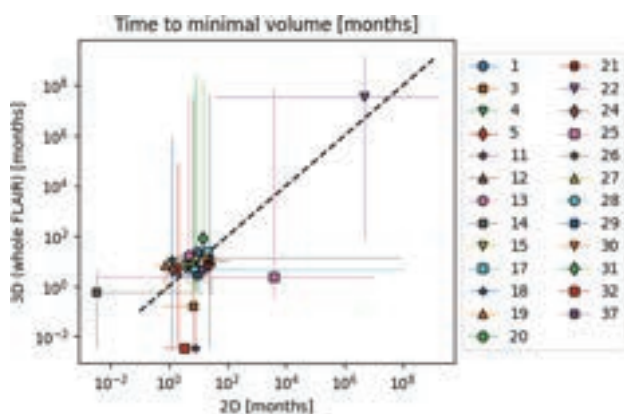


**FIG 2.** ROC curves comparing the performance of 2D area and whole- and solid-volume percent-age change in classifying BT-RADS PR versus non-PR (SD/PD) for 31 participants at the end of treatment or last available scan (A) and time of BT-RADS first-confirmed response or first follow-up scan if no response (B). Vol indicates volume.



**FIG 3.** Examples of agreement (A) and disagreement (B) in 2D and whole-tumor volume trajectories of 2 participants based on model-determined time to best response. 2D measurements and segmentations on 1 representative slice of the whole tumor are shown. Line colors indicate the model-based assignment to either PR (blue), SD (yellow), or PD (red), and background colors reflect BT-RADS labels relative to the pretreatment scan at each follow-up. Model fits are shown with 95% confidence intervals, and the obtained distribution for minimal tumor volume, as medians with 95% confidence intervals are given in the legend.





**FIG 4.** Correlation in model-determined median time to best response with 95% confidence intervals between 2D and whole-tumor volume trajectories for 25 participants. We observed a nonsignificant correlation of the estimated time to best response between 2D and whole-tumor assessments (Spearman rank correlation coefficient = 0.39,  $P$  value = .054).

transition of  $\geq 90$  days in 2D or whole-tumor trajectories. Of these, 2D showed a transition from SD to PD, while the whole volume did not show a transition in 5 participants. Two participants showed a later SD-to-PD transition in 2D.

Given Bonferroni multiple testing correction (number of tests = 7) and the comparably small sample size, the Wilcoxon rank-sum test showed no statistically significant difference in any of the assessed imaging or goodness-of-fit features. The lowest  $P$  value was observed for the ratio of the 2D and whole-volume nRMSE between participants with agreement (median 2D/whole-volume nRMSE = 4.2; full range, 1.1–12.6) and disagreement (median 2D/whole-volume nRMSE = 7.4; full range, 4.7–12.4) in the PR-to-SD transition ( $P = .01$ ). Participants with disagreement in 2D and whole-volume trajectories had a larger observed error in 2D compared with whole-tumor model fits. The best-performing imaging feature was the mean slice thickness ( $P = .04$  for PR-SD agreement), indicating a trend in which a larger slice thickness may lead to disagreeing trends (median FLAIR slice thickness in agreement group = 4.0 with full range, 1.0–5.0; median in disagreement group = 4.3 with full range, 4.0–4.9).

## DISCUSSION

The main findings of this study are the following: 1) a moderate congruence between BT-RADS and both 2D and whole-volume assessments using standard and volume-extrapolated RAPNO thresholds, respectively; 2) a significant difference in the time to response between BT-RADS and volumetrics using volume-extrapolated RAPNO thresholds; 3) no significant difference in the AUC between 2D and volumetric ROC curves for BT-RADS PR classification; and 4) a strong variation in model-based estimation of the time to best response between 2D and whole-volume trajectories. Among a subset of 25 participants, 32% had a difference in time to transition from PR to SD of  $\geq 90$  days between 2D and whole-volume trajectories and 50% had a difference of  $\geq 90$  days in the time to transition from SD to PD.

The results contrast with our hypothesis in that volumetrics did not perform significantly better than 2D in the classification

of response based on BT-RADS, our predefined reference standard in this study. One reason may be insufficient statistical power to detect a significant difference, given our limited sample size. Another reason may be that 2D measurements were performed on the sequence/plane that best represented tumor burden after a multiplanar assessment of tumor response. Subtle changes in growth dynamics on multiple planes may be better represented in a carefully chosen 2D plane, whereas the nature of 3D might introduce volume averaging of subtle findings that diminish sensitivity. On the other hand, selecting a 2D plane that represents subtle changes well relies on the precision of the radiologist. Moreover, the heterogeneity of tumors in the participant cohort, which included both low- and high-grade gliomas, may have contributed to the lack of a significant difference seen among volumetrics and 2D in detecting a partial response. Further investigation into the role of volumetrics based on tumor imaging characteristics, location, and subtype may reveal differences among response methods. Volumetric segmentation accuracy may have been affected by thicker image slices because isotropic 1-mm slice imaging protocols for several participants were unavailable. Given the low number of participants who had PD according to BT-RADS criteria, we could not assess the relative performance of 2D versus volumetrics in classifying PD.

Longitudinal modeling showed different trajectories in 2D and whole volume in several participants based on the model-estimated time to best response. While visualization provides a qualitative motivation for this observation, we quantified this agreement in terms of the agreement to transition in labels, either PR to SD (32% disagreement) or SD to PD (50% disagreement). Four participants had a stable whole-volume trajectory throughout treatment, while 2D showed changes in response labels in these same 4 participants. Three other participants showed tumor response followed by SD in both whole and 2D trajectories. However, these same 3 participants showed tumor regrowth following SD in 2D but not whole-volume trajectories. These findings are consistent with those from Ellingson et al<sup>11</sup> showing that volumetrics had more stable growth rates than 2D measurements in adult low-grade gliomas. Furthermore, we found that participants with disagreement in 2D-versus-whole-volume trajectories had a higher error in 2D than in the volumetric model fit, which was taken as a proxy for measurement uncertainty.

Taken together, our findings suggest that variability and uncertainty in 2D measurements may contribute to more fluctuations in 2D trajectories compared with volumetrics. This finding may lead to unwarranted changes in clinical management, making it essential to confirm 2D changes with volumetric ones. On the basis of preliminary observations from our study data, we hypothesize that fluctuations observed in 2D but not volumetric trajectories may be due to the following reasons: 1) Nonuniform tumor changes may not be fully captured by 2D measurements, leading to falsely classifying a tumor as having a response or progression based on a single plane; 2) variation in slice thickness between follow-up images may lead to inconsistency in the location of 2D measurement planes; and 3) in tumors with diffuse borders, 2D measurements may be more prone to measurement fluctuations than volumetrics. Thus, we recommend the use of volumetrics, particularly in pediatric

gliomas with diffuse borders or nonuniform changes occurring on multiple planes.

Our study has several strengths. While previous studies on volumetric assessment of gliomas<sup>6,11</sup> have compared 2D and volumetrics using existing literature thresholds, our study compared 2D and volumetric change with the visual assessment of the response of neuroradiologists based on the structured BT-RADS criteria, which are closely correlated with clinical management. We found that most participants had a partial response when assessed by BT-RADS but SD based on 2D and volumetric RAPNO criteria. Furthermore, while there have been other studies that tracked 2D-versus-volumetric longitudinal treatment response in adult and pediatric low-grade gliomas,<sup>11,18</sup> our study used a mathematic model that accounts for both treatment response and resistance<sup>16</sup> when tracking 2D and volumetric changes. We also investigated how imaging protocols and measurement uncertainty are related to agreement between 2D and volumetric trajectories. The choice of 2D or volumetrics in clinical practice may depend on tumor characteristics such as heterogeneity and shape, which can affect the reliability of 2D and volumetric trajectories, and both trajectories should be considered in decision-making.

Our study also had limitations. The main ones were the small participant sample size and inherent interrater variability associated with manual volumetric segmentation,<sup>19</sup> for which we tried to account by assuming a 10% contouring uncertainty in our longitudinal models. Interrater differences in 2D measurements, particularly for more irregularly shaped and multifocal tumors, may also affect the comparison of 2D and volumetrics. A second limitation was the unavailability of functional treatment outcomes to correlate with radiographic response assessment. Instead, we used the best reference standard we had available (ie, neuroradiologist assessment based on BT-RADS) to compare 2D and volumetric methods, though this falls short of the criterion standard of clinical outcomes. BT-RADS criteria are holistic and structured, with scores corresponding to recommendations for clinical monitoring and management, which can be more useful than RAPNO criteria in the clinical setting.

Finally, we assigned a BT-RADS score at each timepoint compared with the pretreatment image. A future avenue of research would be to correlate longitudinal 2D and volumetric trends to trends in BT-RADS scores, which are assigned at each timepoint compared with the most recent prior scans instead of the pretreatment image alone. Future research could also aim to quantify interrater variability in volumetric measurements, which can be incorporated into longitudinal volumetric trajectories. With the advent of pediatric glioma-segmentation algorithms, volumetrics may be more efficiently incorporated into clinical practice and research.<sup>20</sup> Finally, prospective research in larger pediatric glioma cohorts is necessary to determine whether volumetrics or 2D better predicts clinical outcomes.

## CONCLUSIONS

Our study compared the 2D and volumetric response assessment with the reference standard of neuroradiologists' visual assessments using BT-RADS in *BRAF* V600E-mutant pediatric gliomas from the PNOC-002 trial. While our study did not detect a significant difference between 2D and whole-volume change in

classifying the BT-RADS partial response either at the end of treatment or at the first timepoint of BT-RADS partial response detection, we found a significant difference in the time to first response between BT-RADS and volumetrics using volume-extrapolated RAPNO criteria. Moreover, we did not find a significant correlation in the time to best response between longitudinal models of 2D and volumetric trajectories, suggesting that trends in individual participant responses may greatly differ between 2D and volumetrics. Because several participants had large differences in the time to transition between response classifications based on 2D and volumetric trajectories, we suggest that volumetrics be incorporated into clinical practice because the additional information provided may influence treatment monitoring and management.

## ACKNOWLEDGMENTS

We would like to thank all Pacific Pediatric Neuro-Oncology Consortium investigators/staff, Schuyler Stoller and Aubrie Drechsler, for help in image transfer, and all trial participants and their families.






Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

1. Adel Fahmideh M, Scheurer ME. **Pediatric brain tumors: descriptive epidemiology, risk factors, and future directions.** *Cancer epidemiology. Cancer Epidemiol Biomarkers Prev* 2021;30:813–21 CrossRef Medline
2. Ryall S, Zapotocky M, Fukuoka K, et al. **Integrated molecular and clinical analysis of 1,000 pediatric low-grade gliomas.** *Cancer Cell* 2020;37:569–83.e5 CrossRef Medline
3. Lhermitte B, Wolf T, Chenard MP, et al. **Molecular heterogeneity in BRAF-mutant gliomas: diagnostic, prognostic, and therapeutic implications.** *Cancers (Basel)* 2023;15:1268 CrossRef Medline
4. D'Arco F, Culleton S, De Cocker L, et al. **Current concepts in radiologic assessment of pediatric brain tumors during treatment, part 1.** *Pediatr Radiol* 2018;48:1833–43 CrossRef Medline
5. Fangusaro J, Witt O, Hernáiz Driever P, et al. **Response assessment in paediatric low-grade glioma: recommendations from the Response Assessment in Pediatric Neuro-Oncology (RAPNO) working group.** *Lancet Oncol* 2020;21:e305–16 CrossRef Medline
6. D'Arco F, O'Hare P, Dashti F, et al. **Volumetric assessment of tumor size changes in pediatric low-grade gliomas: feasibility and comparison with linear measurements.** *Neuroradiology* 2018;60:427–36 CrossRef Medline
7. Chappell R, Miranpuri SS, Mehta MP. **Dimension in defining tumor response.** *J Clin Oncol* 1998;16:1234 CrossRef Medline
8. Gahrman R, van den Bent M, van der Holt B, et al. **Comparison of 2D (RANO) and volumetric methods for assessment of recurrent glioblastoma treated with bevacizumab—a report from the BELOB trial.** *Neuro Oncol* 2017;19:853–61 CrossRef Medline
9. Gahrman R, Smits M, Vernhout RM, et al. **The impact of different volumetric thresholds to determine progressive disease in patients with recurrent glioblastoma treated with bevacizumab.** *Neurooncol Adv* 2022;4:vdac032 CrossRef Medline
10. Huang RY, Unadkat P, Bi WL, et al. **Response assessment of meningioma: 1D, 2D, and volumetric criteria for treatment response and tumor progression.** *Neuro Oncol* 2019;21:234–41 CrossRef Medline
11. Ellingson BM, Kim GH, Brown M, et al. **Volumetric measurements are preferred in the evaluation of mutant IDH inhibition in non-enhancing diffuse gliomas: evidence from a phase I trial of ivosidenib.** *Neuro Oncol* 2022;24:770–78 CrossRef Medline

12. Nicolaides T, Nazemi KJ, Crawford J, et al. **Phase I study of vemurafenib in children with recurrent or progressive BRAFV600E mutant brain tumors: Pacific Pediatric Neuro-Oncology Consortium study (PNOC-002).** *Oncotarget* 2020;11:1942–52 CrossRef Medline
13. Weinberg BD, Gore A, Shu HK, et al. **Management-based structured reporting of posttreatment glioma response with the Brain Tumor Reporting and Data System.** *J Am Coll Radiol* 2018;15:767–71 CrossRef Medline
14. Aboian M, Bousabarah K, Kazarian E, et al. **Clinical implementation of artificial intelligence in neuroradiology with development of a novel workflow-efficient picture archiving and communication system-based automated brain tumor segmentation and radiomic feature extraction.** *Front Neurosci* 2022;16:860208 CrossRef Medline
15. Robin X, Turck N, Hainard A, et al. **pROC: an open-source package for R and S+ to analyze and compare ROC curves.** *BMC Bioinformatics* 2011;12:77 CrossRef Medline
16. Brüningk SC, Peacock J, Whelan CJ, et al. **Intermittent radiotherapy as alternative treatment for recurrent high grade glioma: a modeling study based on longitudinal tumor measurements.** *Sci Rep* 2021;11:20219 CrossRef Medline
17. Glazar DJ, Grass GD, Arrington JA, et al. **Tumor volume dynamics as an early biomarker for patient-specific evolution of resistance and progression in recurrent high-grade glioma.** *J Clin Med* 2020;9:2019 CrossRef
18. Tsai JW, Choi JJ, Ouaalam H, et al. **Integrated response analysis of pediatric low-grade gliomas during and after targeted therapy treatment.** *Neurooncol Adv* 2023;5:vdac182 CrossRef Medline
19. Bø HK, Solheim O, Jakola AS, et al. **Intra-rater variability in low-grade glioma segmentation.** *J Neurooncol* 2017;131:393–402 CrossRef Medline
20. Kazerooni AF, Khalili N, Liu X, et al. **The Brain Tumor Segmentation (BraTS) Challenge 2023: Focus on Pediatrics (CBTN-CONNECT-DIPGR-ASNR-MICCAI BraTS-PEDs).** *ArXiv* 2024 Jan 4. [Epub ahead of Print] arXiv:2305.17033v2 Medline

# Expanding the Imaging Spectrum of Polymorphous Low-Grade Neuroepithelial Tumor of the Young in Children

 Carmen Cerron-Vela,  Fabricio Guimarães Gonçalves,  Angela N. Viaene,  Luis Octavio Tierradentro-García, and  Arastoo Vossough



## ABSTRACT

**SUMMARY:** Polymorphous low-grade neuroepithelial tumors of the young (PLNTY) are rare brain tumors first described in 2017 and recently included in the 2021 5th World Health Organization Classification of Tumors of the Central Nervous System. They typically affect children and young adults. Few pediatric cases have been reported in the literature. The most common imaging features described, include location within the temporal lobe, involvement of the cortical/subcortical region, coarse calcifications, and well-defined margins with solid and cystic morphology, with slight-or-no enhancement. However, there is limited information on imaging features in children. We present the imaging spectrum of neuroimaging features in a series of pediatric patients with a histologically and molecularly proved PLNTY diagnosis. Coarse calcifications are uncommon in children compared with the adult literature, and they may develop with time. The transmantle-like sign can be observed, and adjacent cortical dysplasia may be seen. Seizure recurrence may occur despite gross total resection of the tumor.

**ABBREVIATIONS:** DNET = dysembryoplastic neuroepithelial tumor; IQR = interquartile range; PLNTY = polymorphous low-grade neuroepithelial tumor of the young; TLS = transmantle-like sign; WHO5 = World Health Organization Classification of Tumors of the Central Nervous System, 5th edition

Polymorphous low-grade neuroepithelial tumor of the young (PLNTY) was described in 2017 by Huse et al<sup>1</sup> as an epileptogenic tumor with infiltrative growth, oligodendroglioma-like components, frequent calcification, strong diffuse aberrant CD34 expression, and genetic alterations in the mitogen-activated kinase pathway, including *FGFR2*, *FGFR3*, and *BRAF*.<sup>1-3</sup>

PLNTY is 1 of the 6 newly recognized pediatric low-grade gliomas and glioneuronal tumors in the 2021 5th edition of the World Health Organization Classification of Tumors of the Central Nervous System (WHO5).<sup>4</sup> It primarily affects children and young adults<sup>2</sup> but can also be seen in individuals older than 30 years of age.<sup>5,6</sup>

Few cases have been reported, including adult and pediatric cases.<sup>7,8</sup> The imaging features of this entity have been shown to be diverse,<sup>3,9</sup> and calcifications are recognized as an important imaging characteristic of PLNTY.<sup>8,10</sup> However, there is still scarce information on the imaging features in children.

PLNTY has a benign clinical course with few exceptions.<sup>8,11</sup> Correct identification of PLNTY can affect patient treatment and

prognosis by directing surgical management and medical treatment. We aimed to describe the neuroimaging features in a series of pediatric patients with a histologic and molecularly-proved PLNTY diagnosis.

## CASE SERIES

This single-center, retrospective study was reviewed and approved by our institutional review board and complied with the Health Insurance Portability and Accountability Act. We obtained a waiver of informed consent. We searched our databases (mPower by Nuance Communications, <https://www.nuance.com/healthcare/diagnostics-solutions/radiology-performance-analytics/mpower-clinical-analytics.html>), and Illuminate InSight, <https://www.illuminate.ai/illuminate/solutions/insight>) for pathology-proved PLNTY cases. We extracted demographic data from an electronic chart system (Epic Systems, <https://www.forbes.com/companies/epic-systems/?sh=627bbf05ca11>). Inclusion criteria were 0–18 years of age, availability of preoperative brain MR imaging, and a molecular diagnosis based on immunohistochemical or genetic analysis.

We evaluated 10 patients. Five (50%) were female. The median age at diagnosis was 9 years (range, 0–14 years). The most common symptoms at presentation were seizures ( $n = 8$ , 80%). The median duration of symptoms was 32 months (interquartile range [IQR], 5.75–85 months). All tumors were surgically resected. The median time interval between MR imaging and

Received October 17, 2023; accepted after revision December 16.

From the Departments of Radiology (C.C.-V., F.G.G., L.O.T.-G., A.V.) and Pathology and Laboratory Medicine (A.N.V.), Children's Hospital of Philadelphia, Philadelphia, Pennsylvania.

Please address correspondence to Carmen Cerron-Vela, MD, Department of Radiology, The Children's Hospital of Philadelphia, 3401 Civic Center Blvd, Philadelphia, PA 19104; e-mail: [cerronvelc@chop.edu](mailto:cerronvelc@chop.edu)



Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8145>



surgery was 9.5 months (IQR, 2.75–59.5 months). Recurrence of seizure symptoms was seen in 3 patients (30%), who required surgical reintervention. Demographics, clinical data, and pathologic features of this series are summarized in the Online Supplemental Data.

### Imaging

Patients were referred from different health centers, and initial imaging protocols were not fully standardized. The patients had undergone either 1.5T or 3T brain MR imaging (Various scanners; Siemens). Sequences performed included sagittal and axial T1WI, axial T2WI, T2 FLAIR, DWI, and gadolinium-enhanced postcontrast T1WI. Gradient-echo, SWI, and arterial spin-labeling sequences were variably obtained depending on the study and scanner. All imaging from our institution included 0.9- to 1.0-mm isotropic T1-weighted MPRAGE and high-resolution 2-mm (no gap) T2-weighted and 3-mm T2 FLAIR images in the axial and coronal planes.

Two pediatric neuroradiologists systematically evaluated all imaging in consensus to assess the following: tumor location, margins, and morphology; T1WI, T2WI, susceptibility imaging, T2 FLAIR signal; and diffusion characteristics and contrast enhancement. If available, CT attenuation was also evaluated.

All tumors were supratentorial. Most tumors involved the temporal lobe ( $n = 7$ , 70%) with a mixed cortical-subcortical location ( $n = 8$ , 80%). The most common tumor morphology was a mixed solid and cystic pattern ( $n = 8$ , 80%). The median cyst size was 5.5 mm (IQR, 3.2–8.7 mm). The largest measured tumor was 17.2 mm. Most tumors had no ( $n = 6$ , 60%) or mild ( $n = 4$ , 40%) contrast enhancement. Margins were operationally defined on T2WI as a clear demarcation between the tumor and the surrounding brain tissue for most of the circumference of the tumor margin. Nine (90%) tumors had ill-defined margins. No restricted (reduced) diffusion was seen in the tumors. Calcifications were determined by a combination of T1-weighted, T2-weighted, SWI with unwrapped phase images, gradient recalled-echo, and CT on initial and follow-up imaging. Two (20%) cases had initial CT. Eight tumors (80%) had visible calcifications, predominately with a punctate pattern ( $n = 5$ , 50%) (Figure and Table). One case (case 4) showed a gradual development of a large central calcification during 10 years (Online Supplemental Data). A transmantle-like sign (TLS) was defined as a linear T2 or T2 FLAIR hyperintense signal in the white matter from the mass that tapers toward the lateral ventricles. This finding was seen in 4 (40%) patients (Fig 1J, -L). When available, a mixed pattern of low or only mildly elevated perfusion was seen on arterial spin-labeling. Imaging findings are summarized in the Online Supplemental Data.

### Pathologic Features

Tumors diagnosed as PLNTY were reviewed by a board-certified neuropathologist, blinded to clinical and radiologic data. The evaluation included a review of both hematoxylin & eosin-stained slides and immunohistochemical stains. In addition, next-generation sequencing was performed on all tumors. Tumors that fulfilled the WHO5 “essential criteria” for PLNTY<sup>4</sup> were selected for inclusion. The WHO criteria are summarized in

the Online Supplemental Data. Two additional “desirable criteria” (calcifications and absence of 1p/19q codeletion)<sup>4</sup> and the presence of adjacent cortical dysplasia on histology were also assessed.

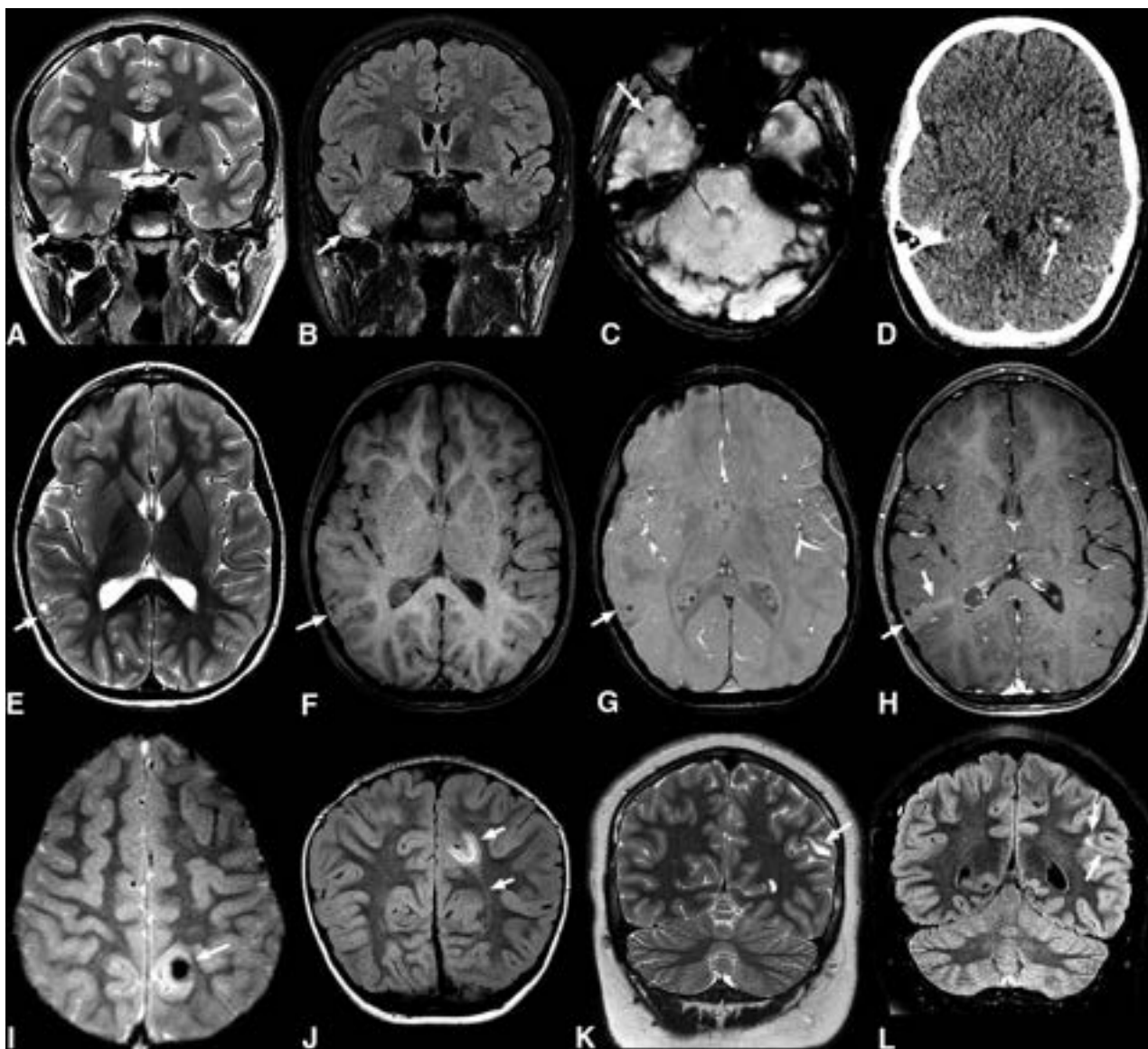
All tumors met all the WHO5 essential criteria as well as the desirable criterion of lacking 1p/19 codeletion. Seven cases (70%) showed calcifications on histology. All tumors showed at least focal diffuse growth and oligodendroglial-like components. All tumors showed areas of positive staining for glial fibrillary acidic protein and/or the oligodendrocyte transcription factor. Intense, diffuse CD34 staining was present in all cases, as well as regions of ramified staining; and the Ki-67 proliferation indices ranged from <1% to 3%–5%. Mitotic activity was typically absent, though 1 tumor demonstrated 3 mitoses per 10 high-power fields. Necrosis and microvascular proliferation were not seen in any case. Calcifications ranged from diffuse to sparse or absent. Histologic features of cortical dysplasia were present in 3 cases (30%). We had 5 cases of *FGFR2* fusions and 5 cases of *BRAF* p.V600E mutations. Pathologic findings are shown in the Online Supplemental Data.

### DISCUSSION

To our knowledge, this is the largest series of imaging features of the rare tumor entity PLNTY in children. In our study, all tumors were in the supratentorial compartment, and most were in the temporal lobe in a cortical-subcortical location, in keeping with prior literature.<sup>5,7,9</sup> Kurokawa et al,<sup>9</sup> in a systematic review, reported 2 main imaging patterns for PLNTY: 1) a well-circumscribed, solid, and cystic mass with scarce mass effect surrounded by normal parenchyma; and 2) an ill-defined, small mass with tiny cysts and TLS. Although the first pattern was reported as the most common in that article, in this study, the second pattern was predominant.

Prominent coarse and central calcifications have been suggested as characteristic features of PLNTY.<sup>3</sup> However, extensive calcifications were not a common imaging feature in this series. Calcifications showed, predominately, a punctate pattern, also called “salt and pepper pattern.”<sup>12</sup> Despite these differences, 1 case gradually developed a sizeable central calcification during a follow-up period of 10 years, exhibiting the classic prominent central calcification. This finding suggests that there may be an evolving imaging appearance of these tumors, in which small calcifications develop in younger patients and grow or coalesce with time to become dense coarse calcifications in older patients. Due to the small number of cases, we cannot assess a correlation between the age of presentation or the duration of symptoms with the presence or development of prominent central calcifications. For example, our case 6 with a 4-year history of seizures had a prominent calcification on the initial MR imaging, while our case 10 did not demonstrate calcifications on the initial MR imaging or develop calcifications during 7 years.

Kurokawa et al<sup>9</sup> associated TLS with a longer duration of symptoms and found that cortical dysplasia was more frequent in patients with TLS. Other reports have also described this association.<sup>10,13</sup> In our study, the duration of seizures for cases that showed TLS was variable, and pathology-proved cortical dysplasia in association with TLS was found in only 25% (1 of 4) of



**FIGURE.** Common imaging features of pediatric PLNTY tumors. A–C, A 9-year-old boy (case 5) with a 1-month history of headaches. Coronal T2WI (A) and FLAIR (B) sequences show a vague area of increased signal intensity in the cortical-subcortical region of the right temporal lobe (arrows in A and B). Axial SWI shows a small calcification (C). D, An 11-year-old girl with seizures (case 1). Nonenhanced axial CT shows a focal lesion in the left temporal lobe with intralésional calcifications exhibiting a punctate calcification pattern (arrow). E–H, A 6-year-old girl with complex focal seizures (case 3). Axial T2WI (E) image depicts a mixed cortical-subcortical lesion with solid and cystic components in the right parietal lobe. Pre- and postcontrast axial T1 TSE (F and H) reveals mild postcontrast enhancement (arrows in E and F). Axial SWI (G) demonstrates a punctate pattern of calcification (arrow). I and J, A 4-year-old girl with a 4-year history of seizures (case 6). Axial T2\* gradient-echo image (I) reveals an ill-defined lesion in the left parietal lobe with a large central susceptibility in keeping with a chunky central calcification (arrow). Coronal FLAIR image (J) shows a related high-signal-intensity area in the subcortical white matter extending toward the left lateral ventricle, consistent with a TLS (arrows). K and L, A 14-year-old boy with new-onset seizures (case 8). Coronal T2WI (K) shows a hyperintense lesion in the left parietal region (arrow). Coronal FLAIR (L) shows an associated high signal area extending medially, in keeping with a TLS (arrows).

patients. These findings may be due to pathologic sampling confounding. It is also possible that cortical dysplasia may be seen without a clear TLS, or alternatively, not all TLSs are truly indicative of cortical dysplasia. The proximity of temporal lobe tumors to the margins of the ventricles may also cause difficulties in appreciating the TLS. A transmantle appearance may cause diagnostic confusion on imaging, with the possibility of assigning simple focal cortical dysplasia type IIb as the diagnosis, or in cases in which a mass is readily apparent, the tumor may be diagnosed

as dysembryoplastic neuroepithelial tumor (DNET). However, the features of PLNTY as described here can indicate a correct diagnosis or a more targeted differential consideration of PLNTY in children. Close attention to the presence of small tumoral features on multiple planes can prevent this potential misdiagnosis.

Isolated reports describe PLNTY cases with difficult therapeutic control or disease recurrence.<sup>11,14</sup> Armocida et al<sup>8</sup> found that contrast enhancement in PLNTY was associated with recurrence or poor control of epileptic symptoms. We found a recurrence of

## Distribution of imaging and pathologic features

	Percentage, Number
Location	
Temporal lobe	70%, <i>n</i> = 7
Parietal lobe	20%, <i>n</i> = 2
Occipital lobe	10%, <i>n</i> = 1
Cortical-subcortical	80%, <i>n</i> = 8
Subcortical only	20%, <i>n</i> = 2
Morphology	
Mixed solid and cystic	80%, <i>n</i> = 8
Solid	10%, <i>n</i> = 1
Cystic	10%, <i>n</i> = 1
Calcifications	
Punctate	50%, <i>n</i> = 5
Chunky	30%, <i>n</i> = 3
No definite Ca <sup>2+</sup> on imaging	20%, <i>n</i> = 2
Margins	
Ill-defined	90%, <i>n</i> = 9
Well-defined	10%, <i>n</i> = 1
Contrast enhancement	
Mild	40%, <i>n</i> = 4
None	60%, <i>n</i> = 6
TLS	
Yes	40%, <i>n</i> = 4
None	60%, <i>n</i> = 6
Pathology proven cortical dysplasia	30%, <i>n</i> = 3
Pathology proven calcification	70%, <i>n</i> = 7
Mutation	
<i>FGFR2</i> alteration	50%, <i>n</i> = 5
<i>BRAF</i> mutation	50%, <i>n</i> = 5

seizures after surgery in 3 cases; none showed postcontrast enhancement. We suspect that the presence of contrast enhancement is not necessarily the main predictor of poor seizure control.

Nearly all the PLNTYs have either *BRAF* p.V600E mutations or fusion events involving *FGFR2/FGFR3*.<sup>1</sup> The *BRAF* p.V600E mutation is the most common.<sup>5</sup> We had an equal proportion of *BRAF* p.V600E mutations and *FGFR2* fusions. There are no clear associations between genetic mutations and clinical outcomes in patients with PLNTY.<sup>8</sup> We found that recurrence of symptoms after surgery occurred similarly in patients with both mutations. The *FGFR3-TACC3* fusion gene in PLNTY has shown malignant transformation in some reports.<sup>15</sup> This mutation was not found in our series.

The differential diagnosis of PLNTY in children includes DNET, ganglioglioma, and pleomorphic xanthoastrocytoma, all of which are commonly seen in the temporal lobe and can commonly present with seizures.<sup>6,9,10</sup> Cystic changes and calcifications have been reported to be more common in PLNTY than in gangliogliomas,<sup>16</sup> though we do not find that this distinction helped guide a differential diagnosis. Nevertheless, larger unilocular cysts are not typical of PLNTY and were not seen in our series. Vasogenic edema is uncommon in all these tumors, and they may be seen in cortical/subcortical locations. PLNTY may show mild or no contrast enhancement, differing from the typical moderate-to-marked enhancement in parts of ganglioglioma or pleomorphic xanthoastrocytoma. Intratumoral T2 FLAIR hyperintense septations with multiple cystlike central T2 FLAIR hyperintense areas (FLAIR rim sign) are much more frequent in DNET.<sup>9</sup> Both DNET and ganglioglioma may have associated

cortical dysplasia on histopathology. Neurocysticercosis also has cystic areas with punctate calcifications but typically does not have a clear associated mass.

## CONCLUSIONS

Unlike previous reports, which have mostly been documented in adult cases, large, dense calcifications and well-defined margins are not commonly seen in pediatric cases. We also suggest that these tumors may have a slowly evolving imaging appearance. PLNTY can show associated cortical dysplasia on pathology, not always associated with a TLS. Seizure recurrence may occur despite gross total resection.

**Disclosure forms** provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

- Huse JT, Snuderl M, Jones DTW, et al. **Polymorphous low-grade neuroepithelial tumor of the young (PLNTY): an epileptogenic neoplasm with oligodendroglioma-like components, aberrant CD34 expression, and genetic alterations involving the MAP kinase pathway.** *Acta Neuropathol* 2017;133:417–29 CrossRef Medline
- Ida CM, Johnson DR, Nair AA, et al. **Polymorphous low-grade neuroepithelial tumor of the young (PLNTY): molecular profiling confirms frequent MAPK pathway activation.** *J Neuropathol Exp Neurol* 2021;80:821–29 CrossRef Medline
- Johnson DR, Giannini C, Jenkins RB, et al. **Plenty of calcification: imaging characterization of polymorphous low-grade neuroepithelial tumor of the young.** *Neuroradiology* 2019;61:1327–32 CrossRef Medline
- Louis DN, Perry A, Wesseling P, et al. **The 2021 WHO Classification of Tumors of the Central Nervous System: a summary.** *Neuro Oncol* 2021;23:1231–51 CrossRef Medline
- Fei X, Zhao J, Wei W, et al. **Clinical, radiological, pathological features and seizure outcome with surgical management of polymorphous low-grade neuroepithelial tumor of the young associated with epilepsy.** *Front Oncol* 2022;12:863373 CrossRef Medline
- Benson JC, Summerfield D, Carr C, et al. **Polymorphous low-grade neuroepithelial tumor of the young as a partially calcified intra-axial mass in an adult.** *AJNR Am J Neuroradiol* 2020;41:573–78 CrossRef Medline
- Bale TA, Rosenblum MK. **The 2021 WHO Classification of Tumors of the Central Nervous System: an update on pediatric low-grade gliomas and glioneuronal tumors.** *Brain Pathol* 2022;32:e13060 CrossRef Medline
- Armocida D, Berra LV, Frati A, et al. **Radiological and surgical aspects of polymorphous low-grade neuroepithelial tumor of the young (PLNTY).** *Acta Neurol Belg* 2023;123:327–40 CrossRef Medline
- Kurokawa M, Kurokawa R, Capizzano AA, et al. **Neuroradiological features of the polymorphous low-grade neuroepithelial tumor of the young: five new cases with a systematic review of the literature.** *Neuroradiology* 2022;64:1255–64 CrossRef Medline
- Chen Y, Tian T, Guo X, et al. **Polymorphous low-grade neuroepithelial tumor of the young: case report and review focus on the radiological features and genetic alterations.** *BMC Neurol* 2020;20:123 CrossRef Medline
- Bitar M, Danish SF, Rosenblum MK. **A newly diagnosed case of polymorphous low-grade neuroepithelial tumor of the young.** *Clin Neuropathol* 2018;37:178–81 CrossRef Medline
- Paredes-Aragón EM, AlKhaldi NA, Burneo JG, et al. **Pearls & Oysters: salt and pepper sign, PLNTY for drug-resistant epilepsy.** *Neurology* 2023;100:791–95 CrossRef Medline

13. Gupta VR, Giller C, Kolhe R, et al. **Polymorphous low-grade neuroepithelial tumor of the young: a case report with genomic findings.** *World Neurosurg* 2019;132:347–55 CrossRef Medline
14. Bale TA. **FGFR-gene family alterations in low-grade neuroepithelial tumors.** *Acta Neuropathol Commun* 2020;8:21–21 Medline
15. Bale TA, Sait SF, Benhamida J, et al. **Malignant transformation of a polymorphous low grade neuroepithelial tumor of the young (PLNTY).** *Acta Neuropathol* 2021;141:123–25 CrossRef Medline
16. Adachi Y, Yagishita A. **Gangliogliomas: characteristic imaging findings and role in the temporal lobe epilepsy.** *Neuroradiology* 2008;50:829–34 CrossRef Medline



# Prognostic Significance of Central Skull Base Remodeling in Chiari II Malformation

John T. Freiling, Niles K. Desai, Stephen F. Kralik, William E. Whitehead, and Thierry A.G.M. Huisman

## ABSTRACT

**BACKGROUND AND PURPOSE:** Outward convexity of the basiocciput and posterior atlanto-occipital membrane are common in patients with Chiari II malformation associated with an open neural tube defect. We aimed to determine if the severity of these findings correlated with the need for future hydrocephalus treatment.

**MATERIALS AND METHODS:** A retrospective chart and imaging review identified patients who underwent open neural tube defect repair at a quaternary care pediatric hospital from July 2014 through September 2022. Patients were classified by the need for hydrocephalus treatment and whether they received prenatal or postnatal neural tube defect repair. Measurements of imaging parameters related to posterior fossa maldevelopment and skull base remodeling were performed.

**RESULTS:** Compared with 65 patients who did not require hydrocephalus treatment, 74 patients who required treatment demonstrated statistically significantly greater mean basiocciput convexity ( $P < .001$ ). While the mean basiocciput length in the hydrocephalus treatment group was smaller ( $P < .001$ ), the ratio of basiocciput convexity to length was larger ( $P < .001$ ). Notably, 100% of patients with a basiocciput convexity of  $\geq 4$  mm required hydrocephalus treatment. The mean posterior atlanto-occipital membrane convexity was significantly greater for patients who required hydrocephalus treatment in the postnatal group ( $P = .02$ ), but not the prenatal group ( $P = .09$ ).

**CONCLUSIONS:** Pediatric patients with Chiari II malformation who ultimately required surgical hydrocephalus treatment had greater outward convexity of the basiocciput but had greater posterior atlanto-occipital membrane outward convexity only if the repair was performed postnatally. Together these measurements may be useful in predicting the need for hydrocephalus treatment.

**ABBREVIATIONS:** CM2 = Chiari II malformation; FOD = frontal occipital diameter; HT = hydrocephalus treatment; MMC = myelomeningocele; ONTD = open neural tube defect; PAOM = posterior atlanto-occipital membrane; PFD = posterior fossa diameter

Neural tube defects remain common, occurring at an approximate prevalence of 5–7 per 10,000 live births.<sup>1</sup> Neural tube defects have been classified as closed (skin-covered) or open (nonskin-covered), with open neural tube defects (ONTDs) including myelomeningocele (MMC) and myelocele (also known as myeloschisis).<sup>2</sup> ONTDs are typically associated with Chiari II malformation (CM2), and this relationship has been explained by the unified theory of McLone and Knepper,<sup>3</sup> which suggests that the leakage of CSF into the amniotic fluid through the defect leads to inadequate expansion of the posterior fossa embryonic vesicles and consequently characteristic findings of CM2.<sup>4</sup> On the

basis of this current understanding, fetal surgery with intrauterine repair of the ONTD is performed to reduce the degree of CM2 and possibly the postoperative necessity for ventriculoperitoneal shunting.

Hydrocephalus is a common complication of CM2 and frequently requires surgical intervention. ONTDs and CM2 are associated with significant morbidity and mortality,<sup>5</sup> and hydrocephalus seen with CM2 is a significant contributor. Surgical hydrocephalus treatment (HT) has traditionally encompassed ventriculoperitoneal shunting, the complications of which can also contribute to poor outcomes. More recently, endoscopic third ventriculostomy and choroid plexus cauterization are increasingly being used in HT as an alternate strategy to ventricular shunting.<sup>6</sup>

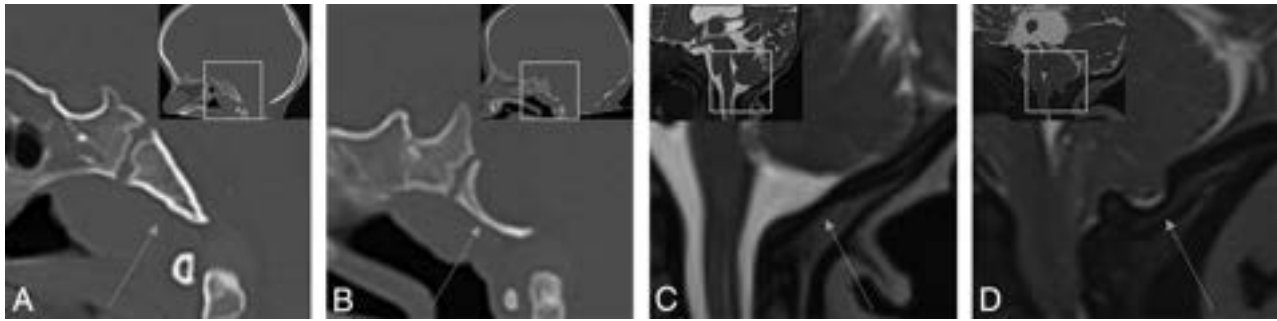
The pathophysiologic mechanisms leading to hydrocephalus in CM2 are complex and incompletely understood, making predicting which patients will need surgical HT difficult. Identifying patients at greatest risk for future hydrocephalus development

Received October 8, 2023; accepted after revision December 11.

From the Edward B. Singleton Department of Radiology (J.T.F., N.K.D., S.F.K., T.A.G.M.H.), Department of Neurosurgery (W.E.W.), Texas Children's Hospital & Baylor College of Medicine, Houston, Texas.

Please address correspondence to John T. Freiling, MD, Edward B. Singleton Department of Radiology, Texas Children's Hospital and Baylor College of Medicine, 6701 Fannin St, Suite 470, Houston, TX 77030; e-mail: jtfreili@texaschildrens.org

<http://dx.doi.org/10.3174/ajnr.A8142>



**FIG 1.** Comparison of normal and remodeled basiocciput and posterior atlanto-occipital membranes. A and B, Sagittal CT images demonstrate the normal appearance of the basiocciput (arrow in A) and outward convexity of the basiocciput (arrow in B). C and D, Sagittal T2-weighted images demonstrate the normal appearance of the PAOM (arrow in C) and outward convexity of the PAOM (arrow in D). Image insets demonstrate larger FOV sagittal images and the location of corresponding zoomed-in FOV.

and the need for HT may help in clinical management and counseling. Previous studies have identified multiple characteristics correlated with the need for future HT. The more cephalad the level of the MMC defect, the higher will be the incidence of ventricular shunting.<sup>7</sup> For patients who underwent prenatal repair of ONTDs, smaller ventricle size and defects located at or below L4 were less likely to require shunting.<sup>8</sup> The Management of Myelomeningocele Study (MOMS) investigators, in a report of the 1-year outcomes from the MOMS trial, demonstrated that larger ventricle size on imaging before prenatal repair was associated with an increased need for future shunting.<sup>9</sup> Zarutskie et al<sup>10</sup> demonstrated the predictive value of multiple imaging parameters with the need for HT in patients who underwent prenatal MMC repair, with the best predictors being presurgery ventricular width, persistent hindbrain herniation postrepair, and growth in ventricular volume.

However, the prognostic value of imaging parameters related to posterior fossa maldevelopment and skull base remodeling in CM2 and their correlation with the need for future HT remain unreported, despite their possible contributing component in hydrocephalus development. Several osseous skull base changes occur with posterior fossa maldevelopment in patients with CM2, including reduced posterior fossa volume, reduced supraocciput and exocciput lengths,<sup>11</sup> a small dorsum sellae,<sup>12</sup> petrous scalloping,<sup>13</sup> as well as hypoplasia and altered morphology of the occipital bone. The severity of scalloping, or outward convexity of the clivus, was found to be significantly greater in patients with MMC and CM2,<sup>11</sup> and in our experience, it is a frequently seen finding. Additionally, outward convexity of the posterior atlanto-occipital membrane (PAOM) is frequently seen in CM2 (Fig 1), though it is of uncertain prognostic significance. The goal of our study was to determine if outward convexity of the basiocciput and PAOM are associated with the need for future HT, which may provide prognostic value as well as possible insight into the mechanisms of hydrocephalus development.

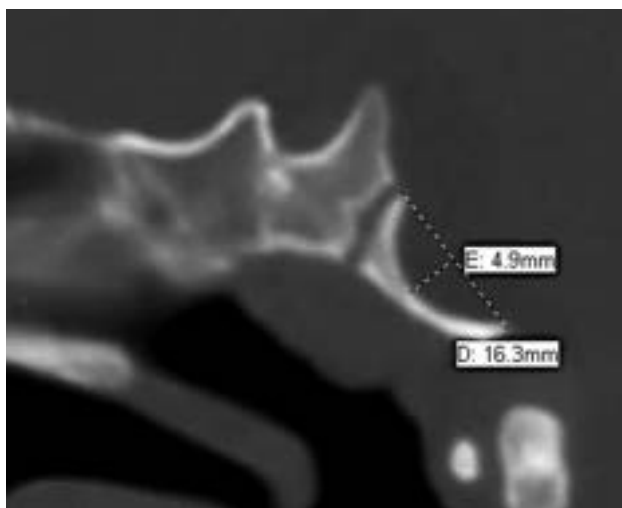
## MATERIALS AND METHODS

The study was approved by the Baylor College of Medicine institutional review board. A retrospective chart and imaging review was performed on initial brain CT and/or MR imaging studies

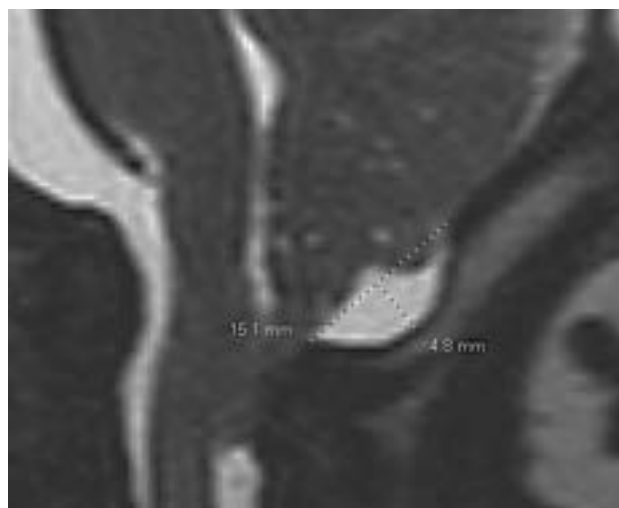
obtained among patients diagnosed with ONTD (MMC or myelocoele) who underwent either in utero or postnatal repair from July 2014 through September 2022 at a large quaternary care pediatric hospital with an integrated dedicated fetal treatment center. Patients were identified from a surgical database and were included if a postnatal head CT or brain MR imaging was performed. Patients were grouped according to those that required neurosurgical HT and those who did not and were additionally stratified according to whether they received prenatal or postnatal surgical repair of their ONTD. The presence or absence of HT was determined by chart review assessing a history of prior ventriculoperitoneal shunting or endoscopic third ventriculostomy/choroid plexus cauterization. For the group that did not require HT, patients were excluded if <12 months of clinical follow-up was available. Additionally, patients were excluded from the analysis if no head CT or MR imaging was available for review before HT, if the performed imaging was nondiagnostic, as well as if they demonstrated severe cerebellar parenchymal volume loss or had a known history of an acute or chronic CNS injury other than the neural tube defect (eg, history of intraventricular hemorrhage or meningitis).

Radiology assessment was performed by a neuroradiologist with additional fellowship training in pediatric neuroradiology. All measurements were performed on the initial/earliest postnatal CT or MR imaging study and before HT. MR imaging sequences were chosen on the basis of the best subjective evaluation of the posterior fossa structures within the limitation of motion and artifacts, and high-resolution fluid-sensitive MR imaging sequences were used when available. Measurements were obtained before knowledge of the future shunt status.

The basiocciput length was measured by placing a line along the dorsal aspect of the basiocciput from the basion to the dorsal spheno-occipital synchondrosis. The severity of basiocciput outward convexity was then measured as the maximum distance to the dorsal basiocciput orthogonal to the basiocciput length line (Fig 2). Because the convexity predominately occurred within the basiocciput with relative sparing of the basisphenoid, the measurements were limited to the convexity of the basiocciput. Distance, rather than angle,<sup>11</sup> was used for reproducibility and simplicity. To control for differences in the total size of the basiocciput affecting the convexity measurements, we compared the ratio of the basiocciput convexity with the basiocciput length.



**FIG 2.** Sagittal CT of the skull base demonstrating outward convexity of the basioccipt in a 10-month-old infant with CM2. Basioccipt length is measured from the basion to the dorsal speno-occipital synchondrosis. Convexity is measured as the maximum distance orthogonal to the *length* line.



**FIG 3.** Sagittal T2-weighted image of the posterior craniocervical junction in a 1-month-old infant with CM2. CI-opisthion length is measured from the posterior C1 arch to the opisthion. Posterior atlanto-occipital membrane convexity is measured as the maximum distance orthogonal to the *length* line.

The severity of the PAOM outward convexity was measured in a similar fashion. First, the C1 opisthion line was drawn from the superior aspect of the posterior C1 arch to the superior aspect of the opisthion. Then, the largest distance from the convexity of the PAOM orthogonal to the distance line was measured (Fig 3).

To estimate posterior fossa size, we measured the posterior fossa diameter (PFD) using the Twining line drawn between the dorsum sellae and the internal occipital protuberance.<sup>14</sup> The PFD was then compared with the frontal occipital diameter (FOD), measured as the largest dimension between the frontal and occipital calvaria in the midsagittal plane.<sup>15</sup> The ratio of these measurements was then calculated (FOD/PFD) to estimate the relative posterior fossa size, using linear measurements for simplicity and ease of measurement.<sup>15,16</sup>

To assess the severity of ventriculomegaly, we measured the ventricular width with a line drawn at the maximum width of the atria of the lateral ventricles in the coronal plane, using the widely accepted method described for measurement of fetal ventriculomegaly.<sup>17</sup>

Statistical comparisons between the 2 groups were performed using an unpaired Student *t* test with GraphPad Prism software (Version 9.4.1, GraphPad Software) and Excel Microsoft (Office Professional Plus 2016). Values of  $P < .05$  were considered statistically significant.

## RESULTS

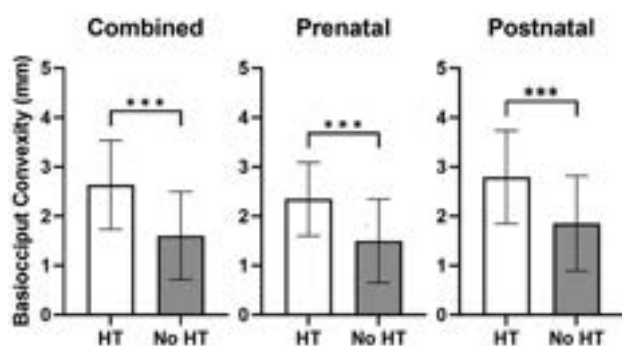
A total of 318 patients underwent surgical ONTD repair; 168 patients had an available head CT or brain MR imaging and were included in the study. Of these patients, there were 29 total patients excluded from the study: 17 were excluded for lack of a head CT or brain MR imaging before HT, and 12 were excluded for nondiagnostic imaging or if they demonstrated severe cerebellar parenchymal volume loss or had a known history of an acute or chronic CNS injury other than the neural tube defect. A total of 139 patients were identified for study analysis, of which 70

patients underwent prenatal ONTD repair and 69 patients underwent postnatal repair. Ultimately, 74 of 139 patients required HT. The mean gestational age at birth was 37.1 (SD, 2.8) weeks for patients who required HT and 37.2 (SD, 3.2) weeks for patients who did not require HT ( $P = .82$ ). The mean age at the time of the first head CT or MR imaging was 3.2 (SD, 3.9) months (minimum = 1 month, maximum = 20 months) in patients who required HT, and 12.1 (SD, 8.3) months (minimum = 1 month, maximum = 35 months) for patients who did not require HT ( $P < .001$ ). For patients who required HT, the mean age at the time of HT was 4.7 (SD, 6.9) months (minimum = 1 month, maximum = 44 months). For patients who did not require HT, the mean clinical follow-up was 39.5 (SD, 22) months (minimum = 12 months, maximum = 103 months).

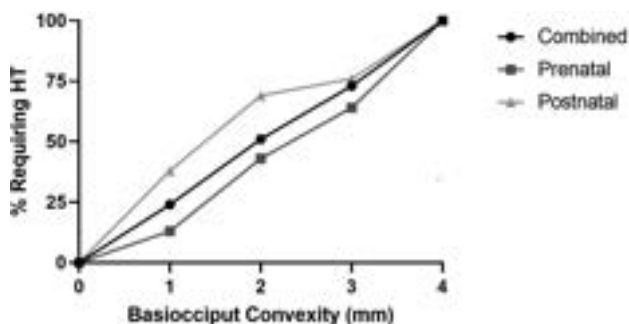
The mean basioccipt outward convexity on the first diagnostic CT or MR imaging was significantly greater in patients who ultimately received HT compared with patients in whom HT was not required in both the prenatal (2.3 mm; range, 1–4 mm compared with 1.5 mm; range, 0–3 mm;  $P < .001$ ) and postnatal (2.8 mm; range, 1–4 mm compared with 1.9 mm; range, 0–3 mm;  $P < .001$ ) repair groups, as well as both groups combined (2.6 mm; range, 1–4 mm compared with 1.6 mm; range, 0–3;  $P < .001$ ) (Fig 4). The percentage of patients requiring HT increased with increasing severity of the basioccipt convexity. Notably, 100% of patients with basioccipt convexity of  $\geq 4$  mm required HT in our cohort (Fig 5).

Compared with patients who did not require HT, patients who required HT demonstrated significantly greater mean basioccipt convexity-to-length ratio ( $P < .001$ ) and FOD/PFD ratio ( $P < .001$ ), as well as significantly less mean basioccipt length (prenatal,  $P < .001$ ; postnatal,  $P = .002$ ), and mean PFD ( $P < .001$ ) (Table 1).

The mean PAOM convexity was significantly greater for patients who required HT in the postnatal group ( $P = .02$ ), but not the prenatal group ( $P = .09$ ). The mean lateral ventricle width



**FIG 4.** Basiocciput convexity is increased in patients requiring HT. Basiocciput convexity (millimeters) was measured in all patients requiring HT and not requiring HT. Patients were then stratified by prenatal or postnatal repair of the ONTD and basiocciput convexities and were compared. Bars represent the mean (SD) with triple asterisks  $P < .001$  representing statistical significance as determined by an unpaired Student  $t$  test.



**FIG 5.** An increase in basiocciput convexity was associated with a higher percentage of patients requiring HT. The percentage of patients requiring HT at varying basiocciput convexities (1–4+ mm) was determined in the combined prenatal and postnatal cohorts.

(average width of the right and left lateral ventricles for each patient) was significantly greater for patients who required HT in the prenatal group ( $P = .02$ ), but not in the postnatal group ( $P = .54$ ).

When we compared the prenatal and postnatal repair groups without consideration of the need for HT, there were significant differences between groups (Table 2). Compared with the postnatal repair group, the prenatal repair group demonstrated significantly smaller mean basiocciput convexity and mean basiocciput convexity/length ratio. There was no significant difference in the mean convexity of the PAOM between the prenatal and postnatal repair groups.

## DISCUSSION

Previous studies have evaluated the correlation of several imaging characteristics with the need for future HT in patients with CM2 associated with ONTDs.<sup>7–10</sup> To our knowledge, correlation between imaging characteristics related to posterior fossa maldevelopment, specifically the frequently seen outward convexity of the basiocciput and the need for future HT, has not yet been reported. Our findings indicate that the severity of basiocciput convexity is significantly associated with the future need for HT in both prenatal and postnatal repair groups. We found that 100% of patients required HT with a basiocciput convexity of  $\geq 4$  mm on

initial head CT or MR imaging performed in the first 3 years of life, suggesting possible use of this assessment in clinical practice.

The mechanism for the development of this outward convexity is unclear. The lengths of the basiocciput were smaller in patients requiring HT in both prenatal and postnatal repair groups, suggesting more significant hypoplasia of the basiocciput in the HT group. This suggestion is consistent with our finding that the posterior fossa diameter is also smaller in the HT group and is consistent with the known occurrence of occipital bone hypoplasia in the setting of CM2.<sup>11</sup> However, our findings indicate that when we controlled for differences in the length of the basiocciput using the basiocciput convexity-to-length ratio, the convexity was still significantly correlated with the need for HT, possibly implying a mechanism unrelated to occipital bone hypoplasia. One possible explanation for the outward convexity is that as the CSF pressure increases in the setting of worsening hydrocephalus, there is pressure-induced outward remodeling of the basiocciput. However, the unified theory suggests that CSF leakage through the ONTD leads to low CSF pressure, which subsequently leads to posterior fossa maldevelopment.<sup>3</sup> Therefore, though not directly assessed in the present study, documentation of consistent basiocciput outward convexity before ONTD repair would suggest that just abnormal CSF pressure is unlikely to be the cause.

Another possible mechanism could be mass effect related to accommodation of the posterior fossa contents in the setting of a small posterior fossa in CM2. The proximity of the belly of the pons to the dorsal basiocciput in the setting of hindbrain herniation before ONTD repair may be responsible for the outward convexity of the basiocciput. Interestingly, there were some cases noted during this retrospective review in which the hindbrain herniation was so extreme that the pons descended to the level of the foramen magnum; in these cases, the foramen magnum appeared expanded and there was notably less basiocciput curvature. This feature would provide support for the mass effect from the pons as a primary mechanism. If true, the degree of basiocciput convexity may function as a surrogate and quantitative measure for more severe initial hindbrain herniation before repair, but not for extreme hindbrain herniation. Further studies may help to establish a causative mechanism, explain the correlation with future HT, and provide insight into the pathophysiology of hydrocephalus.

While the outward convexity of the PAOM was significantly correlated with the need for HT in the postnatal repair group, it was not significantly correlated in the prenatal repair group, though there was a trend toward significance ( $P = .08$ ). A possible cause of the lack of correlation in the evaluation of the PAOM convexity is differences in flexion/extension positioning of the head, because flexion of the head relative to the cervical spine causes a lengthening of the PAOM and may result in a reduction of the measurable outward convexity. Additionally, while a larger mean ventricle size was significantly correlated with the need for HT in the prenatal repair group, it was not significantly correlated in the postnatal group.

Study limitations include the retrospective nature, image evaluation performed by a single radiologist, lack of head CT or MR



**Table 1: CT and MR imaging data of patients who underwent either prenatal or postnatal surgical repair of an ONTD, according to whether they required HT<sup>a</sup>**

	HT	No HT	P Value
Combined prenatal and postnatal			
Mean basiocciput convexity (mm)	2.6 (SD, 0.9)	1.6 (SD, 0.9)	<.001
Mean basiocciput length (mm)	15.2 (SD, 1.8)	17.3 (SD, 2.2)	<.001
Mean basiocciput convexity/length ratio	0.17 (SD, 0.06)	0.1 (SD, 0.06)	<.001
Mean convexity of PAOM (mm)	2.6 (SD, 1.6)	1.9 (SD, 1.4)	.003
Mean lateral ventricle width (mm)	35.1 (SD, 26.2)	29.4 (SD, 9.9)	.10
Mean FOD (mm)	134.8 (SD, 17.2)	152.2 (SD, 15.5)	<.001
Mean PFD (mm)	51.4 (SD, 9.2)	65.9 (SD, 9.6)	<.001
Mean FOD/PFD ratio	2.7 (SD, 0.4)	2.3 (SD, 0.2)	<.001
Prenatal only			
Mean basiocciput convexity (mm)	2.3 (SD, 0.7)	1.5 (SD, 0.8)	<.001
Mean basiocciput length (mm)	15.6 (SD, 2.1)	17.7 (SD, 2.2)	<.001
Mean basiocciput convexity/length ratio	0.15 (SD, 0.05)	0.08 (SD, 0.05)	<.001
Mean convexity of PAOM (mm)	2.6 (SD, 1.9)	1.9 (SD, 1.4)	.09
Mean lateral ventricle width (mm)	35 (SD, 10.9)	28.8 (SD, 9.9)	.02
Mean FOD (mm)	140.5 (SD, 19.0)	154.2 (SD, 12.6)	<.001
Mean PFD (mm)	56 (SD, 10.4)	67.3 (SD, 8.0)	<.001
Mean FOD/PFD ratio	2.6 (SD, 0.3)	2.3 (SD, 0.2)	<.001
Postnatal only			
Mean basiocciput convexity (mm)	2.8 (SD, 0.9)	1.9 (SD, 1.0)	<.001
Mean basiocciput length (mm)	15 (SD, 1.6)	16.5 (SD, 2.0)	.002
Mean basiocciput convexity/length ratio	0.19 (SD, 0.06)	0.12 (SD, 0.07)	<.001
Mean convexity of PAOM (mm)	2.7 (SD, 1.5)	1.8 (SD, 1.5)	.02
Mean lateral ventricle width (mm)	35.1 (SD, 31.2)	30.7 (SD, 10.0)	.53
Mean FOD (mm)	131.6 (SD, 15.5)	148 (SD, 20.0)	<.001
Mean PFD (mm)	48.9 (SD, 7.5)	62.8 (SD, 11.8)	<.001
Mean FOD/PFD ratio	2.7 (SD, 0.4)	2.4 (SD, 0.2)	<.001

<sup>a</sup> Data are presented as mean (SD). P values were determined by an unpaired Student t test.

**Table 2: CT and MR imaging data of all patients regardless of whether they received HT, according to whether they received prenatal or postnatal surgical repair of an ONTD<sup>a</sup>**

All patients	Prenatal	Postnatal	P Value
Mean basiocciput convexity (mm)	1.8 (SD, 0.9)	2.5 (SD, 1.0)	<.001
Mean basiocciput length (mm)	16.9 (SD, 2.4)	15.4 (SD, 1.9)	<.001
Mean basiocciput convexity/length ratio	0.11 (SD, 0.06)	0.17 (SD, 0.07)	<.001
Mean convexity of PAOM (mm)	2.2 (SD, 1.6)	2.4 (SD, 1.5)	.35
Mean lateral ventricle width (mm)	31.1 (SD, 10.6)	33.8 (SD, 27.0)	.44
Mean FOD (mm)	149.1 (SD, 16.6)	136.6 (SD, 18.4)	<.001
Mean PFD (mm)	63.1 (SD, 10.5)	53.1 (SD, 11.1)	<.001
Mean FOD/PFD ratio	2.4 (SD, 0.3)	2.6 (SD, 0.4)	<.001

<sup>a</sup> Data are presented as mean (SD). P values were determined by an unpaired Student t test.

imaging studies in patients without ventriculomegaly or clinical signs or symptoms or hydrocephalus (instead, head ultrasound and clinical follow-up were used), differences in age at the time of initial head CT or MR imaging between patients who required HT and those who did not (which was due to the differences in clinical need for imaging), and variable duration of follow-up to assess HT. Measurements were performed on either CT or MR imaging, possibly introducing measurement error in comparing the dimensions of an osseous structure, which is better evaluated by CT; further studies may be helpful to assess significant differences in measurements between modalities. In addition, our study does not evaluate absolute intracranial pressure measurements and does not evaluate altered CSF flow pulsatility, which may play a role in the skull base remodeling. The main study strengths include the large sample size as well as all patient management occurring at a single institution, possibly reducing variability in the decision to perform neurosurgical HT.

## CONCLUSIONS

In pediatric patients with CM2 related to an ONTD, the severity of the basiocciput outward convexity was significantly correlated with the need for future neurosurgical HT in both prenatal and postnatal ONTD repair groups. The convexity of the posterior atlanto-occipital membrane was significantly correlated with the need for postnatal hydrocephalus treatment in patients with postnatal but not prenatal spinal neural tube defect repairs. Notably, 100% of patients required HT with a basiocciput convexity of  $\geq 4$  mm. These measurements may be useful in predicting the need for HT.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

- Williams J, Mai CT, Mulinare J, et al; Centers for Disease Control and Prevention. Updated estimates of neural tube defects prevented by mandatory folic acid fortification, United States, 1995-2011. *MMWR Morb Mortal Wkly Rep* 2015;64:1-5 Medline
- Tortori-Donati P, Rossi A, Cama A. Spinal dysraphism: a review of neuro-radiological features with embryological correlations and proposal for a new classification. *Neuroradiology* 2000;42:471-91 CrossRef Medline
- McLone DG, Knepper PA. The cause of Chiari II malformation: a unified theory. *Pediatr Neurosci* 1989;15:1-12 CrossRef Medline
- Miller JL, Huisman TAGM. Spinal dysraphia, Chiari 2 malformation, unified theory, and advances in fetoscopic repair. *Neuroimaging Clin N Am* 2019;29:357-66 CrossRef Medline
- McDowell MM, Blatt JE, Deibert CP, et al. Predictors of mortality in children with myelomeningocele and symptomatic Chiari type II malformation. *J Neurosurg Pediatr* 2018;21:587-96 CrossRef Medline
- Blount JP, Maleknia P, Hopson BD, et al. Hydrocephalus in spina bifida. *Neurol India* 2021;69:S367-71 CrossRef Medline
- Rintoul NE, Sutton LN, Hubbard AM, et al. A new look at myelomeningoceles: functional level, vertebral level, shunting, and the implications for fetal intervention. *Pediatrics* 2002;109:409-13 CrossRef Medline
- Bruner JP, Tulipan N, Reed G, et al. Intrauterine repair of spina bifida: preoperative predictors of shunt-dependent hydrocephalus. *Am J Obstet Gynecol* 2004;190:1305-12 CrossRef Medline
- Tulipan N, Wellons JC 3rd, Thom EA, et al; MOMS Investigators. Prenatal surgery for myelomeningocele and the need for cerebrospinal fluid shunt placement [published correction appears in *J Neurosurg Pediatr* 2022;31:87-89]. *J Neurosurg Pediatr* 2015;16:613-20 CrossRef Medline
- Zarutskie A, Guimaraes C, Yopez M, et al. Prenatal brain imaging for predicting need for postnatal hydrocephalus treatment in fetuses that had neural tube defect repair in utero. *Ultrasound Obstet Gynecol* 2019;53:324-34 CrossRef Medline

11. Calandrelli R, Pilato F, Massimi L, et al. **Posterior cranial fossa maldevelopment in infants with repaired open myelomeningoceles: double trouble or a dynamic process of posterior cranial fossa abnormalities?** *World Neurosurg* 2020;141:e989–97 CrossRef Medline
12. Patel D, Saindane A, Oyesiku N, et al. **Variant sella morphology and pituitary gland height in adult patients with Chiari II malformation: potential pitfall in MRI evaluation.** *Clin Imaging* 2020;64:24–28 CrossRef Medline
13. Naidich TP, Pudlowski RM, Naidich JB, et al. **Computed tomographic signs of the Chiari II malformation, Part I: skull and dural partitions.** *Radiology* 1980;134:65–71 CrossRef Medline
14. Bagci AM, Lee SH, Nagornaya N, et al. **Automated posterior cranial fossa volumetry by MRI: applications to Chiari malformation type I.** *AJNR Am J Neuroradiol* 2013;34:1758–63 CrossRef Medline
15. Roller LA, Bruce BB, Saindane AM. **Demographic confounders in volumetric MRI analysis: is the posterior fossa really small in the adult Chiari 1 malformation?** *AJR Am J Roentgenol* 2015;204:835–41 CrossRef Medline
16. Hashimoto H, Takemoto O, Chiba Y. **Growth patterns and ratios of posterior cranial fossa structures in the Japanese pediatric population: a study utilizing CT scans.** *Neuroradiology* 2023;65:1835–44 CrossRef Medline
17. Nagaraj UD, Kline-Fath BM. **Imaging diagnosis of ventriculomegaly: fetal, neonatal, and pediatric.** *Childs Nerv Syst* 2020;36:1669–79 CrossRef Medline

# Neurofilament Light Chain Levels Interact with Neurodegenerative Patterns and Motor Neuron Dysfunction in Amyotrophic Lateral Sclerosis

Penelope Tilsley, Antoine Moutiez, Alexandre Brodovitch, Mohamed Mounir El Mendili, Benoit Testud, Wafaa Zaaraoui, Annie Verschueren, Shahram Attarian, Maxime Guye, José Boucraut, Aude-Marie Grapperon, and Jan-Patrick Stellmann

## ABSTRACT

**BACKGROUND AND PURPOSE:** Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease involving rapid motor neuron degeneration leading to brain, primarily precentral, atrophy. Neurofilament light chains are a robust prognostic biomarker highly specific to ALS, yet associations between neurofilament light chains and MR imaging outcomes are not well-understood. We investigated the role of neurofilament light chains as mediators among neuroradiologic assessments, precentral neurodegeneration, and disability in ALS.

**MATERIALS AND METHODS:** We retrospectively analyzed a prospective cohort of 29 patients with ALS (mean age, 56 [SD, 12] years; 18 men) and 36 controls (mean age, 49 [SD, 11] years; 18 men). Patients underwent 3T ( $n=19$ ) or 7T ( $n=10$ ) MR imaging, serum ( $n=23$ ) and CSF ( $n=15$ ) neurofilament light chains, and clinical ( $n=29$ ) and electrophysiologic ( $n=27$ ) assessments. The control group had equivalent 3T ( $n=25$ ) or 7T ( $n=11$ ) MR imaging. Two trained neuroradiologists performed blinded qualitative assessments of MR imaging anomalies ( $n=29$  patients,  $n=36$  controls). Associations between precentral cortical thickness and neurofilament light chains and clinical and electrophysiologic data were analyzed.

**RESULTS:** We observed extensive cortical thinning in patients compared with controls. MR imaging analyses showed significant associations between precentral cortical thickness and bulbar or arm impairment following distributions corresponding to the motor homunculus. Finally, uncorrected results showed positive interactions among precentral cortical thickness, serum neurofilament light chains, and electrophysiologic outcomes. Qualitative MR imaging anomalies including global atrophy ( $P=.003$ ) and FLAIR corticospinal tract hypersignal anomalies ( $P=.033$ ), correlated positively with serum neurofilament light chains.

**CONCLUSIONS:** Serum neurofilament light chains may be an important mediator between clinical symptoms and neuronal loss according to cortical thickness. Furthermore, MR imaging anomalies might have underestimated prognostic value because they seem to indicate higher serum neurofilament light chain levels.

**ABBREVIATIONS:** ALS = amyotrophic lateral sclerosis; ALSFRS-R = revised ALS Functional Rating Scale; CMAP = compound muscle action potential; CST = corticospinal tract; GLM = general linear model; LMN = lower motor neuron; MEP = motor-evoked potential; MNI = Montreal Neurological Institute; NfL = neurofilament light chains; TMS = transcranial magnetic stimulation; UMN = upper motor neuron

Amyotrophic lateral sclerosis (ALS) is a lethal neurodegenerative disease involving upper and lower motor neurons (UMN/LMN). Clinical manifestations initially predominate in either limb (~65%) and bulbar (~30%) or respiratory (~5%) territories, with deaths occurring anywhere from as little as a

few months to up to a decade postdiagnosis, often due to respiratory failure.<sup>1-3</sup> Brain MR imaging is an established tool used to map and quantify regional neuronal pathology across a range of neurodegenerative diseases. In ALS, MR imaging alterations primarily occur in the motor cortex and corticospinal tract (CST).<sup>4</sup> In particular, the utility of conventional structural MR

Received March 27, 2023; accepted after revision November 8.

From the Centre de Résonance Magnétique Biologique et Médicale (P.T., M.M.E.M., B.T., W.Z., A.V., M.G., A.-M.G., J.-P.S.), Centre National de la Recherche Scientifique, Institut National de la Santé et de la Recherche Médicale (S.A.), Marseille Medical Genetics Center, and Institut National de la Santé et de la Recherche Médicale (J.B.), Institut de Neurosciences des Systèmes, Aix-Marseille University, Marseille, France; Assistance Publique-Marseille Hospitals (P.T., M.M.E.M., B.T., W.Z., M.G., J.-P.S.), Department of Neuroradiology (A.M., B.T., J.-P.S.), and Referral Centre for Neuromuscular Diseases and ALS (A.V., S.A., A.-M.G.), Assistance Publique-Marseille Hospitals, Hôpital de la Timone, Marseille, France; and Immunology Laboratory (A.B., J.B.), Assistance Publique-Marseille Hospitals, Conception Hospital, Marseille, France.

P. Tilsley and A. Moutiez contributed equally to this work.

This work was supported by Assistance Publique-Marseille Hospitals AORC Junior 2014 program, Association pour la Recherche sur la Sclérose Latérale Fédération pour la Recherche sur le Cerveau, and the European Commission Horizon 2020 Framework Program 945539.

Please address correspondence to Penelope Tilsley, PhD, CRMBM-CEMEREM, Center Hospitalier Régional de Marseille, APHM, 264 rue Sainte-Pierre, Marseille, 13385, France; e-mail: tilsley.penelope@gmail.com

Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8154>

imaging such as T1-weighted images in ALS has been previously demonstrated by evidencing cortical atrophy or cortical thickness loss, a hallmark of neuronal death, primarily across the motor cortex.<sup>5</sup> Furthermore, standard structural T2-weighted FLAIR images also show increased signal anomalies (eg, hyperintensities) at the level of the CST,<sup>5,6</sup> detectable both quantitatively and qualitatively through neuroradiologic assessments.<sup>7</sup> Despite these findings, the widespread clinical implementation of MR imaging in ALS is limited to excluding alternative diagnoses such as ALS mimics and remains to be established as a prognostic tool in ALS.<sup>8</sup>

The association between ALS and brain pathology as detected by MR imaging hence remains moderate and suggests the presence of yet-unknown or incompletely understood pathophysiologic mechanisms mediating this relationship. While progress is being made with more advanced mapping of clinical symptoms according to revised ALS Functional Rating Scale (ALSFRS-R) subscores,<sup>9,10</sup> more precise mapping of correlating biomarkers is, so far, missing. Recent advances have demonstrated that the accumulation of neurofilament light chains (NfL) in serum and CSF, an indicator of axonal damage and neural death,<sup>11</sup> is increased more than 9-fold in ALS compared with controls.<sup>12</sup> NfL seem a highly sensitive, robust biomarker for the diagnosis and prognosis of ALS,<sup>13-15</sup> with relatively good specificity compared with neuroinflammatory<sup>16</sup> and other neurodegenerative diseases.<sup>17</sup> NfL could thus be an informative mediator between cortical neuronal loss, observable with both qualitative neuroradiologic rating and quantitative assessment of MR images, and loss of function in ALS, visible through clinical assessments and electrophysiologic criterion standards such as transcranial magnetic stimulation (TMS), which can depict LMN/UMN dysfunction.<sup>18,19</sup> For example, the TMS compound muscle action potential (CMAP), a standard measure of nerve conduction, progressively decreases in ALS, reflecting LMN degeneration,<sup>20,21</sup> while motor-evoked potentials (MEPs) are reduced or sometimes completely absent in ALS, reflecting UMN dysfunction and CST degradation.<sup>18,22</sup> The ratio between them is thought to provide more information about UMN degeneration by factoring out LMN-related biases.<sup>23</sup> The presence of new biologic and imaging biomarkers gives the opportunity to evaluate patients' evolution in vivo, but the link between what they mean and measure in terms of clinical, functional, and anatomic terms is still lacking.

The goal of the study was, thus, to investigate the link between qualitative neuroradiologic assessments or quantitative neurodegeneration in the precentral cortex and either serum and CSF NfL levels or clinical and electrophysiologic outcome measures in a population of patients with ALS.

## MATERIALS AND METHODS

### Participants

We included retrospective data from 29 patients with ALS prospectively recruited between 2013 and 2020 from the Referral Center for Neuromuscular Diseases and ALS, Marseille, France, having both MR imaging and either a serum or CSF NfL assessment. Inclusion criteria were patients who fulfilled the revised El Escorial diagnostic criteria:<sup>24</sup> definite ALS ( $n = 5$ ), probable ALS ( $n = 15$ ), probable laboratory-supported ALS ( $n = 2$ ), and

possible ALS ( $n = 7$ ). Exclusion criteria were comorbid neurologic conditions or coexisting frontotemporal dementia, which was excluded on the basis of the *Diagnostic and Statistical Manual of Mental Illnesses*-5th edition dementia criteria and frontotemporal dementia criteria.<sup>25</sup> All patients were receiving medical treatment with riluzole, and most patients were right-handed, with only 2 patients having a laterality score below +40 on the Edinburgh Handedness Inventory, below the threshold for right-handedness.<sup>26</sup> Patients underwent clinical and electrophysiologic examinations, 3T or 7T MR imaging, and NfL assessments. Thirty-six healthy controls with an equivalent 3T or 7T MR imaging served as a control group. The inclusion criterion for control participants' data was a minimum age of 25 years, the minimum age of the ALS cohort. The study had ethics approval, and participants signed an informed consent before participating in the study.

### Neurofilament Protocol

Serum and CSF NfL levels (picogram/milliliter) were measured with the R-PLEX Human Neurofilament L Antibody Set (MesoScale Discovery, F217X) according to the manufacturer's instructions.

### Clinical Examination

Clinical disability was quantified using ALSFRS-R and ALSFRS-R subscores. The bulbar subscore corresponded to items 1–3 (speech, salivation, and swallowing functions); arm subscore, to items 4–6 (fine motor function: handwriting, cutting food, and dressing and hygiene); leg subscore, to items 7–9 (gross motor function: turning in bed, walking, and climbing stairs); and respiratory subscore, to items 10–12 (dyspnea, orthopnea, and respiratory insufficiency). Disease duration since symptom onset, date of diagnosis, onset site (bulbar, left leg, right leg, left arm, or right arm), and forced vital capacity were measured for all patients. The disease progression rate (ALSFRS-R slope) was also calculated on the basis of ALSFRS-R subscores sampled at 3, 6, and 12 months. Within the current article, sample size limited grouped analyses across different subgroups, such as patients with differing onset sites. The analyses presented in this article were restricted to associations across ALSFRS-R scores and subscores.

### Electrophysiologic Protocol

TMS was performed using a Magstim 200 stimulator (Magstim) with a circular coil (9-cm diameter). The electromyography machine, Nicolet Viking EDX (Natus Medical) was used to record surface electromyography signals with a band-pass filter of 3–10,000 Hz and performed electrical stimulation. MEPs were recorded on both sides for the arm (abductor digiti minimi muscles) and leg (tibialis anterior muscles). MEP ratios were calculated by dividing the MEP amplitude elicited by TMS by the amplitude of the CMAP of the same muscles evoked by supramaximal electrical stimulation of the corresponding peripheral nerve (ulnar or fibular). TMS was not performed if the CMAP amplitude was <1 mV. The central motor conductance time was not measured because it cannot be measured if MEP values are too low.



### MR Imaging Protocol

The 3T protocol included a 3D T1-weighted MPRAGE sequence (TR/TE/IT = 2300/2.98/900 ms, 160 sagittal slices, 1-mm<sup>3</sup> isotropic) and FLAIR (TR/TE/IT = 5000/395/1800 ms, 160 sagittal slices, 1-mm<sup>3</sup> isotropic) and was acquired using a 32-channel (32Rx/1Tx) head coil (Siemens). The 7T MR imaging protocol included a 3D T1-weighted MP2RAGE (TR/TE/TI1/TI2 = 5000/3/900/2750 ms, 256 sagittal slices, 0.6 mm<sup>3</sup> isotropic), FLAIR (TR/TE/IT = 8500/466/2400 ms, 192 sagittal slices, 0.8 mm<sup>3</sup> isotropic), and a sagittal 2D magnetization-prepared turbo FLASH B1+ mapping<sup>27</sup> (TR/TE = 2000/14 ms, 14 axial sections, section thickness = 5 mm, in plane resolution = 3.9 × 3.9 mm<sup>2</sup>) and was acquired using a 32-channel (32Rx/1Tx) head coil.

### Image Processing

7T MP2RAGE images were denoised.<sup>28</sup> 3T and 7T T1-weighted images were then skull-stripped using HD-BET (<https://github.com/MIC-DKFZ/HD-BET>).<sup>29</sup> FreeView (<https://medimng.agfa.com/main/freelview-technology>) was used to manually correct brain masks. FreeSurfer recon-all pipeline (<https://surfer.nmr.mgh.harvard.edu/fswiki/recon-all>) was used to perform surface-based analysis and estimate whole-brain cortical thickness in conjunction with *mris\_preproc* and *mri\_surf2surf*. Individual data were smoothed with a full width at half maximum of 10 mm and registered to the fsaverage brain.

### Blinded Neuroradiologic Rating

A blinded assessment of patient and control MRIs was performed by 2 trained radiologists. The radiologists had 4 years (rater A) and 8 years (rater B) of experience. The radiologists were each given access to anonymized folders containing T1 and FLAIR images when available. Each folder contained either 3T or 7T images from 1 participant (either a control or patient). The radiologists were not informed as to the field strength of each folder or whether the folder contained control or patient images. Atrophy for the whole-brain (global), precentral, frontal, parietal, and temporal lobes was reported as “no,” “mild,” or “marked” for each hemisphere. The occipital lobe was used as a reference, given the relative lack of atrophy in this area in ALS. Atrophy of the precentral gyrus was reported only when very marked, given that it is the primary atrophic zone in ALS. CST atrophy was determined on the basis of the peduncles and pons. Atrophy to both was rated as “yes,” only when one was “undetermined” and neither was “no.”

FLAIR CST anomalies were reported using a window of 0/210 at 3T and 50/400 at 7T. MR images with substantial movement artifacts were rated as “artifacts” and removed from analyses. Any missing data were rated as “not applicable” and removed from analyses in a pair-wise fashion. Following the initial rating, the radiologists reviewed the corresponding images and provided a consensus rating when discordance occurred. Disparities were resolved by the 2 radiologists reviewing the images together and deciding which rating was the most accurate and submitting this rating as the final consensus rating. This process involved reducing ratings of CST atrophy to “yes” for atrophy to either the pons or peduncles, or “no” when no atrophy was present. This rating was used for investigating statistical associations. Participants

were also grouped into those with no anomalies or at least 1 anomaly across all ratings.

### Statistics

Demographic and clinical data were analyzed using R (Version 1.2.5033, packages: tidyverse [<https://www.tidyverse.org/>] and lmerTest [<https://rdrr.io/cran/lmerTest/man/lmerTest-package.html>]). Differences between control and patient characteristics and control and patient blind test ratings across raters were determined using 2-sample *t* tests or Wilcoxon *t* tests, depending on data normality. Associations were examined with Pearson correlations. When multiple ordered groups were present, an increase or decrease in mean values over ordered groups was analyzed with the Jonckheere-Terpstra tests, allowing, for example, detection of a trend with increasing atrophy across no-versus-mild-versus-marked atrophy groups. Any missing values marked as “not applicable” were omitted from analyses in a pair-wise fashion.

Spatial patterns of cortical thickness between patients and controls were analyzed using the surface-based general linear model (GLM) analyses (*mri\_glmfit* and *mri\_glmfit-sim*) of FreeSurfer. Between-patient analyses were restricted to the precentral area from the Desikan-Killiany atlas (?apar.annot: precentral).<sup>30</sup> Associations among patients’ cortical thickness and ALSFRS-R subscores, MEP ratios, and serum or CSF NfL levels were also examined using GLMs. All models were adjusted for age, sex, and scanner (3T or 7T) and allowed random intercepts. An unadjusted *P* value below < .05 was used to explore the results. False discovery rate cluster correction was run with a *P* value < .05, using 1000 permutations.

## RESULTS

### Demographic and Clinical Data

Data from 29 patients with ALS and 36 healthy controls were analyzed. All patients participated in an MR imaging assessment, blinded assessment, and clinical assessment and had either a serum or CSF NfL assessment. Patients’ data included 3T (*n* = 19) or 7T (*n* = 10) MR imaging, clinical examinations including ALSFRS-R scores (*n* = 29), electrophysiological assessments (*n* = 27), blinded assessments by 2 raters (*n* = 29), and NfL serum (*n* = 23) or CSF (*n* = 15) samples. Two patients did not have all electrophysiologic measures due to refusal to undergo certain TMS procedures and were removed from corresponding analyses, and in a further 2 patients, it was not possible to acquire MEP data for either the right arm (*n* = 1) or lower legs (*n* = 1). Within the NfL assessments, 1 patient’s sample was removed from the serum NfL assessment because it was classed as an outlier, 2 patients had no serum sample, 2 patients had no CSF sample, and 15 samples were removed due to problems during the quantification process. Control data included a 3T (*n* = 25) or 7T (*n* = 11) MR imaging and blinded assessments by 2 raters (*n* = 36). As hypothesized, we observed decreased cortical thickness in patients compared with controls (Online Supplemental Data).

Patients’ demographic, control group demographic and clinical details can be found in the Table. The control group had a mean age of 49 (SD, 11) years, with median and range of 51 [27–66] years, and a sex ratio of 1:1 male to female. Two-sample *t* tests comparing age and sex between controls and patients showed no

difference for sex ( $t=585$ ,  $P=.34$ ), but a difference for age ( $t=372$ ,  $P=.048$ ). Log NfL values were used for analyses due to the non-normal distribution of raw serum NfL levels according to the Shapiro-Wilk test ( $W$  statistic = 0.89,  $P=.020$ ). The Pearson correlation showed a positive correlation between serum

#### Demographic and clinical details of patients with ALS and the control cohort<sup>a</sup>

Variable	Central Tendency	Spread	Count (No.)
Patient cohort			
Male/female ratio	38%	[0–1]	29
Age (yr)	56	(SD, 12)	29
Duration of disease (mo)	13	[5–84]	29
Age at disease onset (yr)	55	(SD, 12)	29
Spinal: Bulbar ratio	75%	[0–1]	28
Vital capacity (%)	95%	(SD, 20)	28
Weight (kg)	69	(SD, 14)	29
ALSFRS-R (/48)	38	(SD, 5.2)	29
ALSFRS-R bulbar (/12)	12	[4–12]	29
ALSFRS-R arm (/12)	10	[3–12]	29
ALSFRS-R leg (/12)	6	[2–12]	29
ALSFRS-R slope (1 year)	0.53	[0.12–3.6]	29
MEP ratio left arm	43	[0–76]	27
MEP ratio right arm	37	[0–72]	26
MEP ratio left leg	0	[0–63]	26
MEP ratio right leg	0	[0–69]	26
Serum NfL	400	[65–1800]	23
Log serum NfL	6	(SD, 0.86)	23
CSF NfL	32,000	(SD, 19,000)	15
Log CSF NfL	10	(SD, 0.7)	15
Control cohort			
Male/female ratio	50%	[0–1]	36
Age (yr)	51	[27–66]	36

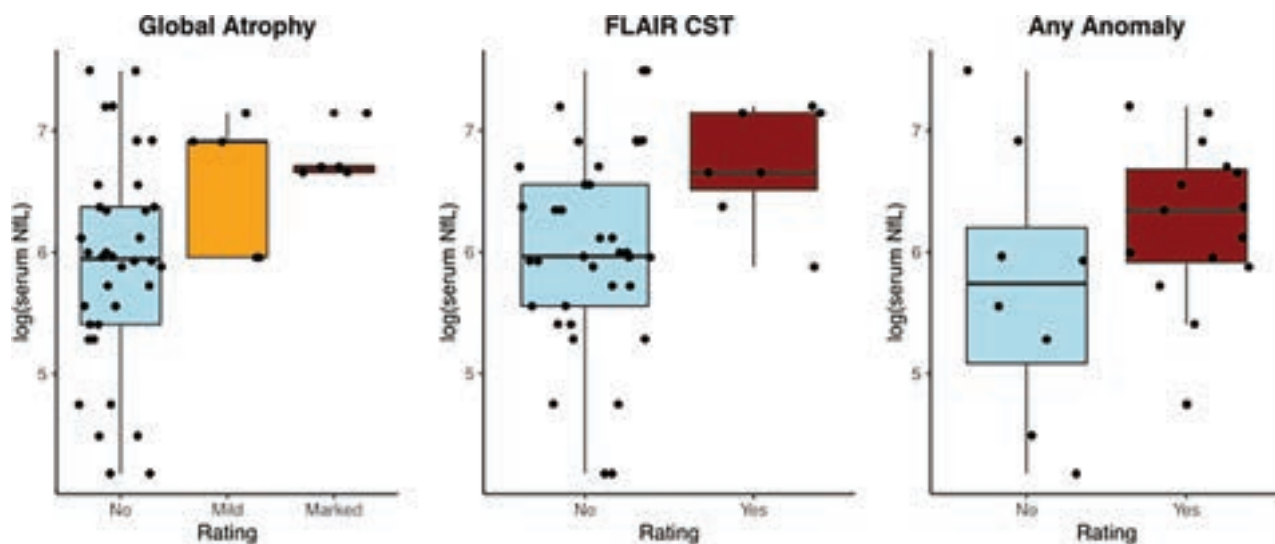
**Note:**— Brackets in the first Variable column refer to units (eg, months, years), or the maximum value for the given scale (eg, /12 being out of 12).

<sup>a</sup> Central tendency and spread are shown as mean (SD) for normally distributed data or median and [range] for non-normally distributed data. Quantitative variables are shown as percentages. The number of participants within each variable analysis is indicated in the column count.

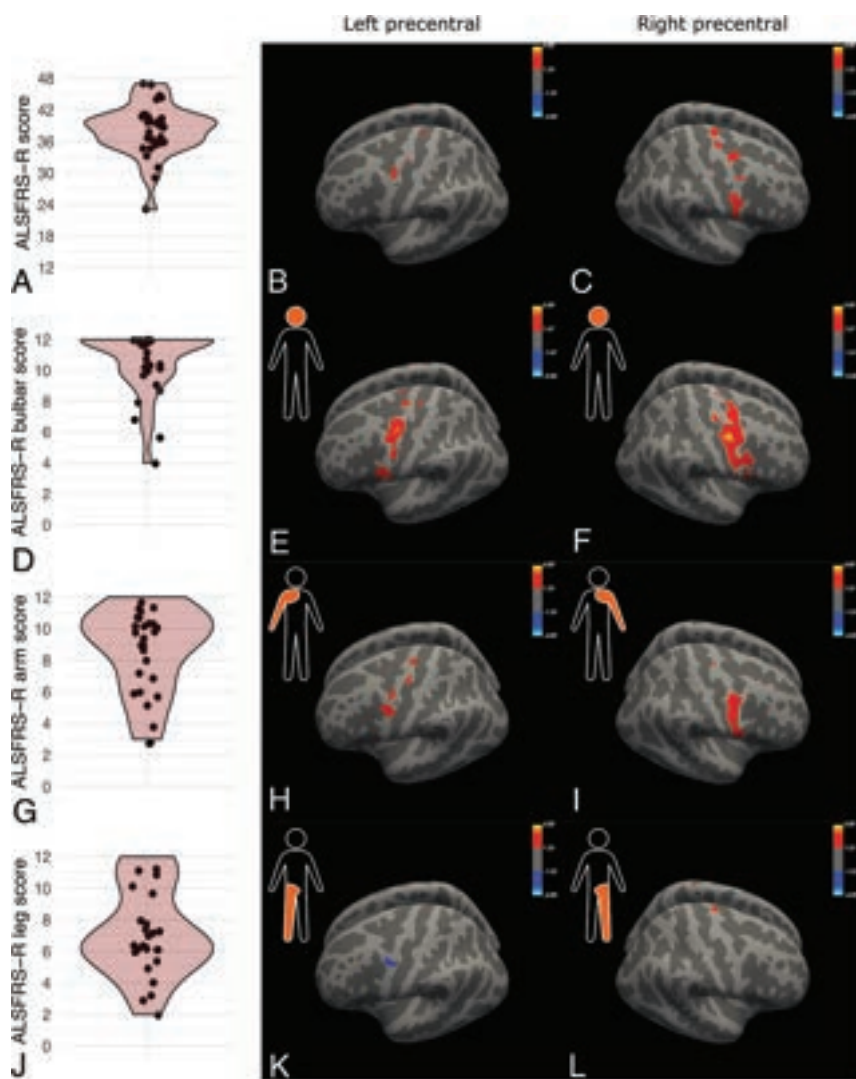
and CSF NfL levels ( $r=0.73$ ,  $P=.003$ ), corroborating previous results in the literature showing a strong correlation between NfL in serum and CSF in ALS<sup>31</sup> and other diseases such as MS<sup>32</sup> and Parkinson disease.<sup>33</sup> A significant negative correlation was also observed between log serum NfL and clinical scores such as ALSFRS-R ( $r=-0.63$ ,  $P=.001$ ). Spearman correlations showed a significant negative correlation between log serum NfL and ALSFRS-R leg subscores ( $r=-0.54$ ,  $P=.007$ ), but not with ALSFRS-R arm subscores ( $r=-0.09$ ,  $P=.67$ ) or bulbar subscores ( $r=-0.10$ ,  $P=.66$ ). Similarly, progression rates indicated by the ALSFRS-R slope calculated across 1 year showed a significant positive correlation with log NfL serum, with quicker progression correlating with greater log serum NfL ( $r=0.48$ ,  $P=.022$ ).

#### Neuroradiologic Rating

Radiologic assessments were performed for 36 controls and 29 patients, with a total of 1020 images rated. FLAIR scans for 9 controls and 1 patient rated as “not applicable” by radiologists were removed from corresponding analyses. One further control had imaging artifacts on CST atrophy and precentral signal and was also removed from the corresponding analyses. Concordance between the raters was 92.9% for the control hemispheres and 83.5% for patient hemispheres ( $P=.036$ ). Global atrophy was rated “no” in 66 control and 43 patient hemispheres, “mild” in 6 control and 10 patient hemispheres, and “marked” in 0 control and 5 patient hemispheres. FLAIR anomalies were rated as “yes” in 0 control and 7 patient hemispheres and “no” in 54 control and 49 patient hemispheres. At least 1 anomaly was found in 5 controls and 24 patients. Subsequent analyses exploring the link between serum NfL and consensus ratings found higher serum NfL levels in patients with FLAIR hypersignal ( $W$  statistic = 62.5,  $P=.033$ ) and an association with global atrophy (Jonckheere-Terpstra test = 292.5,  $P=.003$ ; alternative hypothesis: increasing serum NfL levels with increasing anomalies) (Fig 1). Nonetheless, anomalies were



**FIG 1.** Association between log serum NfL levels and anomalies based on neuroradiologic rating consensus for each hemisphere in patients with ALS. Graphs show consensus ratings compared with log serum NfL values for A, Global atrophy; B, FLAIR in CST; and C, Instances of at least 1 anomaly found on any image. Individual scatterpoints show each individual patient's data, with 2 points per patient corresponding to the consensus ratings of the left or right hemisphere and their corresponding log serum NfL value.



**FIG 2.** ALSFRS-R score and subscore distributions (violin graphs) and corresponding GLM comparison for patients' precentral cortical thickness and ALSFRS-R scores for A–C, global score. D–F, Bulbar. G–I, Arm. J–L, Leg region subscores. Note that while ALSFRS-R subscores are not lateralized, GLM correlations for the left precentral cortex were presumed to be related to right-limb symptoms, and vice versa for the right precentral cortex, as depicted by the orange shading on the silhouette. The color scale (signed  $\log_{10} P$  values) indicates decreased cortical thickness with decreasing (less normal) ALSFRS-R scores in red (and vice versa, negative associations in blue). An uncorrected threshold of 1.3 corresponds to a  $\log_{10}$  of  $P = .05$ . Significant clusters surviving cluster-wise  $P$  value correction ( $P < .05$ ) are outlined in white.

rare and serum NfL values did not differ between patients with no or at least 1 anomaly ( $W$  statistic = 51,  $P = .59$ ). Similar statistical results were found with a simple  $t$  test and grouping global atrophy scores of “mild” and “marked” into a “yes” category. Furthermore, a positive association was also seen between ALSFRS-R slope, an indicator of progression rate, and global atrophy (Jonckheere-Terpstra test = 362.5,  $P$  value  $\leq .001$ ; alternative hypothesis: increasing across no, to mild, to marked atrophy) and FLAIR CST ( $W$  statistic = 70.5,  $P$  value = .034). As with previous results, no significant association was found with any anomaly ( $W$  statistic = 82,  $P = .35$ ).

#### Mapping Disability Scores with Cortical Atrophy

To determine whether patterns of cortical atrophy were associated with limb-specific clinical symptoms or electrophysiologic

results in expected precentral areas, we performed associations between ALSFRS-R subscores and MEP ratios. Figure 2 shows the association between ALSFRS-R scores and cortical thickness. For ALSFRS-R bulbar scores, 1 cluster in the right (precentral gyrus: size = 2395 mm<sup>2</sup>, Montreal Neurological Institute [MNI] [x] = 44.2, MNI [y] = -9.2, MNI [z] = 45.6,  $P = .002$ ) and left (precentral gyrus: size = 927 mm<sup>2</sup>, MNI [x] = -45.6, MNI [y] = -9.5, MNI [z] = 50.5,  $P = .008$ ) hemispheres survived cluster-wise correction. ALSFRS-R arm scores survived cluster-wise correction in the right hemisphere (precentral gyrus: size = 1209 mm<sup>2</sup>, MNI [x] = 60.1, MNI [y] = 2.2, MNI [z] = 19.0,  $P = .010$ ). Results from MEP ratios (Fig 3) showed a similar distribution, with decreased cortical thickness in the upper-leg regions of the precentral cortex being associated with decreased leg MEP ratios.

#### Association among NFL Levels, Electrophysiology, and Cortical Thickness

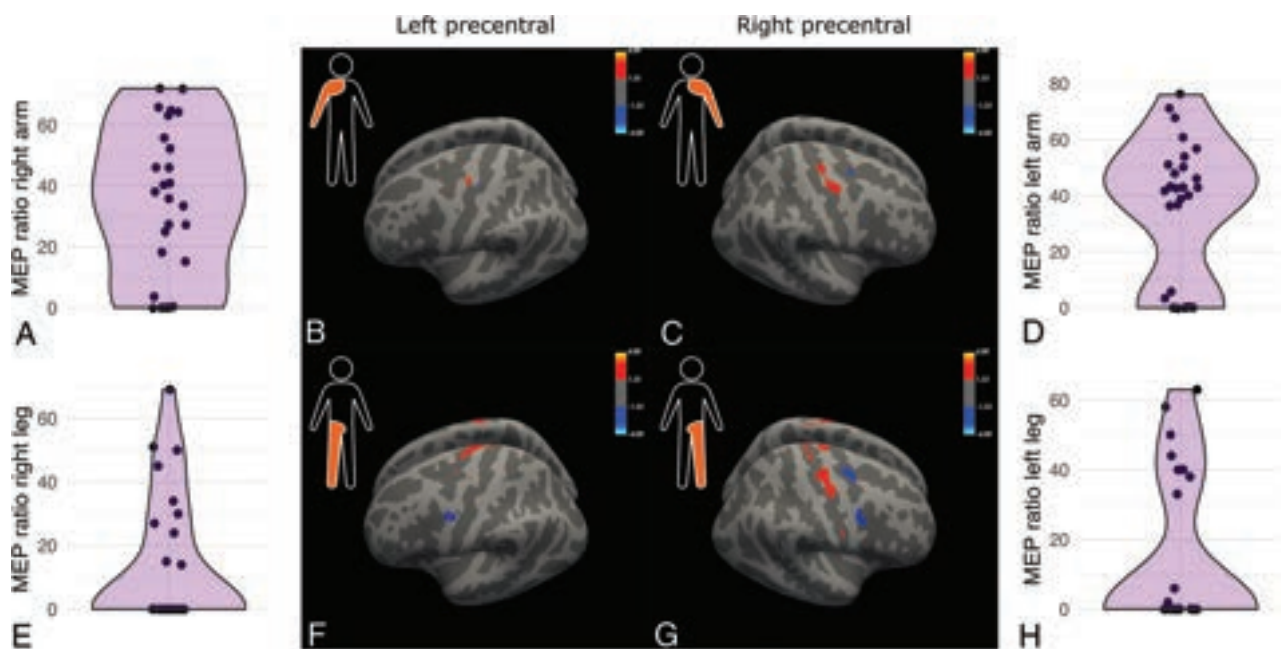
The spatial origin of NfL associated with cortical neurodegeneration is shown in Fig 4. Results showed a negative correlation between NfL levels in CSF or serum with cortical thickness (left hemisphere:  $r = -0.46$ ,  $P = .03$ ; right hemisphere:  $r = -0.46$ ,  $P = .03$ ), ie, reduced cortical thickness with increasing NfL. Furthermore, we explored the interaction between serum NfL levels and MEP ratios for each limb compared with cortical thickness in the contralateral hemisphere (Fig 5). Overall, we found positive interactions, indicating that the inverse effect of serum NfL values

on cortical thickness is stronger (ie, the slope of the linear model is steeper) in patients with lower MEP values (higher disability). These interactions were found across arm and bulbar premotor and motor areas.

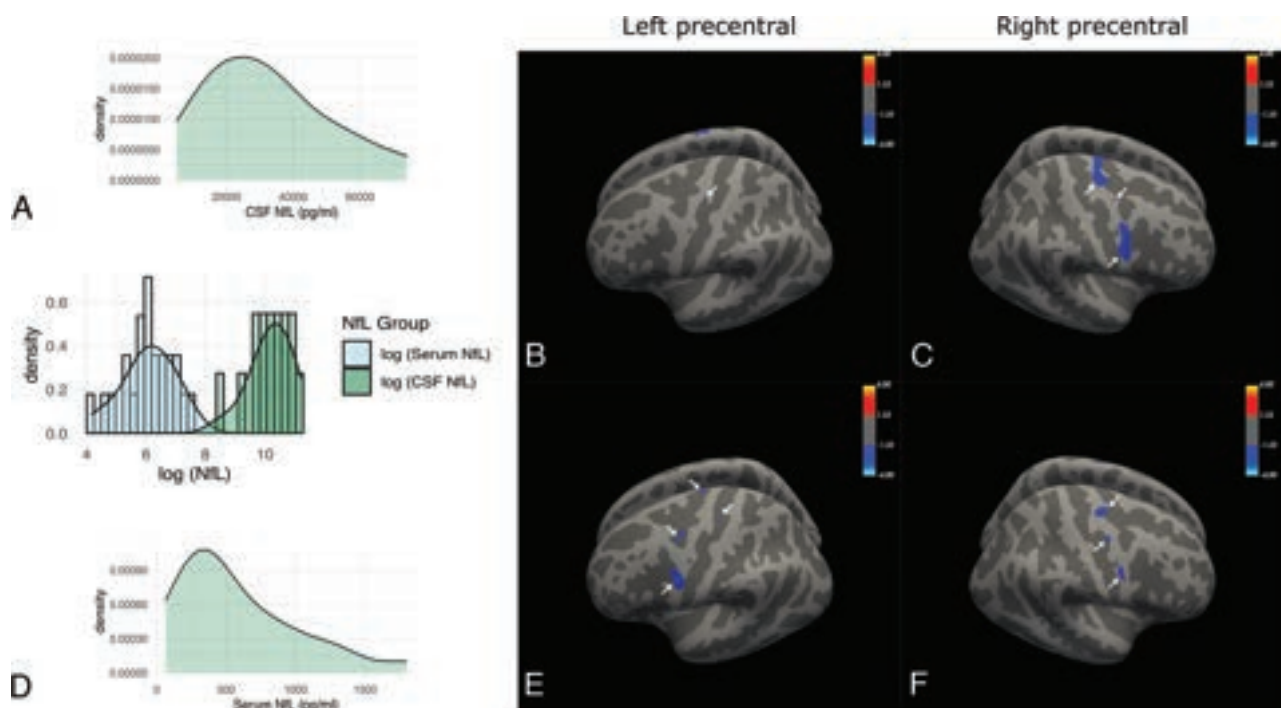
#### DISCUSSION

Our study investigated clinical and research grade MR imaging outcomes as bridging techniques between disability and neuronal loss in a group of patients with ALS. Serum NfL levels, currently the most promising blood biomarker in neurodegenerative diseases, served as a marker of ongoing neuronal loss and were associated with precentral area cortical thickness. While with a small sample size, uncorrected results suggest an interaction between cortical thickness and NfL with electrophysiologic outcomes, an



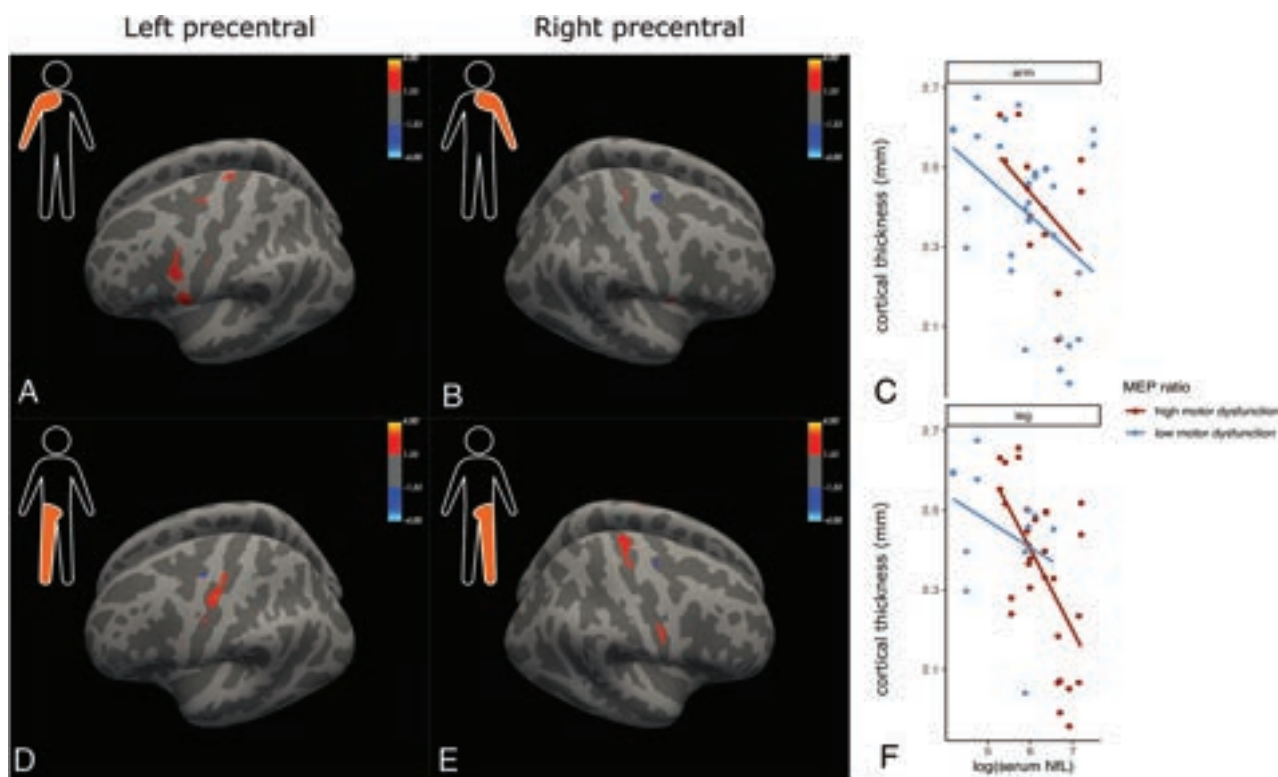


**FIG 3.** MEP ratio distributions (*violin plots*) across patients and corresponding GLM comparison for patients' precentral cortical thickness and electrophysiologic MEP ratios for A and B, Right arm; C and D, Left arm; E and F, Right leg; and G and H, Left leg. Correlations for the left precentral area were performed with right-limb measures, and vice versa for the right precentral area, following the *orange shading* on the silhouettes. The *color scale* (signed  $\log_{10} P$  values) indicates decreased cortical thickness with decreasing (less normal) MEP ratios in red (and vice versa, negative associations in blue). Uncorrected threshold of 1.3 corresponding to a  $\log_{10}$  of  $P = .05$ . Significant clusters surviving cluster-wise  $P$  value correction ( $P < .05$ ) are outlined in white.



**FIG 4.** A and D, Distribution of CSF and serum NfL levels across patients, with corresponding log CSF and serum values represented as a combined histogram-density graph. B and C, GLM comparison for patients' precentral cortical thickness and log CSF NfL levels. E and F, GLM comparison for patients' precentral cortical thickness and log serum NfL levels. The *color scales* in B and C and E and F (signed  $\log_{10} P$  values) indicate decreasing cortical thickness with increasing NfL scores in blue (and vice versa, positive associations in red). An uncorrected threshold of 1.3 corresponds to a  $\log_{10}$  of  $P = .05$ . Note that the negative correlation in the left precentral area is around the premotor cortex. Significant clusters surviving cluster-wise  $P$  value correction ( $P < .05$ ) are outlined in white.





**FIG 5.** GLM demonstrating the interaction among patients' precentral cortical thickness, NfL levels and MEP ratios in A, Right arm; B, Left arm; D, Left leg; and E, Right leg. Panels on the right depict scatterplots and linear model estimations for the interaction between serum NfL and cortical thickness according to MEP ratios for C, the arm and D, leg split into either high motor dysfunction (low MEP ratios) or low motor dysfunction (high MEP ratios) using the mean. The color scale (signed log10 *P* values) where red indicates that higher MEP values were associated with a lower effect of serum NfL on cortical thickness. Uncorrected threshold of 1.3 corresponding to a log10 of *P* = .05. Significant clusters surviving cluster-wise *P* value correction (*P* < .05) are outlined in white.

objective measure of upper motor neuron dysfunction. Finally, despite low observations of anomalies during neuroradiologic rating, our results suggest that the visibility of FLAIR anomalies increases with higher serum NfL levels, which are linked to poorer prognosis.

#### Radiologic Ratings Linked to Serum Biomarkers in ALS

To date, qualitative MR imaging ratings have usually been found to be neither sensitive nor specific for the prognosis of ALS, resulting in a tendency to use clinical and electrophysiologic scores without necessarily having additional MR imaging data.<sup>5</sup> We investigated the link between qualitative and quantitative MR imaging and a highly specific biomarker in ALS, NfL, as well as the link to the clinical and electrophysiologic outcomes. As in other studies in ALS<sup>31</sup> and other diseases such as MS,<sup>32</sup> Parkinson disease,<sup>33</sup> X-linked adrenoleukodystrophy,<sup>34</sup> we confirmed a strong correlation between NfL in serum and CSF. In particular, serum NfL represents an easily-accessible biomarker with a much higher elevation in ALS compared with other diseases and even ALS mimics,<sup>35</sup> and are thought to represent axonal degeneration and death. Here, while detection of anomalies was rare, perhaps due to MR imaging being performed at the beginning of patients' disease courses, we found a meaningful association between serum NfL levels and FLAIR CST signal anomalies and global atrophy. Furthermore, we found a positive association between clinical markers for disease states and

progression with FLAIR CST and global atrophy anomalies. These results are in line with a previous study suggesting that CST hyperintensities on FLAIR images represent a marker of axonal degeneration,<sup>6</sup> the same pathophysiologic process thought to cause the increase in NfL.<sup>11</sup>

Previous studies have also demonstrated similar results, with levels of NfL shown to be positively correlated with the disease progression rate in X-linked adrenoleukodystrophy<sup>34</sup> and ALS,<sup>36</sup> while quantitative MR imaging profiles of patients with ALS vary in their disease-progression rates.<sup>37</sup> These results may support the prognostic value of qualitative MR imaging assessments in ALS. Larger studies with longitudinal quantitative and qualitative MR imaging data and NfL data might be necessary to clearly define the link between atrophy and NfL and could examine associations across subgroups according to the progression rate or genetic mutations, which we did not address in our study. Additionally, our study focused on precentral areas, yet changes in subcortical and extramotor structures, the most frequent longitudinal changes in ALS,<sup>38</sup> or spinal cord atrophy could warrant exploration. It has already been demonstrated, for instance, that the rate of serum NfL change is associated with brain volume changes across several widespread brain areas in frontotemporal dementia, a similar disease thought to be on the same spectrum, with NfL levels able to capture rates of brain atrophy for certain mutation carriers.<sup>39</sup>

### **Spatial Distribution of ALS Symptoms**

We demonstrated that the association between precentral cortical thinning and limb ALSFRS-R subscores and MEP ratios followed a distribution generally respecting the motor homunculus. As with previous results, the spatial distribution was observed bilaterally,<sup>4,6</sup> and associations were prominent for bulbar scores.<sup>9,10</sup> We also observed overlap of subscores into other motor homunculus regions. One explanation could be the nonspecific nature of ALSFRS-R subscores,<sup>40,41</sup> yet we also observed overlap for MEP ratios. Alternatively, studies have suggested apparent spreading of symptoms to adjacent motor regions.<sup>6,9</sup> Our results may thus be in line with recent challenges to the somewhat misinterpreted strict segmentation of the homunculus across sensorimotor areas,<sup>42</sup> with increasing suggestions of a more distributed, overlapping profile, particularly among body regions with higher physical proximity.<sup>43,44</sup>

### **Interaction between Motor Dysfunction and NfL**

Finally, while results did not survive cluster correction, likely due to small sample size, uncorrected results for cortical thickness, MEPs, and NfL showed an overall positive interaction, with a negative association between cortical thickness and serum NfL levels that was stronger in patients with decreased MEP ratios. This result could be interpreted as demonstrating axonal loss primarily at the level of motor cortex neurons or CST. Reduced MEPs, or higher disability, is observed in patients with axonal loss across the motor cortex, while NfL levels reflect a more global axonal loss. Serum NfL may thus be a useful marker of ongoing neurodegeneration, which is associated with cortical atrophy and electrophysiologic measures of motor neuron function. A longitudinal study in patients with ALS showed that serum NfL levels indicated CST degeneration.<sup>31</sup> Furthermore, NfL increase even before symptoms are present, and remain stable and elevated throughout the disease course.<sup>45</sup> Here, we extend arguments for multimodal assessments in ALS<sup>41</sup> by presenting the opportunity to combine noninvasive markers of neurodegeneration and muscle function. It has already been shown, for example, from other cross-sectional studies that the combination of noninvasive measures from DTI and electrophysiology could represent a useful biomarker in ALS during initial disease phases,<sup>46</sup> while, examination of iron abnormalities from susceptibility weighted images represented a marker of UMN degeneration based on ALSFRS-R clinical scores.<sup>6</sup> Further studies with larger sample sizes and longitudinal evolutions of these markers alongside CST degradation or quantitative DTI would be essential in fully detailing the link among NfL, symptoms, and neurodegeneration to create standardized maps of cortical degradation in disease states.<sup>47,48</sup>

### **Limitations**

Our main limitation was reduced power due to a small data set. While we used pair-wise deletion which maximized the amount of data within each analysis, certain measures such as CSF NfL were more heavily impacted by missing values, which reduced our ability to examine associations. Given the strong correlations between serum NfL and CSF NfL, correlations based on the larger serum NfL sample aimed to reduce the potential impact on this article. Ultimately a larger study with serum and CSF NfL data

could examine more deeply the links and differences between these 2 measures in terms of atrophy and clinical outcomes. We also used 3T and 7T MR imaging data, which have differing sensitivities to anomalies and cortical thickness when spatial resolution differs.<sup>49</sup> The use of identical, standardized surface-based postprocessing using FreeSurfer and inclusion of an MR imaging scanner as a covariate in statistical analyses aimed to reduce the associated potential bias.

Within the blinded rating assessment, complete anonymization across participant groups (control or patient) and field strength (3T or 7T) also aimed to reduce similar biases. Performing analyses on a subset of the data without 7T images did not significantly change the presented results. Furthermore, due to the low levels of patients with bulbar onset ( $n = 6$ ), our cohort was largely patients with spinal onset with higher leg disability according to ALSFRS-R subscores and MEP ratios. Ultimately, this feature could have impacted the current results. Additionally, NfL may also be linked to neurodegeneration within the spinal cord, which we did not explore. Finally, while age was incorporated into statistical analyses, FLAIR anomalies and global atrophy become more apparent with age, even in healthy controls. The blinded rating assessments may thus have been influenced by age.

### **CONCLUSIONS**

Our study evidenced the benefit of using radiologic qualitative and quantitative MR imaging assessment of cortical thickness in ALS, alongside noninvasive markers such as serum NfL and clinical and electrophysiologic evaluations for a detailed mapping of ongoing neurodegeneration. While quantitative MR imaging measures have been increasingly promoted, particularly in the context of multimodal assessment of ALS, whether qualitative assessments are able to reflect the same link between neurodegeneration and clinical outcomes has been less demonstrated. Here we provide preliminary evidence toward the link between NfL and both quantitative and qualitative measures to improve our understanding of ALS pathophysiology and extend arguments promoting multimodal standardized approaches in advancing the prognostic capabilities of these methods.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

### **REFERENCES**






1. Hardiman O, Al-Chalabi A, Chio A, et al. **Amyotrophic lateral sclerosis.** *Nat Rev Dis Primers* 2017;3:17071–79 CrossRef Medline
2. Grad LI, Rouleau GA, Ravits J, et al. **Clinical spectrum of amyotrophic lateral sclerosis (ALS).** *Cold Spring Harb Perspect Med* 2017;7:a024117 CrossRef Medline
3. Kiernan MC, Vucic S, Cheah BC, et al. **Amyotrophic lateral sclerosis.** *Lancet* 2011;377:942–55 CrossRef Medline
4. Thorns J, Jansma H, Peschel T, et al. **Extent of cortical involvement in amyotrophic lateral sclerosis: an analysis based on cortical thickness.** *BMC Neurol* 2013;13:148 CrossRef Medline
5. Mazón M, Costa JF, Ten-Estève A, et al. **Imaging biomarkers for the diagnosis and prognosis of neurodegenerative diseases: the example of amyotrophic lateral sclerosis.** *Front Neurosci* 2018;12:784 CrossRef Medline
6. Vázquez-Costa JF, Mazón M, Carreres-Polo J, et al. **Brain signal intensity changes as biomarkers in amyotrophic lateral sclerosis.** *Acta Neurol Scand* 2018;137:262–71 CrossRef Medline

7. Schweitzer AD, Liu T, Gupta A, et al. **Quantitative susceptibility mapping of the motor cortex in amyotrophic lateral sclerosis and primary lateral sclerosis.** *AJR Am J Roentgenol* 2015;204:1086–92 CrossRef Medline
8. Verstraete E, Turner MR, Grosskreutz J, et al; attendees of the 4th NiSALS meeting. **Mind the gap: the mismatch between clinical and imaging metrics in ALS.** *Amyotroph Lateral Scler Frontotemporal Degener* 2015;16:524–29 CrossRef Medline
9. Jin J, Hu F, Zhang Q, et al. **Dominant heterogeneity of upper and lower motor neuron degeneration to motor manifestation of involved region in amyotrophic lateral sclerosis.** *Sci Rep* 2019;9:20059 CrossRef Medline
10. Schuster C, Kasper E, Machts J, et al. **Focal thinning of the motor cortex mirrors clinical features of amyotrophic lateral sclerosis and their phenotypes: a neuroimaging study.** *J Neurol* 2013;260:2856–64 CrossRef Medline
11. Khalil M, Teunissen CE, Otto M, et al. **Neurofilaments as biomarkers in neurological disorders.** *Nat Rev Neurol* 2018;14:577–89 CrossRef Medline
12. Forgrave LM, Ma M, Best JR, et al. **The diagnostic performance of neurofilament light chain in CSF and blood for Alzheimer's disease, frontotemporal dementia, and amyotrophic lateral sclerosis: a systematic review and meta-analysis.** *Alzheimers Dement (Amst)* 2019;11:730–43 CrossRef Medline
13. Gaiani A, Martinelli I, Bello L, et al. **Diagnostic and prognostic biomarkers in amyotrophic lateral sclerosis: neurofilament light chain levels in definite subtypes of disease.** *JAMA Neurol* 2017;74:525–32 CrossRef Medline
14. Rossi D, Volanti P, Brambilla L, et al. **CSF neurofilament proteins as diagnostic and prognostic biomarkers for amyotrophic lateral sclerosis.** *J Neurol* 2018;265:510–21 CrossRef Medline
15. Benatar M, Wu J, Turner MR. **Neurofilament light chain in drug development for amyotrophic lateral sclerosis: a critical appraisal.** *Brain* 2022;146:2711–16 CrossRef Medline
16. Brodovitch A, Boucraut J, Delmont E, et al. **Combination of serum and CSF neurofilament-light and neuroinflammatory biomarkers to evaluate ALS.** *Sci Rep* 2021;11:703 CrossRef Medline
17. Verde F, Steinacker P, Weishaupt JH, et al. **Neurofilament light chain in serum for the diagnosis of amyotrophic lateral sclerosis.** *J Neurol Neurosurg Psychiatry* 2019;90:157–64 CrossRef Medline
18. Grapperon AM, Verschueren A, Jouve E, et al. **Assessing the upper motor neuron in amyotrophic lateral sclerosis using the triple stimulation technique: a multicenter prospective study.** *Clin Neurophysiol* 2021;132:2551–57 CrossRef Medline
19. Geevasinga N, van den Bos M, Menon P, et al. **Utility of transcranial magnetic stimulation in studying upper motor neuron dysfunction in amyotrophic lateral sclerosis.** *Brain Sci* 2021;11:906 CrossRef Medline
20. de Carvalho M, Swash M. **Lower motor neuron dysfunction in ALS.** *Clin Neurophysiol* 2016;127:2670–81 CrossRef Medline
21. Rutkove SB. **Clinical measures of disease progression in amyotrophic lateral sclerosis.** *Neurotherapeutics* 2015;12:384–93 CrossRef Medline
22. Vucic S, Stanley Chen KH, Kiernan MC, et al. **The clinical diagnostic utility of transcranial magnetic stimulation: report of an IFCN committee.** *Clin Neurophysiol* 2008;150:131–75 CrossRef Medline
23. de Carvalho M, Chio A, Dengler R, et al. **Neurophysiological measures in amyotrophic lateral sclerosis: markers of progression in clinical trials.** *Amyotroph Lateral Scler Other Motor Neuron Disord* 2005;6:17–28 CrossRef Medline
24. Brooks BR, Miller RG, Swash M, et al; World Federation of Neurology Research Group on Motor Neuron Diseases. **El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis.** *Amyotroph Lateral Scler Other Motor Neuron Disord* 2000;1:293–99 CrossRef Medline
25. Rascovsky K, Hodges JR, Knopman D, et al. **Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia.** *Brain* 2011;134:2456–77 CrossRef Medline
26. Oldfield RC. **The assessment and analysis of handedness: the Edinburgh inventory.** *Neuropsychologia* 1971;9:97–113 CrossRef Medline
27. Fautz H, Vogel M, Gross P, et al. **B1 mapping of coil arrays for parallel transmission.** In: *Proceedings of the International Society for Magnetic Resonance in Medicine*. May 3–9, 2008;1247. Toronto, Ontario, Canada
28. O'Brien KR, Kober T, Hagmann P, et al. **Robust T1-weighted structural brain imaging and morphometry at 7T using MP2RAGE.** *PLoS One* 2014;9:e99676 CrossRef Medline
29. Isensee F, Schell M, Pflueger I, et al. **Automated brain extraction of multisequence MRI using artificial neural networks.** *Hum Brain Mapp* 2019;40:4952–64 CrossRef Medline
30. 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
31. Menke RAL, Gray E, Lu CH, et al. **CSF neurofilament light chain reflects corticospinal tract degeneration in ALS.** *Ann Clin Transl Neurol* 2015;2:748–55 CrossRef Medline
32. Piehl F, Kockum I, Khademi M, et al. **Plasma neurofilament light chain levels in patients with MS switching from injectable therapies to fingolimod.** *Mult Scler* 2018;24:1046–54 CrossRef Medline
33. Hansson O, Janelidze S, Hall S, et al; Swedish BioFINDER study. **Blood-based NfL: a biomarker for differential diagnosis of parkinsonian disorder.** *Neurology* 2017;88:930–37 CrossRef Medline
34. Weinhofer I, Rommer P, Zierfuss B, et al. **Neurofilament light chain as a potential biomarker for monitoring neurodegeneration in X-linked adrenoleukodystrophy.** *Nat Commun* 2021;12:1816 PMC] CrossRef Medline
35. Gille B, De Schaepdryver M, Goossens J, et al. **Serum neurofilament light chain levels as a marker of upper motor neuron degeneration in patients with amyotrophic lateral sclerosis.** *Neuropathol Appl Neurobiol* 2019;45:291–304 CrossRef Medline
36. Vacchiano V, Mastrangelo A, Zenesini C, et al. **Plasma and CSF neurofilament light chain in amyotrophic lateral sclerosis: a cross-sectional and longitudinal study.** *Front Aging Neurosci* 2021;13:753242 CrossRef Medline
37. El Mendili MM, Verschueren A, Ranjeva JP, et al. **Association between brain and upper cervical spinal cord atrophy assessed by MRI and disease aggressiveness in amyotrophic lateral sclerosis.** *Neuroradiology* 2023;65:1395–403 CrossRef Medline
38. Menke RA, Körner S, Filippini N, et al. **Widespread grey matter pathology dominates the longitudinal cerebral MRI and clinical landscape of amyotrophic lateral sclerosis.** *Brain* 2014;137:2546–55 CrossRef Medline
39. van der Ende EL, Meeter LH, Poos JM, et al; Genetic Frontotemporal dementia Initiative (GENFI). **Serum neurofilament light chain in genetic frontotemporal dementia: a longitudinal, multicentre cohort study.** *Lancet Neurol* 2019;18:1103–11 CrossRef Medline
40. Bede P, Bokde A, Elamin M, et al. **Grey matter correlates of clinical variables in amyotrophic lateral sclerosis (ALS): a neuroimaging study of ALS motor phenotype heterogeneity and cortical focality.** *J Neurol Neurosurg Psychiatry* 2013;84:766–73 CrossRef Medline
41. Wirth AM, Khomenko A, Baldaranov D, et al. **Combinatory biomarker use of cortical thickness, MUNIX, and ALSFRS-R at baseline and in longitudinal courses of individual patients with amyotrophic lateral sclerosis.** *Front Neurol* 2018;9:614 CrossRef Medline
42. Catani M. **A little man of some importance.** *Brain* 2017;140:3055–61 CrossRef Medline
43. Muret D, Root V, Kieliba P, et al. **Beyond body maps: Information content of specific body parts is distributed across the somatosensory homunculus.** *Cell Rep* 2022;38:110523 CrossRef Medline
44. Schellekens W, Bakker C, Ramsey NF, et al. **Moving in on human motor cortex: characterizing the relationship between body parts with non-rigid population response fields.** *PLoS Comput Biol* 2022;18:e1009955 CrossRef Medline

45. Bergman J, Dring A, Zetterberg H, et al. **Neurofilament light in CSF and serum is a sensitive marker for axonal white matter injury in MS.** *Neurol Neuroimmunol Neuroinflamm* 2016;3:e271 CrossRef Medline
46. Geraldo AF, Pereira J, Nunes P, et al. **Beyond fractional anisotropy in amyotrophic lateral sclerosis: the value of mean, axial, and radial diffusivity and its correlation with electrophysiological conductivity changes.** *Neuroradiology* 2018;60:505–15 CrossRef Medline
47. Wagner J, Weber B, Urbach H, et al. **Morphometric MRI analysis improves detection of focal cortical dysplasia type II.** *Brain* 2011;134:2844–54 CrossRef Medline
48. Wang ZI, Alexopoulos AV, Jones SE, et al. **Linking MRI postprocessing with magnetic source imaging in MRI-negative epilepsy.** *Ann Neurol* 2014;75:759–70 CrossRef Medline
49. Lüsebrink F, Wollrab A, Speck O. **Cortical thickness determination of the human brain using high resolution 3T and 7T MRI data.** *Neuroimage* 2013;70:122–31 CrossRef Medline



# Visualizing the Habenula Using 3T High-Resolution MP2RAGE and QSM: A Preliminary Study

 BingYang Bian,  Lin Hou, YaTing Chai,  YueLuan Jiang,  XingChen Pan, Yang Sun, HongChao Wang, DongDong Qiu, ZeChen Yu, Hua Zhao, HuiMao Zhang, FanYang Meng, and  Lei Zhang



## ABSTRACT

**BACKGROUND AND PURPOSE:** The habenula is a key node in the regulation of emotion-related behavior. Accurate visualization of the habenula and its reliable quantitative analysis is vital for the assessment of psychiatric disorders. To obtain high-contrast habenula images and allow them to be compatible with clinical applications, this preliminary study compared 3T MP2RAGE and quantitative susceptibility mapping with MPRAGE by evaluating the habenula segmentation performance.

**MATERIALS AND METHODS:** Ten healthy volunteers were scanned twice with 3T MPRAGE and MP2RAGE and once with quantitative susceptibility mapping. Image quality and visibility of habenula anatomic features were analyzed by 3 radiologists using a 5-point scale. Contrast assessments of the habenula and thalamus were also performed. The reproducibility of the habenula volume from MPRAGE and MP2RAGE was evaluated by manual segmentation and the Multiple Automatically Generated Template brain segmentation algorithm (MAGeTbrain). T1 values and susceptibility were measured in the whole habenula and habenula geometric subregion using MP2RAGE T1-mapping and quantitative susceptibility mapping.

**RESULTS:** The 3T MP2RAGE and quantitative susceptibility mapping demonstrated clear boundaries and anatomic features of the habenula compared with MPRAGE, with a higher SNR and contrast-to-noise ratio (all  $P < .05$ ). Additionally, 3T MP2RAGE provided reliable habenula manual and MAGeTbrain segmentation volume estimates with greater reproducibility. T1-mapping derived from MP2RAGE was highly reliable, and susceptibility contrast was highly nonuniform within the habenula.

**CONCLUSIONS:** We identified an optimized sequence combination (3T MP2RAGE combined with quantitative susceptibility mapping) that may be useful for enhancing habenula visualization and yielding more reliable quantitative data.

**ABBREVIATIONS:** CNR = contrast-to-noise ratio; COV = coefficient of variation; FGATIR = fast gray matter acquisition T1 inversion recovery; GRE = gradient recalled-echo; Hb = habenula; ICC = intraclass correlation coefficient; MAGeTbrain = Multiple Automatically Generated Template brain segmentation algorithm; MDD = major depressive disorder; QSM = quantitative susceptibility mapping; UNI = uniform T1-weighted image

The habenula (Hb) in the human brain is a small (5- to 9-mm diameter) strategic nucleus located at the posterior dorsal-medial end of the thalamus, in front of the pineal body. The Hb commissure connects the left and right Hbs of the 2 hemispheres and locally forms the Hb trigone in front of the

posterior commissure. The Hb receives input mainly from the limbic forebrain and basal ganglia by the stria medullaris and output through the fasciculus retroflexus to the ventral tegmental area, substantia nigra, medial raphe, and raphe nucleus (Online Supplemental Data). It has been implicated in regulating behaviors such as anxiety, reward, depression, and circadian rhythm.<sup>1,2</sup> Recent research indicates that the lateral Hb plays a major role in depression, and ketamine inhibits the *N*-methyl-*D*-aspartate receptor-dependent lateral Hb, leading to a quick relief of depressive symptoms.<sup>3</sup> Moreover, the medial Hb is functionally important in regulating behavior and emotions.<sup>4</sup>

Received July 15, 2023; accepted after revision December 18.

From the Department of Radiology (B.B., L.H., Y.C., X.P., Y.S., H.W., D.Q., H. Zhang, F.M., L.Z.), Jilin Provincial Key Laboratory of Medical Imaging and Big Data, Radiology and Technology Innovation Center of Jilin Province, Jilin Provincial International Joint Research Center of Medical Artificial Intelligence, The First Hospital of Jilin University, Changchun, Jilin, China; MR Scientific Marketing, Diagnostic Imaging (Y.J.), Siemens Healthineers Ltd, Beijing, China; Siemens Healthineers Digital Technology (Shanghai) Co Ltd (Z.Y.), Shanghai, China; and Department of Physiology (H. Zhao), College of Basic Medical Sciences, Jilin University, Changchun, Jilin, China.

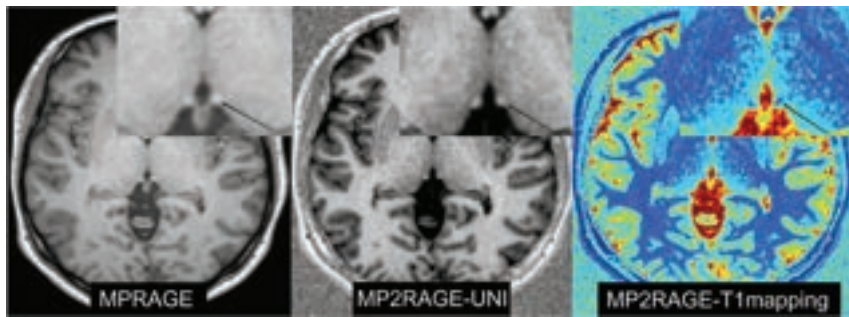
This study received funding from the National Natural Science Foundation of China (grant No. 82102143); Natural Science Foundation of Jilin Province (grant No. YDZJ20201ZYTS019); Jilin Province Science and Technology Department (grant No. YDZJ202201ZYTS679); National Natural Science Foundation of China (grant No. 32171145); Jilin Provincial Key Laboratory of Medical Imaging & Big Data (grant No. 20200601003JC); and Radiology and Technology Innovation Center of Jilin Province (grant No. 20190902016TC).

Please address correspondence to Lei Zhang, PhD, The First Hospital of Jilin University Department of Radiology, Jilin Provincial Key Laboratory of Medical Imaging and Big Data, Radiology and Technology Innovation Center of Jilin Province, Jilin Provincial International Joint Research Center of Medical Artificial Intelligence, Changchun, Jilin 130021, China; e-mail: zlei99@jlu.edu.cn



Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8156>



**FIG 1.** Representative MPRAGE, MP2RAGE-UNI, and T1-mapping images of the Hb from a healthy man 30 years of age (black arrow). Visual contrast between the Hb and the thalamus in UNI is better than that in MPRAGE.

However, its accurate depiction and reliable quantitative analysis remain challenging with current clinical MR imaging sequences. Volume changes in the Hb have long been the focus of attention in neuropsychiatric diseases (Online Supplemental Data). Evaluation of Hb volume is widely used to investigate various disorders, including schizophrenia, bipolar disorder, and major depressive disorder (MDD).<sup>5-11</sup> A 3D T1-weighted high-resolution anatomic sequence is typically used to display and manually delineate the Hb. Among these techniques, MPRAGE imaging is a representative sequence that has been extensively used for manual Hb volume segmentation. Herein, on MPRAGE, the Hb manifests as slight hyperintensity with a surface landmark that extends medially from the thalamus. However, we found that the Hb on MPRAGE had poor contrast compared with the thalamus, with the lateral boundaries separating the Hb from the thalamus, which remains poorly marked, especially after being magnified many times, possibly causing differences in the volume of the Hb when manually delineated by physicians. Furthermore, MPRAGE is vulnerable to the inhomogeneity of both transmission ( $B_1+$ ) and reception ( $B_1-$ ) radiofrequency fields at both 3T and 7T, leading to a decrease in contrast between the gray and white matter.<sup>12</sup> Meanwhile, the inhomogeneity and distortion of the main magnetic field ( $B_0$ ) may also influence image quality.<sup>13</sup> Therefore, identifying a sequence that can improve the contrast of Hb compared with other structures is critical to enhance the performance of Hb segmentation and quantitative analysis.

In recent years, the emergence of MP2RAGE has resulted in higher contrast for gray/white matter, higher reproducibility of deep gray matter, and improved accuracy of the segmentation of brain structures.<sup>14-17</sup> The MP2RAGE sequence is an extension of MPRAGE that requires a single echo and can significantly reduce image bias and prevent the effect of inhomogeneity. In MP2RAGE, 2 rapid gradient recalled-echo (GRE) images, INV1 (GRE TI1) and INV2 (GRE TI2), are acquired at different TIs and combined to produce a uniform T1-weighted image (UNI) that is free from  $B_1$  inhomogeneity.<sup>12,16</sup> Quantitative T1-mapping is also obtained, allowing accurate measurement of tissue properties. Despite the availability of the 7T MP2RAGE sequence to manually segment Hb,<sup>8,18</sup> an increase in field strength also induces nonuniformities in the transmission field

( $B_1+$ ), which can compromise image contrast nonuniformity.<sup>16</sup> Furthermore, 7T MR imaging is not widely implemented in clinical practice.

Quantitative susceptibility mapping (QSM) is another promising technique for visualizing the Hb. QSM can quantify the magnetic susceptibility of tissue, which is highly nonuniform within the Hb because the susceptibility enhancement of the Hb is associated with paramagnetic sources such as iron.

Therefore, our preliminary study aimed to achieve the following: 1) better visualize the Hb anatomic features

using 3T MP2RAGE and QSM, 2) compare the segmentation performance of MP2RAGE and MPRAGE, and 3) quantify Hb volume, T1 relaxation time, and susceptibility. The findings of this study may have important implications for improving the clinical assessment of Hb in neuropsychiatric disorders and, beyond, help guide future research in the development of normative data in this field.

## MATERIALS AND METHODS

### Participants

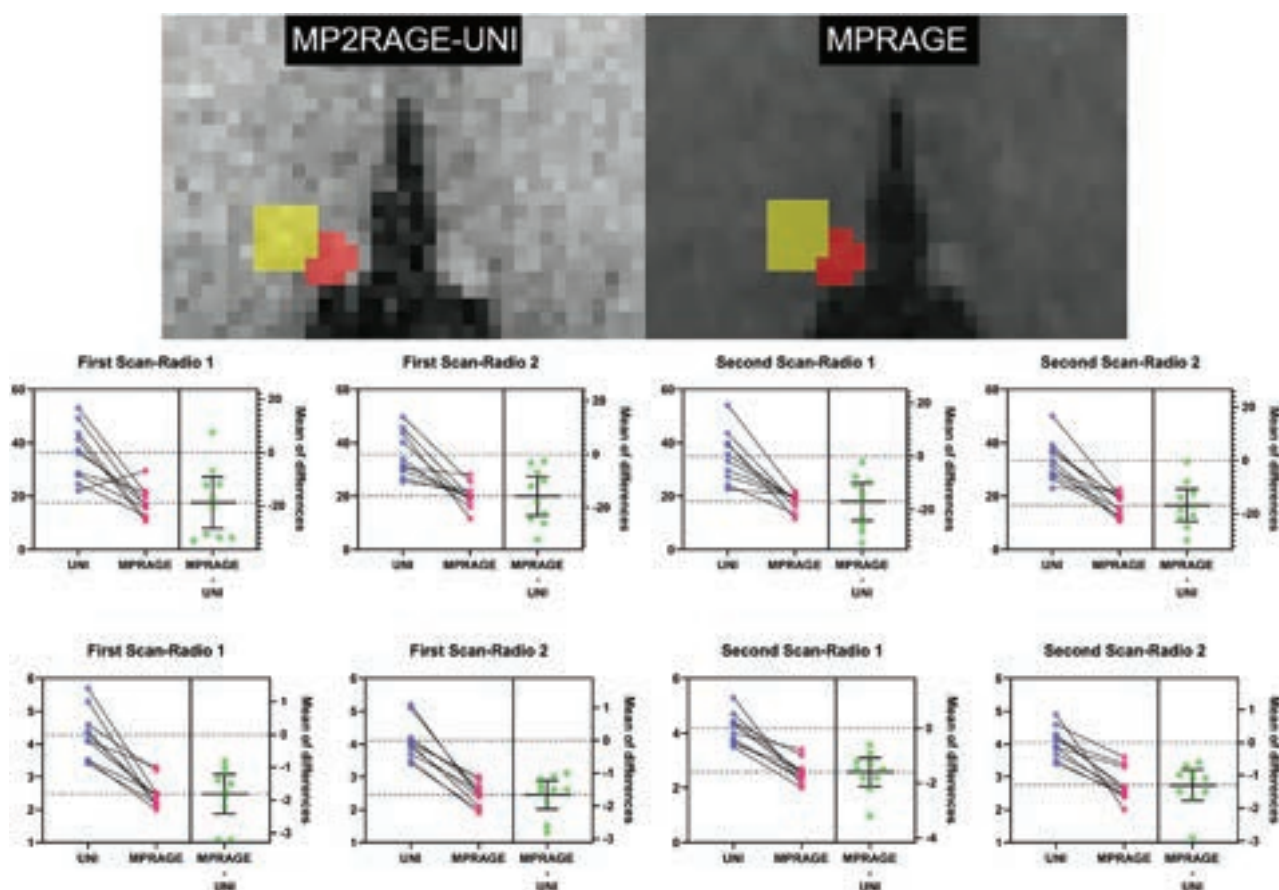
This prospective study was approved by the first hospital of Jilin University institutional review board (approval No. 22K027-001), and written informed consent was obtained from all participants. The study included 10 healthy volunteers (25–40 years of age; mean age, 31.5 [SD, 4.4] years; 5 men and 5 women) with no history of severe head trauma, tumor or neuropsychiatric disorders, large-vessel disease, and/or diseases with white matter lesions.

### Imaging Acquisition

All participants were examined using a 3T MR imaging scanner (Magnetom Vida; Siemens) with a 20-channel head-and-neck coil. The scans were obtained on the sagittal plane parallel to the median sagittal position line. Earplugs and foam cushions were provided to ensure the participants' comfort and reduced motion artifacts. The MPRAGE, MP2RAGE, and QSM acquisition parameters are listed in the Online Supplemental Data. MPRAGE and MP2RAGE were acquired twice for each participant, with a 1-week interval between scans. In contrast, QSM was performed after the second MP2RAGE and MPRAGE scans.

### Image Analysis

MP2RAGE raw data were obtained using a Syngo Via-software MR imaging workstation (Siemens) to calculate UNI and T1 mapping (Fig 1). Manual segmentation of bilateral Hbs on MP2RAGE, MPRAGE, and QSM was performed in ITK-SNAP ([www.itksnap.org](http://www.itksnap.org)). According to previous studies, the Hb boundaries were traced in the region of the triangular depression of the third ventricle, starting with the section above the posterior commissure and moving in the superior direction. The Hb was delineated as long as it protruded into the third ventricle or showed a visible contrast difference from the surrounding thalamus.



**FIG 2.** The Hb ROI was chosen on the basis of the largest axial image on MP2RAGE-UNI and MPRAGE, and the thalamus ROI was defined as twice the size of the Hb ROI, which is located on the anterolateral side of the Hb ROI (*first row*). The CNR (*middle row*) and SNR (*third row*) of UNI is statistically higher than those of MPRAGE (paired samples *t* test). Radio indicates radiologist.

### Subjective Visual Score

The qualitative scores of the Hb in MPRAGE and MP2RAGE-UNI of the first scan, as well as QSM, were assessed by 3 radiologists (B.B., Z.L., and Z.H.M. with 8, 15, and 25 years of experience, respectively) blinded to participant information. To standardize the subjective assessment, 3 radiologists were directed to reach a consensus on the image series. In case of disagreement in the image analysis, a consensus was reached among the 3 observers. The radiologists reported their impressions of the overall image quality on a scale of 1 (unusable) to 5 (excellent). The scores were defined as the following: 1) The Hb was not visible; 2) the Hb was poorly visible with fuzzy borders; 3) the Hb was clearly differentiable from the third ventricle but not from the thalamus; 4) the Hb was better visible with clear borders; and 5) the Hb was well-defined and clearly differentiable.

### Image Quality Analysis

The section in which the Hb showed the largest axial image on MP2RAGE-UNI and MPRAGE was chosen for quantitative analysis. The Hb ROI as defined as the intersection of the manual segmentation results of 2 radiologists. The thalamic ROI was defined as twice the size of the Hb ROI and was in the anterolateral region of the Hb ROI (Fig 2). The SNR of the Hb and the contrast-to-noise ratio (CNR) of the Hb and the thalamus were

calculated on the basis of MP2RAGE-UNI and MPRAGE, using the following equations:

$$\begin{aligned} \text{SNR}_{\text{Hb}} &= \text{SI}_{\text{Hb}} / \text{SD}_{\text{Hb}} \text{ and } \text{CNR}_{\text{Hb-thalamus}} \\ &= |\text{SI}_{\text{Hb}} - \text{SI}_{\text{thalamus}}| / \sqrt{\text{SD}_{\text{Hb}}^2 + \text{SD}_{\text{thalamus}}^2}, \end{aligned}$$

where SI is the signal intensity and noise is the SD. The mean values of the bilateral Hb and thalamus were calculated by 2 radiologists, with 8 and 25 years of experience in using ITK-SNAP.

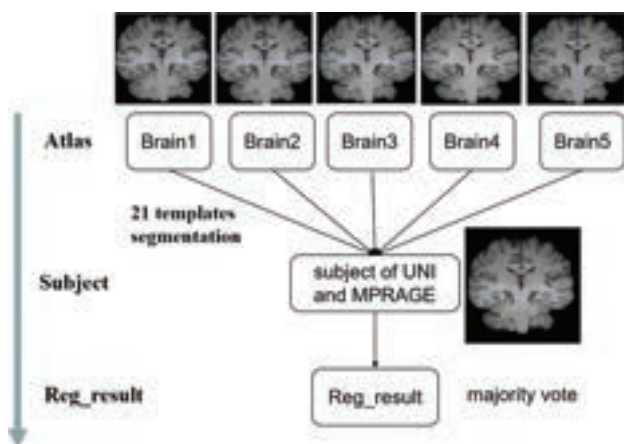
### Segmentation and Validation Evaluation

Manual segmentation of the Hb was performed on MP2RAGE-UNI and MPRAGE zoomed by a factor of 8 using ITK-SNAP software by 2 independent radiologists (B.B. and Z.L.) to obtain an ROI. A consensus segmentation was reached by the radiologists to offer the best delineation. The left and right Hbs were manually labeled using ITK-SNAP on all MPRAGE and UNI images, and volume estimations were calculated from manual labels by calculating the number of labeled voxels for each healthy subject. The segmentation was repeated twice with a 7-day interval for MP2RAGE-UNI and MPRAGE, respectively, and the changes in the number of labeled voxels and absolute volume were recorded each time (Online Supplemental Data). Manual segmentation of the Hbs in these participants was then repeated twice (the interval

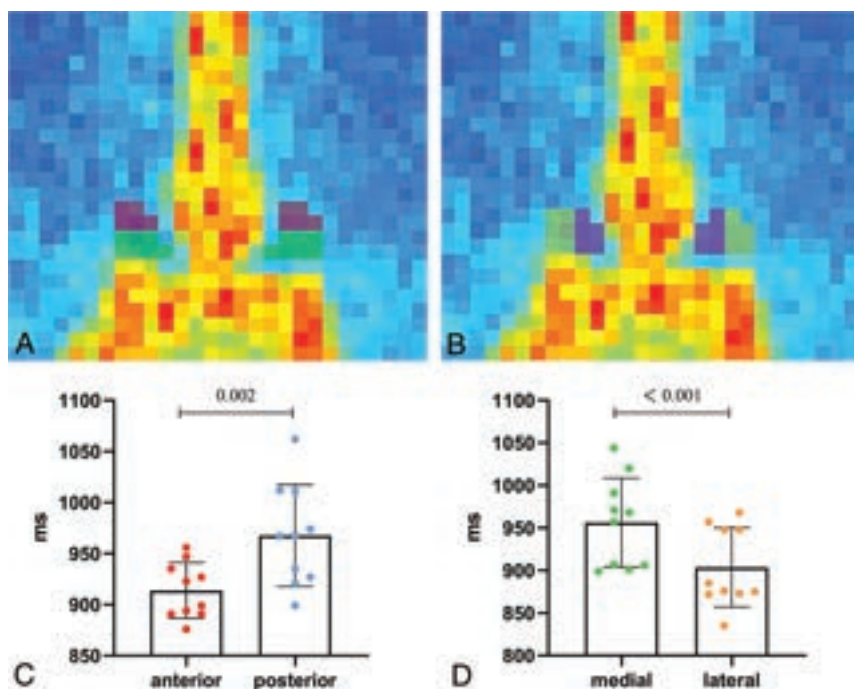


between the 2 times was 7 days) using UNI and MPAGE, and changes in the number of labeled voxels and absolute volume were recorded each time.

A fully automated Hb atlas segmentation based on the Multiple Automatically Generated Template brain segmentation algorithm (MAGeTbrain) (<http://cobralab.ca/software/MAGeTbrain/>) was



**FIG 3.** Segmentations from 5 atlases are propagated via image registration to 21 selected template images before being propagated using image registration to each UNI and MPAGE. This feature creates a large number ( $5 \text{ atlases} \times 21 \text{ templates} = 105$ ) of candidate segmentations that are fused using majority vote. This has 2 advantages: reducing atlas bias and averaging registration errors. Also, because the procedure is based on image registration rather than on contrast-based segmentation, it allows the neuroanatomic variability of the subjects to refine each individual subject's final segmentation. Reg\_result indicates registration result.



**FIG 4.** ROI on the T1 mapping from the left-to-right Hb of a representative participant. A, Anterior (red) and posterior (green) subregions within the habenula. B, Medial (pink) and lateral (yellow) subregions within the Hb. C and D, The T1 values of the posterior halves of the Hb are higher than those in the anterior halves, and the medial halves are higher than those in the lateral halves.

also applied. Full Hb atlases (<https://github.com/CobraLab/atlas>) provide a well-established method for Hb segmentation (Fig 3). The MP2RAGE-UNI and MPAGE images were automatically segmented using the MAGeTbrain. Subsequently, interrater reliability was assessed manually by radiologist 1 (B.B.) versus MAGeTbrain and radiologist 2 (L.Z.) versus MAGeTbrain.<sup>19</sup>

### Quantitative Analysis

For quantitative analysis, the section in which the Hb showed the largest axial image of T1-mapping and QSM was selected. Two methods were used for analysis: The whole Hb was used to measure T1 values and susceptibility, and the Hb was geometrically divided into halves (the anterior-posterior and medial/lateral subregions) (Fig 4A, -B and Online Supplemental Data), to measure the T1 values and the susceptibility of each subregion. The QSM image reconstruction process is illustrated in the Online Supplemental Data.

### Statistical Analysis

All statistical analyses were performed using SPSS software (Version 21.0; IBM). To evaluate the confidence of the Hb ROI, we conducted intraclass correlation coefficient (ICC) analysis for the subjective visual score, SNR, CNR, T1 values, and susceptibility in the 3 sequences. The SNR and CNR of the Hb were compared between the MPAGE and MP2RAGE-UNI groups using paired *t* tests. The agreement between the MP2RAGE-UNI and MPAGE Hb volume estimates from the manual segmentation and MAGeTbrain was assessed using Bland-Altman plots. To assess the reproducibility of the T1 values from MP2RAGE, we calculated the coefficient of variation (COV) of the T1 value of

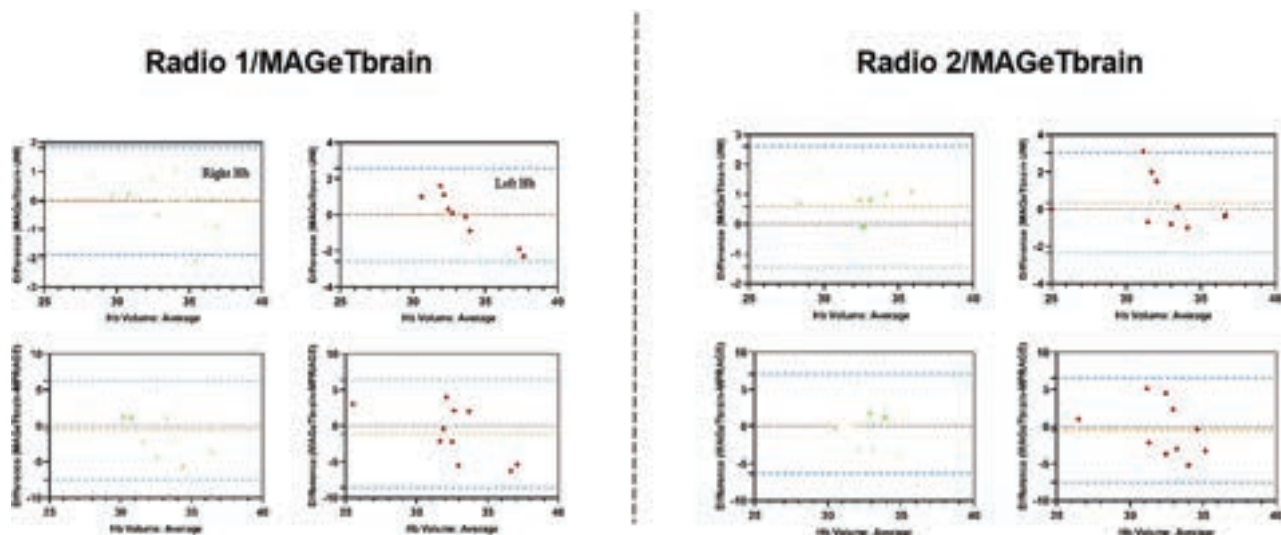
Hb by dividing the SD of each Hb level of the T1 value by the mean (expressed as a percentage). T1 values and the susceptibility of Hb geometric subregions were compared using a paired *t* test. A *P* value < .05 was considered statistically significant.

## RESULTS

### Subjective Visual Assessment

The ICCs of the qualitative scores among the 3 radiologists were 0.870, 0.602, and 0.859 for MP2RAGE-UNI, MPAGE, and QSM, respectively. The Hb appeared slightly hyperintense on MPAGE, as well as hyperintense on MP2RAGE and QSM. In MP2RAGE, stria medullaris and fasciculus retroflexus can be readily distinguished, compared with MPAGE (Online Supplemental Data). The average subjective visual scores for the Hb by the 3 radiologists for the MP2RAGE-UNI, MPAGE, and QSM images were 4.83, 4.20, and 4.63, respectively. Higher image contrast and optimal geometry were obtained from MP2RAGE-UNI





**FIG 5.** Bland-Altman plot showing the volume differences of the right Hb (green dots) and left Hb (red dots) between the MAGEtbrain and manual segmentations in the first scan for MP2RAGE-UNI and MP2RAGE. Radio indicates radiologist.

images than from MP2RAGE and QSM images (Online Supplemental Data), which was unanimously recognized. Furthermore, both MP2RAGE-UNI and QSM contrast images could be easily used to delineate the well-defined boundaries between the Hb and thalamus.

#### **Difference in SNR and CNR of Hb**

The ICCs of the SNR and CNR of the Hb for the first and second scans are presented in the Online Supplemental Data. The mean SNR of the Hb was 36.39 (SD, 10.01) and 17.55 (SD, 5.40) on the MP2RAGE-UNI and the MP2RAGE images, respectively. In addition, the CNRs between the Hb and thalamus were 4.29 (SD, 0.72) and 2.48 (SD, 0.42) on the MP2RAGE-UNI and the MP2RAGE images, respectively. Compared with MP2RAGE, significantly higher  $SNR_{Hb}$  and  $CNR_{Hb-thalamus}$  were detected in MP2RAGE-UNI (all  $P < .05$ ), as shown in Fig 2.

#### **Difference in Hb Volume**

The mean volumes of the right and left Hbs from the 10 volunteers after manual segmentation using the ITK-SNAP software are presented in the Online Supplemental Data. The UNI-based manual volume estimates were highly correlated with MP2RAGE for both the left and right Hbs in the first and second scans ( $P < .001$  for all simple linear regression models). The COVs in the right and left Hb volumes of the MAGEtbrain obtained from the UNI and MP2RAGE images were low at 5.4%, 6.3%, 9.7%, and 8.6%, respectively. Figure 5 depicts a greater similarity in Hb volume measurements, including manual and MAGEtbrain segmentation, between the UNI and MP2RAGE images.

#### **T1 Values Measured by MP2RAGE T1-Mapping and Their COVs**

The ICC of the T1 values of the Hb ranged from 0.884 (right Hb on the first scan) to 0.991 (right Hb on the second scan) (Online Supplemental Data). The measured T1 values of the 10 healthy volunteers are summarized in the Online Supplemental Data. The COVs of the T1 values ranged from 3.4% to 5.8% in the first

and second MP2RAGE scans. The T1 values were not significantly different between the left and right Hbs (all  $P > .05$ ; Online Supplemental Data). However, the bilaterally averaged T1 values obtained from the Hb subregion for the 10 participants varied from 835 to 1044 ms. The difference between the anterior and posterior halves, as well as the lateral and medial halves, was statistically significant ( $P = .002$  and  $< .001$ , paired  $t$  test) (Fig 4).

#### **Susceptibility Measured by QSM**

The ICC for susceptibility to the Hb was 0.989 for the right side and 0.998 for the left side and was not significantly different between the left and right Hbs ( $P = .67$ ). The mean susceptibility for the right and left Hbs was 15.74 (SD, 1.88) ppb and 15.83 (SD, 1.83) ppb, respectively. However, the posterior halves of the Hb were significantly higher than the anterior halves ( $P < .001$ , paired  $t$  test). The difference between the medial and lateral halves was not significant (paired  $t$  test,  $P = .467$ ) (Online Supplemental Data).

## **DISCUSSION**

The preliminary results indicate that MP2RAGE-UNI yielded superior visual scores, SNR, CNR, and reliability of Hb volume segmentation when directly compared with MP2RAGE. MP2RAGE T1-mapping and QSM have the potential to notably advance quantitative analysis aimed at reflecting changes in the Hb composition. For Hb MR imaging, a homogeneous data set is required. Thus, we used a narrow age range for recruiting the healthy volunteers because it is well-documented that with aging, many structures in the brain accumulate iron and undergo atrophy.<sup>20</sup> Future work should aim to collect needed data for different age groups, categorized by decade. This information would be beneficial for current age groups in terms of future references and comparisons with disease processes.

For patients with posttraumatic stress disorder, MDD, schizophrenia, and bipolar disorder, the volume of the Hb is the first concern (Online Supplemental Data). Savitz et al<sup>5</sup> demonstrated

that Hb volume increases in patients with MDD relative to controls by a 3T MP2RAGE sequence. Schmidt et al<sup>6</sup> reported no significant difference in the volume of Hb between unmedicated and medicated patients with MDD by a 7T MP2RAGE sequence. In our study, we used bias-corrected T1-weighted images (MP2RAGE-UNI) and QSM to visualize the Hb of healthy volunteers. The preliminary results demonstrated that the MP2RAGE-UNI sequence provides superior tissue contrast between the thalamus and the Hb compared with MP2RAGE, allowing improved subjective and objective assessment. Additionally, both manual and automatic segmentation methods yielded highly reliable results for Hb volume estimation. With excellent tissue contrast, the Bland-Altman plot shows good consistency between MAGETbrain and manual segmentations. Additionally, COV in the right and left Hb volumes of MAGETbrain obtained from the UNI were low (COV: 5.4% and 6.3%, respectively). Our COV results are better than those of previous semi-/fully automatic segmentations based on the myelin content in T1-weighted and T2-weighted images.<sup>21</sup> Automated segmentation-based AI is accurate enough for use compared with manual segmentation. Future AI applications are expected to support and facilitate the establishment of an accurate Hb automated segmentation model, reducing the time spent by radiologists to manually sketch the Hb and avoiding errors caused by subjectivity.

Our study also found no significant differences in T1 values between the left and right Hb, which was inconsistent with the findings of Cho et al.<sup>22</sup> Significant differences were observed between the anterior and posterior halves as well as the medial and lateral halves. The T1-relaxation time is a measure of how quickly the longitudinal magnetization recovers to its ground state. An increase in the volume fraction of myelin reduces the average T1 of white matter. Quantitative analyses of T1 values in the subregions of the human Hb are limited, and the interpretation of this imbalance remains unclear. The MP2RAGE sequence, which provides purely T1-weighted images, offers a promising foundation for subcortical nucleus segmentation. Our results provide further evidence supporting the use of MP2RAGE for accurate segmentation of Hb and exploring T1 value differences within the Hb and its subregions between healthy subjects and those with neuropsychiatric disorders.

Sands et al<sup>23</sup> found that iron increased in the medial Hbs of cerebral mouse models of MS. However, our study found that susceptibility differences also existed in different subregions of the human Hb, with the posterior subregion showing increased susceptibility. Yoo et al<sup>24</sup> previously reported that an increase in susceptibility is associated with iron deposition rather than a decrease in myelin and that the posterior region tends to show increased susceptibility, a finding consistent with our results. Wang et al<sup>25</sup> observed that late-life depression progression is associated with increased iron deposition in the whole Hb, but not the lateral Hb. Our results contribute to further studies investigating whether depression progression is associated with iron deposition in the Hb, particularly in the posterior subregion.

Although our results are not completely in accordance with those of previous studies, our ICC was generally higher than 0.8, indicating excellent reliability. This can be because MP2RAGE provides ideal tissue contrast and clear boundaries. Previous ex

vivo imaging of the Hb using MP2RAGE at 7T proved the heterogeneity of the lateral Hb.<sup>26</sup>

## Limitations

This study had some limitations: first, the small sample size, with only 10 cases in the 25- to 40-year age range. Furthermore, a multicenter study would have enhanced statistical representativeness. For our future research, we scanned volunteers of different age groups and sexes to obtain a range of Hb values, which will serve as the basis for normative data in a healthy-aging population. Second, MP2RAGE and QSM, respectively, are definitely more than twice and triple the duration to obtain MP2RAGE, possibly limiting its widespread application in psychiatric patients. We are actively applying artificial intelligence to improve the acquired speed of MP2RAGE, QSM, and denoising. Third, the resolution of QSM appears challenging for achieving accurate quantitative evaluation due to Hb being an extremely small brain structure. In addition, veins in and around the Hb interfered with magnetic susceptibility characterization. Fourth, the fast gray matter acquisition T1 inversion recovery (FGATIR) sequence more clearly illustrates boundaries between the subcortical nuclei and is frequently used for electrode placement.<sup>27</sup> We hope to use FGATIR and MP2RAGE to determine the visualization of the optimal sequence of the Hb. Fifth, other structures in the Hb pathway, such as the interpeduncular nucleus, ventral tegmental area of the midbrain, and raphe nuclei, were not visualized. A multiparametric MR imaging protocol to visualize multiple reward-pathway structures may be valuable for functional neurosurgery procedures in the future.

## CONCLUSIONS

Our study demonstrates the utility of MP2RAGE-UNI and QSM sequences for the visualization and quantitative analysis of the Hb in healthy subjects. The high tissue contrast provided by the MP2RAGE-UNI sequence enables reliable and reproducible Hb segmentation. Thus, MP2RAGE and QSM sequences have the potential for clinical applications in the study of psychiatric disorders.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

1. Hu H, Cui Y, Yang Y. **Circuits and functions of the lateral habenula in health and in disease.** *Nat Rev Neurosci* 2020;21:277–95 CrossRef Medline
2. Yang Y, Cui Y, Sang K, et al. **Ketamine blocks bursting in the lateral habenula to rapidly relieve depression.** *Nature* 2018;554:317–22 CrossRef Medline
3. Gold PW, Kadriu B. **A major role for the lateral habenula in depressive illness: physiologic and molecular mechanisms.** *Front Psychiatry* 2019;10:320 CrossRef Medline
4. Viswanath H, Carter AQ, Baldwin PR, et al. **The medial habenula: still neglected.** *Front Hum Neurosci* 2013;7:931 CrossRef Medline
5. Savitz JB, Rauch SL, Drevets WC. **Reproduced from Habenula volume in bipolar disorder and major depressive disorder: a high-resolution magnetic resonance imaging study.** *Mol Psychiatry* 2013;18:523 CrossRef Medline
6. Schmidt FM, Schindler S, Adamidis M, et al. **Habenula volume increases with disease severity in unmedicated major depressive**

- disorder as revealed by 7T MRI. *Eur Arch Psychiatry Clin Neurosci* 2017;267:107–15 CrossRef Medline
7. Carceller-Sindreu M, de Diego-Adelino J, Serra-Blasco M, et al. **Volumetric MRI study of the habenula in first episode, recurrent and chronic major depression.** *Eur Neuropsychopharmacol* 2015;25:2015–21 CrossRef Medline
8. Savitz JB, Nugent AC, Bogers W, et al. **Habenula volume in bipolar disorder and major depressive disorder: a high-resolution magnetic resonance imaging study.** *Biol Psychiatry* 2011;69:336–43 CrossRef Medline
9. Luan SX, Zhang L, Wang R, et al. **A resting-state study of volumetric and functional connectivity of the habenular nucleus in treatment-resistant depression patients.** *Brain Behav* 2019;9:e01229 CrossRef Medline
10. Zhang L, Wang H, Luan S, et al. **Altered volume and functional connectivity of the habenula in schizophrenia.** *Front Hum Neurosci* 2017;11:636 CrossRef Medline
11. Bocchetta M, Gordon E, Marshall CR, et al. **The habenula: an under-recognised area of importance in frontotemporal dementia?** *J Neurol Neurosurg Psychiatry* 2016;87:910–12 CrossRef Medline
12. Marques JP, Kober T, Krueger G, et al. **MP2RAGE, a self bias-field-corrected sequence for improved segmentation and T1-mapping at high field.** *Neuroimage* 2010;49:1271–81 CrossRef Medline
13. Fujimoto K, Polimeni JR, van der Kouwe AJ, et al. **Quantitative comparison of cortical surface reconstructions from MP2RAGE and multi-echo MPRAGE data at 3 and 7 T.** *Neuroimage* 2014;90:60–73 CrossRef Medline
14. Trotier AJ, Dilharreguy B, Anandra S, et al. **The compressed sensing MP2RAGE as a surrogate to the MPRAGE for neuroimaging at 3 T.** *Invest Radiol* 2022;57:366–78 CrossRef Medline
15. Ferraro PM, Gualco L, Costagli M, et al. **Compressed sensing (CS) MP2RAGE versus standard MPRAGE: a comparison of derived brain volume measurements.** *Phys Med* 2022;103:166–74 CrossRef Medline
16. Oliveira I, Roos T, Dumoulin SO, et al. **Can 7T MPRAGE match MP2RAGE for gray-white matter contrast?** *Neuroimage* 2021;240:118384 CrossRef Medline
17. Okubo G, Okada T, Yamamoto A, et al. **MP2RAGE for deep gray matter measurement of the brain: a comparative study with MPRAGE.** *J Magn Reson Imaging* 2016;43:55–62 CrossRef Medline
18. Lim SH, Yoon J, Kim YJ, et al. **Reproducibility of automated habenula segmentation via deep learning in major depressive disorder and normal controls with 7 Tesla MRI.** *Sci Rep* 2021;11:13445 CrossRef Medline
19. Germann J, Gouveia FV, Martinez R, et al. **Fully automated habenula segmentation provides robust and reliable volume estimation across large magnetic resonance imaging datasets, suggesting intriguing developmental trajectories in psychiatric disease.** *Biol Psychiatry Cogn Neurosci Neuroimaging* 2020;5:923–29 CrossRef Medline
20. Dusek P, Hofer T, Alexander J, et al. **Cerebral iron deposition in neurodegeneration.** *Biomolecules* 2022;12:714 CrossRef Medline
21. Kim JW, Naidich TP, Joseph J, et al. **Reproducibility of myelin content-based human habenula segmentation at 3 Tesla.** *Hum Brain Mapp* 2018;39:3058–71 CrossRef Medline
22. Cho SE, Park CA, Na KS, et al. **Left-right asymmetric and smaller right habenula volume in major depressive disorder on high-resolution 7-T magnetic resonance imaging.** *PLoS One* 2021;16:e255459 CrossRef Medline
23. Sands SA, Tsau S, LeVine SM. **The habenula and iron metabolism in cerebral mouse models of multiple sclerosis.** *Neurosci Lett* 2015;606:204–08 CrossRef Medline
24. Yoo S, Kim JW, Schenck JF, et al. **Magnetic susceptibility imaging of human habenula at 3 T.** *Sci Rep* 2020;10:19357 CrossRef Medline
25. Wang F, Zhang M, Li Y, et al. **Alterations in brain iron deposition with progression of late-life depression measured by magnetic resonance imaging (MRI)-based quantitative susceptibility mapping.** *Quant Imaging Med Surg* 2022;12:3873–88 CrossRef Medline
26. Strotmann B, Kogler C, Bazin PL, et al. **Mapping of the internal structure of human habenula with ex vivo MRI at 7T.** *Front Hum Neurosci* 2013;7:878 CrossRef Medline
27. Tao S, Zhou X, Westerhold EM, et al. **Optimization of fast gray matter acquisition T1 inversion recovery (FGATIR) on 7T MRI for deep brain stimulation targeting.** *Neuroimage* 2022;252:119043 CrossRef Medline

# Does CISS MRI Reliably Depict the Endolymphatic Duct in Children with and without Vestibular Aqueduct Enlargement?

Olutayo I. Olubiye, Nicholas Thompson, Thad Benefield, Kassie L. McCullagh, and Benjamin Y. Huang



## ABSTRACT

**BACKGROUND AND PURPOSE:** High-resolution CT is the mainstay for diagnosing an enlarged vestibular aqueduct (EVA), but MR imaging may be an appealing alternative, given its lack of ionizing radiation exposure. The purpose of this study was to determine how reliably MR imaging demonstrates the endolymphatic duct and endolymphatic duct enlargement in hearing-impaired children.

**MATERIALS AND METHODS:** We performed a retrospective review of temporal bone high-resolution CT and MR imaging of hearing-impaired children evaluated between 2017 and 2020. Vestibular aqueduct diameter was measured on high-resolution CT. The vestibular aqueducts were categorized as being enlarged (EVA+) or nonenlarged (EVA-) using the Cincinnati criteria. The endolymphatic ducts were assessed on axial high-resolution CISS MR imaging. We categorized endolymphatic duct visibility into the following: type 1 (not visible), type 2 (faintly visible), and type 3 (easily visible). Mixed-effect logistic regression was used to identify associations between endolymphatic duct visibility and EVA. Interreader agreement for the endolymphatic duct among 3 independent readers was assessed using the Fleiss  $\kappa$  statistic.

**RESULTS:** In 196 ears from 98 children, endolymphatic duct visibility on MR imaging was type 1 in 74.0%, type 2 in 14.8%, and type 3 in 11.2%; 20.4% of ears were EVA+ on high-resolution CT. There was a significant association between EVA+ status and endolymphatic duct visibility ( $P < .01$ ). Endolymphatic duct visibility was type 1 in 87.1%, type 2 in 12.8%, and type 3 in 0% of EVA- ears and type 1 in 22.5%, type 2 in 22.5%, and type 3 in 55.0% of EVA+ ears. The predicted probability of a type 3 endolymphatic duct being EVA+ was 0.997. There was almost perfect agreement among the 3 readers for distinguishing type 3 from type 1 or 2 endolymphatic ducts.

**CONCLUSIONS:** CISS MR imaging substantially underdiagnoses EVA; however, when a type 3 endolymphatic duct is evident, there is a >99% likelihood of an EVA.

**ABBREVIATIONS:** ELD = endolymphatic duct; ELS = endolymphatic sac; EVA = enlarged vestibular aqueduct; HRCT = high-resolution CT; IAC = internal auditory canal; SNHL = sensorineural hearing loss; VA = vestibular aqueduct

Imaging constitutes an important part of the evaluation of children with hearing loss, with the primary imaging modalities being high-resolution CT (HRCT) of the temporal bone and focused MR imaging evaluation of the inner ear and internal auditory canals (IACs).<sup>1</sup> Enlargement of the vestibular aqueduct (VA) is among the most common abnormal imaging findings identified in patients with congenital deafness, and it is implicated in approximately 12% of children with congenital sensorineural hearing loss (SNHL).<sup>2,3</sup>

The diagnosis of an enlarged vestibular aqueduct (EVA) is important to explain audiologic findings, to allow appropriate genetic testing and counseling, to advise patients of the risk of progressive hearing loss from participation in contact sports, to predict the presence of other labyrinthine anomalies on imaging, and to determine potential risks of cochlear implantation such as a perilymphatic gusher.<sup>4</sup> Traditionally, HRCT has been the mainstay for imaging evaluation of patients with hearing loss and for diagnosing EVA. While an MR imaging correlate to EVA, the enlarged endolymphatic duct syndrome, has been described,<sup>5</sup> MR imaging has been largely used as an adjunct tool in evaluating these patients.<sup>6-8</sup> However, MR imaging has several potentially compelling advantages over CT, including its lack of ionizing radiation exposure as well as its improved ability to demonstrate perilymph in the inner ear and structures not directly visualized on HRCT such as the cochleovestibular nerve.<sup>9</sup> Being able to perform a complete imaging assessment of the inner ear with a single

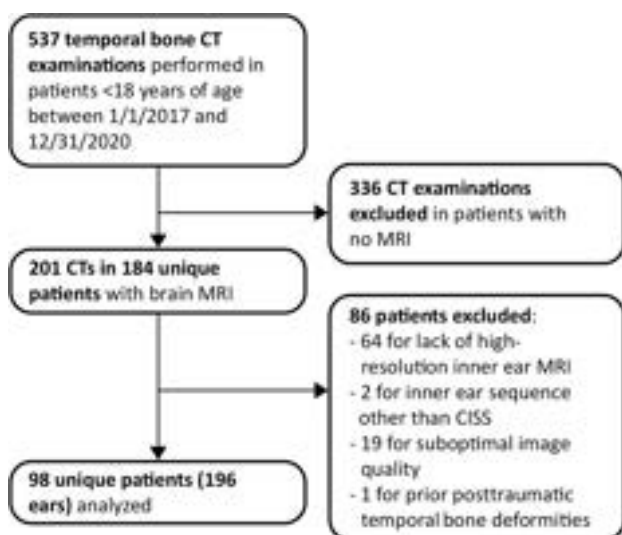
Received May 24, 2023; accepted after revision December 23.

From the Division of Neuroradiology, Department of Radiology (O.I.O.), Department of Otolaryngology (N.T.), and Department of Radiology (T.B., K.L.M., B.Y.H.), University of North Carolina School of Medicine, Chapel Hill, North Carolina; Commonwealth Radiology PC (O.I.O.), Richmond, Virginia; and University of North Carolina Hospitals (K.L.M.), Chapel Hill, North Carolina.

Please address correspondence to Olutayo I. Olubiye, MB, ChB, MPH, Commonwealth Radiology PC, Radiology, St. Mary's Hospital, 5801 Bremono Rd, Richmond, Richmond, VA 23226; e-mail: oio061@mail.harvard.edu

<http://dx.doi.org/10.3174/ajnr.A8158>





**FIG 1.** Flow chart summarizing identification, selection, and exclusion of subjects from this study.

test while avoiding radiation exposure in children with SNHL or suspected EVA would be a highly appealing proposition.

To date, several studies have investigated the utility of MR imaging compared with HRCT for diagnosing EVA;<sup>4,10-13</sup> however, there is relatively sparse data<sup>13</sup> on how reliably MR imaging is able to depict certain endolymph-containing structures, namely the endolymphatic duct (ELD) and the endolymphatic sac (ELS). The purpose of this study was to establish the following: 1) how often the ELD is visualized on routine high-resolution 3D fluid-sensitive MR imaging of the temporal bones, and 2) how reliably MR imaging can detect VA enlargement in children referred for imaging of hearing loss.

## MATERIALS AND METHODS

This retrospective study was approved by our center's institutional review board (University of North Carolina) and approved with a waiver of the requirement for informed consent. Patients younger than 18 years of age who had a temporal bone HRCT on our PACS between January 2017 and December 2020 were initially identified. These patients were screened to identify those who also had a brain MR imaging that included high-resolution 3D fluid-sensitive sequences tailored to assess the inner ear structures. Patients were excluded if there was no MR imaging available for viewing on our PACS, if the MR imaging did not include high-resolution inner ear images, if either the HRCT or MR imaging studies were deemed suboptimal for evaluation due to excessive artifacts, or if there were posttraumatic deformities of the otic capsule. Because our routine temporal bone MR imaging protocol during the study period overwhelmingly used a CISS sequence for assessment of the inner ears, examinations that used sequences other than CISS were also excluded from our analysis to avoid potential heterogeneity related to use of multiple different pulse sequences. The algorithm used to identify subjects for the study is summarized in Fig 1.

A clinical chart review of each subject's electronic medical record was performed by a neurotology fellow, who recorded

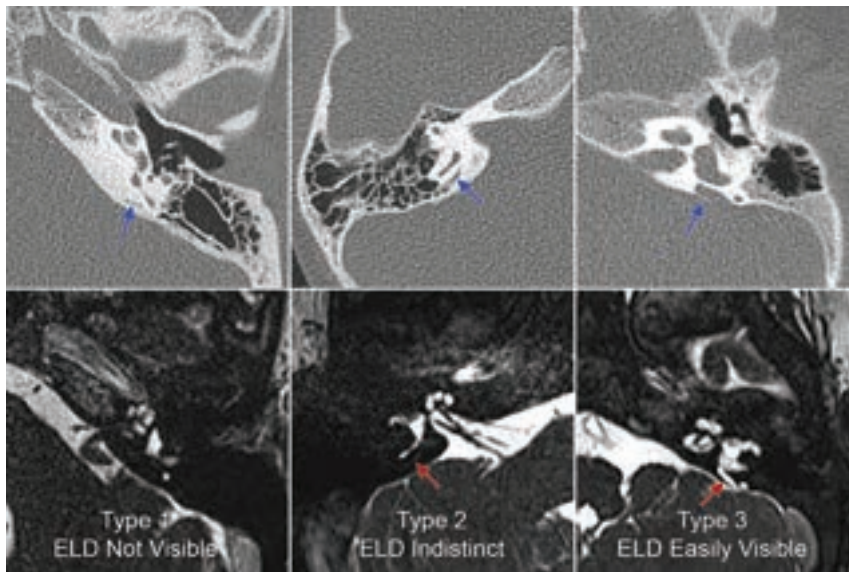
patient demographic and clinical information including available audiometric data for each ear. The HRCT and MR images were initially reviewed by a first-year neuroradiology fellow who was blinded to the clinical data. All the scans and imaging data were subsequently overread by a senior attending neuroradiologist with >15 years of experience who was also blinded to the clinical data. Any discrepancies between the reviewers' interpretations were reconciled at a consensus review session. For the initial MR imaging review performed by the fellow and the senior attending physician, the readers were not blinded to the CT findings.

All temporal bone CT images were acquired axially and reconstructed at a 0.6-mm section thickness. To account for variations in patient head positioning, we reconstructed reformatted images of each temporal bone in the plane parallel to that of the ipsilateral horizontal semicircular canal (termed the "true axial plane") using the built-in Volume Viewing tool on our PACS. A second Pöschl view reconstruction was created along an approximately 45° oblique plane, parallel to the ipsilateral superior semicircular canal. The final VA evaluation was performed on magnified views of these reconstructed image sets in a standard bone window (width = 3900, level = 150). VA diameter in each ear was measured at the midpoint and operculum in the true axial plane according to the method described by Boston et al<sup>14</sup> and at the midpoint in the 45° oblique Pöschl plane, similar to the technique described by Juliano et al.<sup>15</sup> Each VA was then categorized as being either enlarged (EVA+) or normal (EVA-) based on the criteria proposed by Boston et al (the Cincinnati criteria), which considers the VA to be enlarged if its midpoint width in the axial plane is  $\geq 1.0$  mm or its operculum width is  $\geq 2.0$  mm.<sup>14</sup>

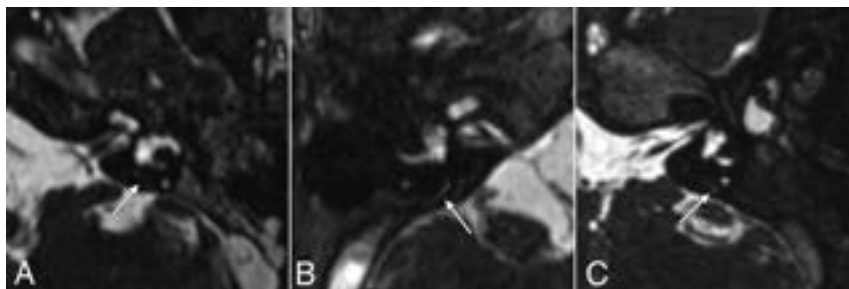
MR imaging examinations included 60 scans (61.2%) performed on a 3T system; and 38 scans (38.8%), on a 1.5T system. All MR images included high-resolution CISS sequences of the temporal bones. The high-resolution sequences featured isotropic or near-isotropic voxel widths ranging from 0.5 to 0.8 mm, 70.4% of which were 0.5 mm, 5.1% were 0.6 mm, and 23.4% were 0.7 mm. The transverse FOV for the CISS images was set to include both temporal bones in a single acquisition and ranged from 109 to 187 mm, with 90.0% being between 135 and 150 mm.

For each ear, the CISS images were subjectively evaluated for the visibility of the ELD, which was categorized into 1 of 3 types: type 1, the ELD not visible; type 2, the ELD faintly visible but indistinct and not subjectively enlarged; and type 3, the ELD distinct and easily visible or subjectively enlarged (Fig 2). Subjective enlargement was determined by the apparent width of the ELD clearly being greater than that of the ipsilateral lateral semicircular duct. Additional examples of type 2 and 3 ELD visibility are demonstrated in Figs 3 and 4 respectively. Because the ELD was not visible or is indistinct in most of our cases, precluding reliable measurements on MR imaging, we elected not to routinely measure ELD width, except in cases in which we thought the ELD could be reliably measured on our PACS, in which cases the axial midpoint diameter of the ELD was measured. When extraosseous ELSs were evident, the greatest orthogonal dimensions of the extraosseous portion of the ELS were measured in the axial plane.

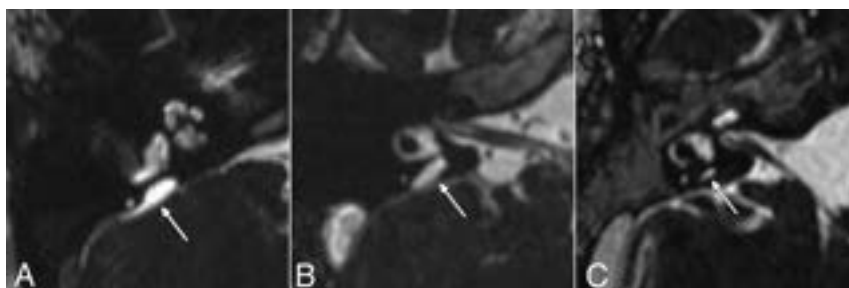
To determine whether MR imaging visibility was associated with VA enlargement, we implemented mixed-effects logistic



**FIG 2.** Examples of the 3 types of ELD visibility observed on MR imaging. *Upper row*, Representative axial-view non-contrast-enhanced temporal bone CT images of 3 different ears from different patients. *Lower row*, Axial temporal bone CISS MR images corresponding to the same ears as the CT images directly above demonstrate type 1 (*lower left image*), type 2 (*lower middle image*), and type 3 ELDs (*lower right image*). *Blue arrows* in the *upper row* indicate the vestibular aqueducts; *red arrows* in the *lower row* indicate the visible ELDs.



**FIG 3.** CISS MR images demonstrating additional examples of type 2 (faintly visible) ELDs (*arrows*) in 3 different patients: *A*, A 4-year-old girl with mild SNHL in the left ear. This patient demonstrated EVA on CT. *B*, A 1-year-old boy with profound SNHL in the right ear. This patient did not have VA enlargement on CT. *C*, An 8-year-old girl with profound SNHL in the left ear. This patient did not have VA enlargement on CT.



**FIG 4.** CISS MR images demonstrating additional examples of type 3 (easily visible) ELDs (*arrows*) in 3 different patients: *A*, A 7-month-old girl with severe SNHL in the right ear. *B*, A 6-year-old boy with severe SNHL in the right ear. *C*, A 4-year-old girl with mild SNHL in the left ear. All patients in these images demonstrated EVA on CT.

regressions with a logit link. In addition, to evaluate whether mean VA widths differ by MR imaging visibility, we implemented mixed-effects regressions, with separate models performed for

each VA measurement. For both sets of models, we included a random R-side compound-symmetric covariance term to account for correlations arising within patients since measurements were collected from both ears. Ear side and MR imaging scanner field strength were included as control variables, as well as the main independent variable, MR imaging visibility. A  $P$  value  $< .05$  for the type 3 test for the effect of MR imaging visibility was considered evidence of an association or mean difference. When an association or mean difference was found, we computed predicted probabilities of enlargement or predicted mean VA widths for each MR imaging visibility category and 95% confidence intervals using observed margins of covariates. We also computed pair-wise tests of predicted probability or mean differences with  $P$  values adjusted using the Tukey method to preserve type 1 error. The initial consensus assessments by the neuroradiology fellow and senior neuroradiologist were used as the criterion standard for these analyses.

Because the fellow and senior attending physician performing the initial imaging review were not blinded to the CT findings, we further assessed interrater agreement in the MR imaging evaluation of ELD visibility by having the MR imaging examinations read separately by 2 additional readers, including a junior attending neuroradiologist with nearly 2 years of experience and the neurotologist who had initially gathered the clinical data and had since joined the faculty. The image review by the neurotologist was performed at least 1 year after the clinical data were collected, and both readers were blinded to the clinical data and CT examinations at the time of MR imaging review. Interrater agreement among the 3 readers (reader 1 = fellow and senior neuroradiologist; reader 2 = junior neuroradiologist; reader 3 = neurotologist) was assessed using the Fleiss  $\kappa$  statistic, including separate statistics by laterality because observations are not independent. The  $\kappa$  statistics, along with a large-sample confidence interval derived using the  $\Delta$  method, were calculated for the 3 readers for the original 3-category outcome (type 1, type 2, or type 3 visibility)

regressions with a logit link. In addition, to evaluate whether mean VA widths differ by MR imaging visibility, we implemented mixed-effects regressions, with separate models performed for

and for a dichotomized outcome of no or faint MR imaging visibility (type 1 or type 2 MR imaging visibility) versus easy MR imaging visibility (type 3 MR imaging visibility). We also repeated the agreement analysis for pair-wise combinations of readers (reader 1 versus reader 2, reader 1 versus reader 3, and reader 2 versus reader 3).

## RESULTS

A total of 537 temporal bone HRCT studies were identified on our PACS and subsequently screened for potential inclusion. For 201/537 HRCT studies, a brain MR imaging was also available for review. Fourteen patients were noted to have undergone multiple HRCT examinations (ranging from 2 to 4), ultimately yielding 184 unique patients with both HRCT and MR imaging for additional screening. Eighty-four patients were eventually excluded due to the MR imaging not including high-resolution inner ear sequences ( $n = 64$ ); the HRCT or MR imaging being technically suboptimal due to significant motion, susceptibility, or other artifacts ( $n = 19$ ); or the presence of significant posttraumatic or

postsurgical deformities of the temporal bones ( $n = 1$ ). Two additional patients were excluded because their IAC MR imaging used a sequence other than CISS for inner ear characterization: T2-weighted, sampling perfection with application-optimized contrast by using different flip angle evolution (T2-SPACE sequence; Siemens) in 1 case and FIESTA in the other. A total of 98 patients and 196 ears were included in the final analysis (Fig 1). For patients with multiple HRCT or MR imaging scans available, the first of each scan type was used for imaging review unless it was deemed technically inadequate or if the initial MR imaging did not include high-resolution images through the inner ears; in the latter situation, the first MR imaging examination including high-resolution images was used.

The mean patient age at the time of HRCT was 6.2 (SD, 4.7) years. Fifty-three subjects (54.1%) were females. The median interval between HRCT and MR imaging was 62 days (interquartile range, 581.8 days). One child did not have audiometric data available for review. Of the remaining 97 children, 66 had bilateral hearing loss and 31 had unilateral hearing loss (Table 1). Of the 194 ears for which audiometry was available, 31 demonstrated normal hearing; 140, SNHL; 18, conductive hearing loss; and 5, mixed hearing loss (Table 2). Of note, the 2 ears in the child for whom audiometric data were unavailable both demonstrated VA enlargement on HRCT.

In 145/196 (74.0%) of the evaluated ears, the high-resolution temporal bone MR imaging sequence failed to demonstrate the ELD (type 1 visualization). In the remaining ears, the ELD was faintly visible (type 2 visualization) in 29/196 (14.8%) and was easily visible (type 3 visualization) in 22/196 (11.2%). Other inner ear and IAC imaging abnormalities were demonstrated in 23.4% of all ears, including dysmorphic cochlear and vestibular structures, the most common of which was an incomplete partition spectrum anomaly, observed in 8.1% of the evaluated ears.

On HRCT, midpoint VA widths measured in the axial plane ranged from 0 to 3.7 mm (median, 0.7 mm; mean, 0.84 [SD, 0.610] mm), while opercular widths ranged from 0 to 5.9 mm (median, 1.0; mean, 1.23 [SD, 0.92] mm). Midpoint VA widths measured in the Pöschl plane ranged from 0 to 3.8 mm (median, 0.6 mm; mean, 0.78 [SD, 0.54] mm).

A summary of rates of ELD visibility on MR imaging based on EVA status is presented in Table 3. Based on HRCT imaging, 40/196 ears (20.4%) in 25 children met the criteria for being EVA+ according to the Cincinnati criteria. There was a significant association between the presence of EVA on HRCT and ELD visibility on MR imaging ( $<.001$ ). In the EVA- ears, ELD visibility on MR imaging was type 1 in 87.2%, type 2 in 12.8%, and type 3 in 0%. In EVA+ ears, ELD visibility was type 1 in 22.5%, type 2 in 22.5%, and type 3 in 55.0% (Table 3).

**Table 1: Baseline demographic, audiometric, and imaging data**

Patient-Level Characteristics	Overall ( $n = 98$ Patients)
Demographic	
Age (mean) (SD) (yr)	6.2 (4.7)
Female sex (No.) (%)	53 (54.1)
Hearing loss laterality (No.) (%)	
Bilateral	66 (67.3)
Unilateral	31 (31.6)
Data not available	1 (1.0)
MR imaging field strength (No.) (%)	
1.5T (%)	38 (38.8)
3T (%)	60 (61.2)

**Table 2: Baseline demographic, audiometric, and imaging data**

Characteristics by Evaluated Ears	Overall ( $n = 196$ Ears)
Type of hearing loss (No.) (%)	
Normal	31 (15.8)
Sensorineural	140 (71.4)
Conductive	18 (9.2)
Mixed	5 (2.5)
Audiometry data not available	2 (1)
Severity of hearing loss (No.) (%)	
Normal	31 (15.8)
Mild	15 (7.7)
Moderate	23 (11.7)
Severe	42 (21.4)
Profound	83 (42.3)
Audiometry data not available	2 (1)
ELD visualization (No.) (%)	
Type 1	145 (74.0)
Type 2	29 (14.8)
Type 3	22 (11.2)

**Table 3: ELD visibility on high-resolution MR imaging based on EVA status**

VA Status on CT <sup>a</sup>	Type of ELD Visualization on MR Imaging			Totals (Column %)
	Type 1, Not Visualized No. (Row %)	Type 2, Faintly Visualized No. (Row %)	Type 3, Easily Visualized No. (Row %)	
EVA –	136 (87.2)	20 (12.8)	0 (0)	156 (79.6)
EVA +	9 (22.5)	9 (22.5)	22 (55.0)	40 (20.4)
Totals (row %)	145 (74.0)	29 (14.8)	22 (11.2)	196 (100)

<sup>a</sup> There was evidence of an association between VA status and ELD visibility ( $P < .0008$ ).



The midpoint ELD width could be measured on MR imaging in all (22/22) of the type 3 ears and in 21/29 (72.4%) of type 2 ears. The mean ELD width was 1.39 (SD, 0.78) mm for all measurable ears, 0.79 (SD, 0.16) mm for type 2 ears, and 1.96 (SD, 0.70) mm for type 3 ears. Among EVA+ ears with measurable ELDs on MR imaging (30 total), 6 (20%) had an ELD width of <1 mm, including 4 ears rated as type 2 and 2 ears rated as type 3. All 24 ears in which the ELD midpoint width was  $\geq 1$  mm (including 4 type 2 ears and 20 type 3 ears) were EVA+ on CT. An enlarged extraosseous ELS was visible in 21 ears, all of which were EVA+ on HRCT and type 3 on MR imaging with axial ELS dimensions ranging from  $2.1 \times 1.6$  mm to  $14.1 \times 5.5$  mm.

An association was observed between the presence of EVA on HRCT and the type of MR imaging visibility ( $P = .0008$ ). The predicted probabilities of an ear being EVA+ based on the type of ELD visibility on MR imaging was 0.053 (95% CI, 0.024–0.114) for type 1 ears, 0.288 (95% CI, 0.141–0.498) for type 2 ears, and 0.997 (95% CI, 0.194–0.999) for type 3 ears (Table 4). There was also evidence of a higher probability of EVA with a 1.5T field strength compared with a 3T field strength ( $P = .031$ ).

There was evidence that the predicted mean VA widths differed significantly on the basis of MR imaging visibility ( $P < .0001$ ), regardless of the location (midpoint or opercular) or the reconstruction plane (axial or Pöschl) used for VA measurement on HRCT (Table 5). Predicted mean axial midpoint VA width was 0.62 (95% CI, 0.55–0.69) mm for type 1 ears, 0.97 (95% CI, 0.83–1.11) mm for type 2 ears, and 2.08 (95% CI, 1.92–2.25) mm for type 3 ears. There was evidence that all the predicted mean midpoint VA widths differed significantly from each other between the different types of ELD visibility on MR imaging (Tukey-adjusted  $P$  value  $< .0001$  for all comparisons).

**Table 4: Predicted probability of VA enlargement based on the type of MR Imaging visibility**

Type of ELD Visualization on High-Resolution MR Imaging	Predicted Probability of EVA	95% Confidence Level
Type 1	0.053	(0.024–0.114)
Type 2	0.288	(0.141–0.498)
Type 3	0.997	(0.194–0.999)

**Table 5: Predicted mean VA width based on MR imaging visibility of the ELD<sup>a</sup>**

Type of ELD Visualization on High-Resolution MR Imaging	VA Midpoint Width (Axial Plane)		VA Opercular Width (Axial Plane)		VA Midpoint Width (Pöschl Plane)	
	Predicted Mean (mm)	95% CI	Predicted Mean (mm)	95% CI	Predicted Mean (mm)	95% CI
Type 1	0.62	(0.55–0.69)	0.92	(0.80–1.03)	0.59	(0.53–0.64)
Type 2	0.97	(0.83–1.11)	1.43	(1.20–1.67)	0.83	(0.72–0.94)
Type 3	2.08	(1.92–2.25)	3.00	(2.73–3.28)	1.98	(1.85–2.11)

<sup>a</sup> All predicted means differed significantly from each other. For comparison of predicted mean VA opercular width between type 1 and type 2 ears, Tukey-adjusted  $P = .0004$ . For comparison of predicted mean VA midpoint width in the Pöschl plane between type 1 and type 2 ears,  $P = .0005$ . For all other comparisons,  $P < .0001$ .

**Table 6: Results of 3-outcome ELD assessment on MR imaging by the 3 readers**

Type of ELD Visualization on High-Resolution MR Imaging	Reader 1 (Fellow and Senior Radiologist)		Reader 2 (Junior Radiologist)		Reader 3 (Neurotologist)	
	Right	Left	Right	Left	Right	Left
Type 1	69	76	49	47	67	73
Type 2	17	12	34	38	20	16
Type 1 or 2	86	88	83	85	87	89
Type 3	12	10	15	13	11	9

The predicted mean axial VA opercular width was 0.92 (95% CI, 0.80–1.03) mm for type 1 ears; 1.43 (95% CI, 1.20–1.67) mm for type 2 ears; and 3.00 (95% CI, 2.73–3.28) mm for type 3 ears. There was evidence that all predicted mean opercular widths differed significantly among MR imaging visibility types (Tukey-adjusted  $P$  value  $= .0004$  for the comparison between type 1 and type 2 ears, and  $< .0001$  for all other comparisons). The predicted mean midpoint VA measurements in the Pöschl plane also differed significantly from each other across all ELD visibility types (Table 5). There was evidence that the 1.5T field strength was associated with higher mean VA widths for all 3 measurements compared with 3T field strength ( $P$  values ranging from .0003 to .0051).

The results of the independent reader assessments of ELD visibility on MR imaging are summarized in Tables 6 and 7. For distinguishing the 3 types of ELD visualization separately, there was substantial agreement among the 3 readers for the right side (Fleiss  $\kappa = 0.691$ , 95% CI, 0.579–0.803) and moderate agreement for the left side (Fleiss kappa  $= 0.531$ ; 95% CI, 0.396–0.666).<sup>16</sup>

When type 1 and type 2 visualizations were combined into a single category resulting in a dichotomized outcome variable (type 1 or type 2 versus type 3), there was almost perfect inter-reader agreement among the 3 readers with Fleiss  $\kappa$  values of 0.849 (95% CI, 0.714–0.984) for the right side and 0.825 (95% CI, 0.667–0.982) for the left side.<sup>16</sup>

Pair-wise analysis of reader agreement showed almost perfect agreement between readers 1 and 3 for both 3-category and 2-category ELD assessment on both sides. Reader 2 demonstrated fair-to-substantial agreement with readers 1 and 3 (Fleiss  $\kappa$ ; range, 0.361–0.615) in the 3-category ELD assessment (Table 7). For the 2-category ELD assessment, reader 2 showed substantial-to-almost perfect agreement with readers 1 and 3 (Fleiss  $\kappa$  range, 0.754–0.822).<sup>16</sup> Overall, reader 2 rated more ears as type 2 or type 3 compared with readers 1 and 3 (Table 6).

## DISCUSSION

Our findings suggest that the high-resolution CISS sequence we commonly use at our institution for inner ear assessment performs poorly for visualizing the ELD and should not be relied on



**Table 7: Summary of interreader agreement analysis for ELD visibility on MR imaging**

Readers	3-Category Outcome: Fleiss $\kappa$ (95% CI)		2-Category Outcome: Fleiss $\kappa$ (95% CI)	
	Right	Left	Right	Left
All 3 readers	0.691 (0.579–0.803)	0.531 (0.396–0.666)	0.849 (0.714–0.984)	0.825 (0.667–0.982)
Pair-wise comparisons:				
Reader 1 vs 2	0.555 (0.402–0.709)	0.361 (0.186–0.537)	0.785 (0.604–0.966)	0.754 (0.547–0.960)
Reader 1 vs 3	0.935 (0.863–1.007)	0.844 (0.724–0.963)	0.951 (0.885–1.047)	0.942 (0.828–1.055)
Reader 2 vs 3	0.615 (0.468–0.762)	0.449 (0.279–0.619)	0.822 (0.654–0.991)	0.795 (0.601–0.989)

to rule out a diagnosis of EVA in children being evaluated for hearing loss. Notably, in nearly three-quarters of our sample, the ELD could not be seen on high-resolution inner ear MR imaging. The percentage of ears in which the ELD was not visible increased to >87% when we focused on EVA– ears, with the ELD being only faintly visible in the remainder. These results are in keeping with a previous study by Sarioglu et al,<sup>13</sup> who reported an overall rate of ELD nonvisualization in 79.2% of all ears and 88.5% of EVA– ears on MR imaging.

However, our results diverge from those of Sarioglu et al<sup>13</sup> when focusing on patients with EVA. Unlike in their report, which found an enlarged VA on MR imaging in 92.9% of EVA cases and a nonvisible VA in the remaining 7.1%, we noted that the ELD was not visible in 22.5% and was only faintly visible in another 22.5% on MR imaging in our population of EVA+ ears.

Several previous studies have reported a high concordance between HRCT and MR imaging for diagnosing EVA; however, in most of those studies, HRCT was more sensitive.<sup>4,7,8,10–12</sup> Recently, Clarke et al<sup>10</sup> found that MR imaging had a sensitivity of 97% for identifying EVA, misdiagnosing only 1 of 38 EVA+ ears. In comparison, we noted that MR imaging did a much poorer job at identifying EVA, failing to depict the ELD in 9/40 EVA+ ears in our sample. While the reason for this discrepancy is unclear, the explanation may rest, in part, on between-study differences in specific sequences used. Our study included only CISS sequences for inner ear characterization, while Clarke et al used 3D FSE sequences (T2 FSE and T2 SPACE). Some centers prefer 3D FSE sequences for inner ear evaluation because they eliminate banding artifacts, have fewer flow artifacts, and have fewer susceptibility artifacts at air-bone interfaces.<sup>17,18</sup> On the other hand, the previously mentioned study by Sarioglu et al<sup>13</sup> used a balanced fast-field echo MR imaging sequence, which, like CISS, is also a balanced steady free precession sequence. Whether 3D TSE sequences or other balanced steady-state gradient-echo sequences are superior to CISS for ELD evaluation is unknown, and future head-to-head comparisons between different high-resolution sequences might be helpful in elucidating this issue.

Unlike some previous studies,<sup>4,6,10,13</sup> we did not specifically compare the diagnostic sensitivities of HRCT and MR imaging for diagnosing EVA, in part because agreed-upon criteria for diagnosing EVA on MR imaging do not exist. We also did not attempt to correlate VA width measurement on HRCT and ELD measurement on MR imaging, because these analyses have already been performed. Despite claims from earlier studies, it has been our observation that accurate measurement of ELD width on MR imaging is frequently problematic because the ELD is usually either not visible or only faintly resolved, making it nearly impossible to reliably measure its width in most cases. In fact, a

recent study by Saeed et al<sup>19</sup> found only moderate agreement among independent readers in the measurement of ELD mid-point width on high-resolution MR imaging. Instead, we elected to focus on a relatively simple subjective classification for evaluating the ELD, which can be easily applied in a busy clinical setting, as was evidenced by high, nearly perfect interrater agreement for distinguishing type 3 ELDs from type 1 and 2 ELDs. The assessments of the fellow and senior radiologists demonstrated almost perfect agreement with the neurotologist in ELD assessment regardless of whether a 3-category (type 1 versus type 2 versus type 3) or 2-category (type 1 or 2 versus type 3) outcome was used, while there was less robust agreement between the junior radiologist (reader 2) and the other 2 readers. Reader 2 rated more ears as type 2 or type 3 compared with readers 1 and 3, possibly due to a lack of calibration training, because the secondary readers were provided only single examples of each ELD visualization type before their MR imaging assessments. Greater training with more exposure to the different ELD types may have resulted in a higher rate of agreement with the other readers.

Despite the limitations of MR imaging for ELD visualization and for ruling out an EVA diagnosis in our sample, type 3 ELD visualization, when present, was highly predictive of the presence of EVA, with easy ELD visualization being associated with a >99% probability of an ear being EVA+. On the basis of this observation, it might be reasonable to suggest that HRCT may not be necessary for confirming EVA when a pre-existing high-resolution MR imaging demonstrates a type 3 ELD.

The significance of type 2 ELD visualization is less clear. Although type 2 visualization was seen in more EVA+ ears than EVA– ears (22.5% versus 12.8%, respectively), the predicted probability of EVA when a type 2 ELD was seen was only 0.288, meaning that most ears demonstrating faint ELD visualization on MR imaging are EVA–. Thus, while type 2 ELD visualization is relatively uncommon in EVA– ears, it is probably normal in most cases when present.

Like previous studies, our study design has the limitations inherent in retrospective analyses. We excluded >80% of the initial screened population, largely due to lack of relevant MR imaging, significantly reducing the sample size and potentially introducing selection biases to the study. At our institution, completion of MR imaging is at the clinician's discretion, and patients in this study were more likely to have SNHL than pure conductive hearing loss because those imaged for conditions such as ear disease due to chronic eustachian tube dysfunction, congenital aural atresia, and cholesteatoma would not routinely undergo MR imaging. Nevertheless, we believe that any selection bias would likely have minimal effect on our observations and conclusions due to the nature of our study aims.

This study being performed at single center and our focusing exclusively on the CISS sequence for inner ear evaluation may limit the generalizability of our findings and explain some of the discrepancies between our results and those of other previous studies. Conversely, this homogeneity of our data conferred by its single-center location and inclusion of only studies that used a CISS MR imaging sequence for the IAC evaluation may also provide some benefits of improved statistical strength to our analysis, which was nearly devoid of confounding effects due to large numbers of differing MR imaging sequences and referring facilities. Finally, increasing application of 7T and higher-field-strength imaging along with continuing advancements in hardware and sequence development will likely provide improvement in resolution and scan quality and afford better MR imaging visualization of the ELD and discrimination between normal and abnormal ELDs in the future.

## CONCLUSIONS

Our findings suggest that the CISS MR imaging technique commonly used for inner ear evaluation performs poorly overall at resolving the ELD and underdiagnoses EVA in a substantial proportion of cases. However, when the ELD is easily seen on CISS MR imaging, there is a >99% probability that the VA is enlarged. On the basis of these observations, we believe that HRCT should remain the first-line imaging choice for ruling out suspected EVA; however, in cases in which an initial MR imaging demonstrates type 3 ELD visibility, HRCT imaging may not be necessary to confirm the diagnosis.

**Disclosure forms** provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

- Joshi VM, Navlekar SK, Kishore GR, et al. **CT and MR imaging of the inner ear and brain in children with congenital sensorineural hearing loss.** *Radiographics* 2012;32:683–98 CrossRef Medline
- Koesling S, Rasinski C, Amaya B. **Imaging and clinical findings in large endolymphatic duct and sac syndrome.** *Eur J Radiol* 2006;57:54–62 CrossRef Medline
- Madden C, Halsted M, Benton C, et al. **Enlarged vestibular aqueduct syndrome in the pediatric population.** *Otol Neurotol* 2003;24:625–32 CrossRef Medline
- Connor SE, Dudau C, Pai I, et al. **Is CT or MRI the optimal imaging investigation for the diagnosis of large vestibular aqueduct syndrome and large endolymphatic sac anomaly?** *Eur Arch Otorhinolaryngol* 2019;276:693–702 CrossRef Medline
- Naganawa S, Koshikawa T, Iwayama E, et al. **MR imaging of the enlarged endolymphatic duct and sac syndrome by use of a 3D fast asymmetric spin-echo sequence: volume and signal-intensity measurement of the endolymphatic duct and sac and area measurement of the cochlear modiolus.** *AJNR Am J Neuroradiol* 2000;21:1664–69 Medline
- Deep NL, Hoxworth JM, Barrs DM. **What is the best imaging modality for diagnosing a large vestibular aqueduct?** *Laryngoscope* 2016;126:302–03 CrossRef Medline
- 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
- Simons JP, Mandell DL, Arjmand EM. **Computed tomography and magnetic resonance imaging in pediatric unilateral and asymmetric sensorineural hearing loss.** *Arch Otolaryngol Head Neck Surg* 2006;132:186–92 CrossRef Medline
- Parry DA, Booth T, Roland PS. **Advantages of magnetic resonance imaging over computed tomography in preoperative evaluation of pediatric cochlear implant candidates.** *Otol Neurotol* 2005;26:976–82 CrossRef Medline
- Clarke RL, Isaacson B, Kutz JW, et al. **MRI evaluation of the normal and abnormal endolymphatic duct in the pediatric population: a comparison with high-resolution CT.** *AJNR Am J Neuroradiol* 2021;42:1865–69 CrossRef Medline
- 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
- 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
- Sarioglu FC, Cetin AC, Guleryuz H, et al; Division of Pediatric Radiology, Department of Radiology, Dokuz Eylul University School of Medicine, Izmir, Turkey. **The diagnostic efficacy of MRI in the evaluation of the enlarged vestibular aqueduct in children with hearing loss.** *Turk Arch Otorhinolaryngol* 2021;58:220–26 CrossRef
- Boston M, Halsted M, Meinen-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
- Juliano AF, Ting EY, Mingkwansook V, et al. **Vestibular aqueduct measurements in the 45 degrees oblique (Pöschl) plane.** *AJNR Am J Neuroradiol* 2016;37:1331–37 CrossRef Medline
- Landis JR, Koch GG. **The measurement of observer agreement for categorical data.** *Biometrics* 1977;33:159–74 Medline
- Byun JS, Kim HJ, Yim YJ, et al. **MR imaging of the internal auditory canal and inner ear at 3T: comparison between 3D driven equilibrium and 3D balanced fast field echo sequences.** *Korean J Radiol* 2008;9:212–18 CrossRef Medline
- Naganawa S, Koshikawa T, Fukatsu H, et al. **MR cisternography of the cerebellopontine angle: comparison of three-dimensional fast asymmetrical spin-echo and three-dimensional constructive interference in the steady-state sequences.** *AJNR Am J Neuroradiol* 2001;22:1179–85 Medline
- Saeed HS, Rajai A, Dixon R, et al. **Can MRI biomarkers for hearing loss in enlarged vestibular aqueduct be measured reproducibly?** *Br J Radiol* 2023;96:20220274 CrossRef Medline

# Sinonasal (Schneiderian) Tumors in the Temporal Bone: Case Series and Systematic Review

Francis Deng, Peter M. Sadow, Nabeeha Khan, Xin Wu, Ivan H. El-Sayed, David H. Jung, Christine M. Glastonbury, and Amy F. Juliano



## ABSTRACT

**BACKGROUND:** Neoplasms derived from the sinonasal epithelium are a rare finding in the temporal bone, and their origins are controversial.

**PURPOSE:** To review the characteristics of sinonasal epithelial (previously known as Schneiderian) tumors occurring in the temporal bone.

**DATA SOURCE:** This was a 2-center case series and systematic review of MEDLINE, EMBASE, and the Web of Science through May 2021.

**STUDY SELECTION:** Patients with clinicopathologic evidence of temporal bone involvement by neoplasms of sinonasal epithelial origin were selected, with or without a history of prior primary sinonasal epithelial tumors.

**DATA ANALYSIS:** Clinical, radiologic, and pathologic data were extracted.

**DATA SYNTHESIS:** The systematic review included 56 studies and our 8 unpublished cases, totaling 76 cases of papillomas or squamous cell carcinomas in the temporal bone. Of these, 51% occurred secondary to sinonasal tumors, and 49% occurred primarily. Secondary tumors were usually metachronous (77%), with a median delay of 1 year from sinonasal-to-temporal bone tumor diagnosis. Most cases were unilateral (90%); bilateral temporal bone involvement occurred only as secondary ("trilateral") tumors. Unilateral secondary tumors had ipsilateral (81%) or bilateral (19%) sinonasal counterparts. Secondary tumors were more likely to be malignant (OR, 6.7,  $P < .001$ ).

**LIMITATIONS:** The review was based on case reports and small case series, which are subject to reporting bias.

**CONCLUSIONS:** The observed tumor patterns support the hypothesis that the Eustachian tube facilitates the spread of sinonasal epithelium-derived neoplasms from the sinonasal cavity to the temporal bone. Transtubal spread of sinonasal epithelium-derived neoplasms should be considered among the rare causes of middle ear masses.

**ABBREVIATIONS:** HPV = human papilloma virus; IQR = interquartile range; WHO = World Health Organization

Sinonasal papillomas are neoplasms that arise from the sinonasal tract lining, the previously eponymous so-called Schneiderian epithelium consisting of ectodermally derived ciliated respiratory epithelium.<sup>1</sup> In some cases, these tumors may undergo malignant transformation or arise de novo from the

same epithelium. In the 2017 World Health Organization (WHO) Classification of Head and Neck Tumors and the subsequent 2022 edition, the word "Schneiderian" has been replaced by "sinonasal" for locally derived papillomas and carcinomas,<sup>2,3</sup> but inclusion of prior terminology is necessary as a reference for historical cases and sites of origin. There are 3 subtypes of sinonasal papillomas: inverted, oncocytic, and exophytic.<sup>2,3</sup> These papillomas may harbor or progress to carcinomas with variable frequency.

Sinonasal (Schneiderian epithelium-derived) tumors may, on occasion, occur in the middle ear. Given that sinonasal (Schneiderian) mucosa is not native to the ear and, in fact, no ciliated respiratory mucosa is located in middle ear or external auditory canal, with the middle ear lined by a simple cuboidal epithelium and the outer tympanic membrane and ear canal lined by squamous epithelium, their pathogenesis has been debated by various authors for decades. The only ciliated respiratory epithelium in the vicinity of the middle ear is along the

Received October 18, 2023; accepted after revision December 7.

From the Russell H. Morgan Department of Radiology and Radiological Science (F.D.), Johns Hopkins University, Baltimore, Maryland; Department of Pathology (P.M.S.), Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts; Departments of Otolaryngology-Head and Neck Surgery (P.M.S., D.H.J.), and Radiology (A.F.J.), Massachusetts Eye and Ear, Harvard Medical School, Boston, Massachusetts; University of South Carolina School of Medicine (N.K.), Greenville, South Carolina; Departments of Radiology and Biomedical Imaging (X.W., C.M.G.) and Otolaryngology-Head and Neck Surgery (I.H.E.), University of California, San Francisco, San Francisco, California.

Please address correspondence to Amy F. Juliano, MD, Massachusetts Eye and Ear, 243 Charles St, Boston, MA, 02114; e-mail: amy\_juliano@meei.harvard.edu; @amyfjuliano



Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8146>

Eustachian tube, which also contains mucinous cells. A considerable number of patients with middle ear sinonasal-type tumors also had a history of tumors in the sinonasal region, whether synchronously or metachronously, and it has been postulated that the tumor seeds or spreads through the Eustachian tube, as has been shown to occur, rarely, in nasopharyngeal carcinomas.<sup>4-6</sup>

We herein present the largest primary case series, to our knowledge, of patients with middle ear sinonasal-type tumors, all of whom had a history of primary sinonasal tumors, along with a review of the current literature of these tumor types, with particular emphasis on the characterization of middle ear sinonasal-type tumors.

## MATERIALS AND METHODS

### Current Case Series

The radiology report database and teaching files at the authors' 2 affiliated institutions (Massachusetts Eye and Ear and the University of California, San Francisco) were searched for cases of temporal bone imaging (CT or MR imaging) with neoplastic involvement of the middle ear, the medical records of which were further reviewed for a history of sinonasal neoplasms and other clinicopathologic information. All cases found with middle ear involvement are reported. There was 1 patient in the institutional records with sinonasal-type tumors of both the middle ear and sinonasal region that were previously reported and not included here.<sup>7</sup> The study was approved by each of the 2 local institutional review boards.

### Search Strategy

A systematic review was performed and reported in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines (Online Supplemental Data). A literature search was conducted on PubMed (US National Library of Medicine), EMBASE (Elsevier), and Web of Science (Clarivate Analytics) from inception through May 27, 2021. The search string consisted of a combination of the following groups of terms: papilloma or carcinoma; Schneiderian, inverted, oncocytic, or exophytic; and temporal bone, middle ear, ear, or Eustachian tube. No additional search filters were used. Additional articles were identified by backward and forward pearl growing from relevant articles, in which the reference lists and citing articles on Google Scholar (<https://scholar.google.com/>) were manually reviewed. The protocol was not preregistered.

### Selection Process

Eligible articles were observational studies of patients with a tumor involving the middle ear that had pathology featuring Schneiderian (sinonasal) epithelium or pathology consistent with that of a concurrent or prior epithelial tumor in the sinonasal cavity in the same patient. After we removed duplicates, records were screened for the availability of full-text articles, excluding conference abstracts or posters. Two authors (N.K. and A.F.J.) independently evaluated all full-text articles for eligibility. Exclusion criteria included non-English articles, reviews and guidelines, animal studies, meeting/conference proceedings, and studies not about Schneiderian or sinonasal neoplasms in the ear.

### Data Extraction

Three authors (F.D., N.K., and A.F.J.) extracted data from the eligible full-text articles. The extracted data included study design, patient demographics (age and sex), symptoms, imaging modalities, all pathologic findings/diagnoses, and laterality of involvement. Pathologic diagnoses were categorized as benign or malignant. "Severe" or "high-grade" dysplasia, a high-risk lesion, was included in the malignant group along with carcinoma in situ and invasive carcinomas. The presence, laterality, and pathologic diagnoses of synchronous or preceding sinonasal neoplasm were also recorded. In cases in which no such data were described, it was assumed to be absent. No data were obtained from study authors directly.

### Quality Assessment

The level of evidence was assessed according to the Oxford Center for Evidence-Based Medicine "Levels of Evidence 1" pertaining to etiology studies.<sup>8</sup> Specifically, cohort study designs (level 2) were a higher level than case-control studies (level 3), which, in turn, were a higher level than case series (level 4). Potential sources of bias within and across studies were qualitatively described.

### Statistical Analysis

Descriptive statistics were calculated for continuous variables, while frequencies and percentages were calculated for categorical variables. Associations between categorical variables were analyzed using  $\chi^2$  test. ORs with 95% CI were computed to evaluate the association between clinicoradiologic factors and tumor malignancy. A 2-tailed *P* value < .05 was considered statistically significant. Statistical analyses were performed using GraphPad Prism (GraphPad Software).

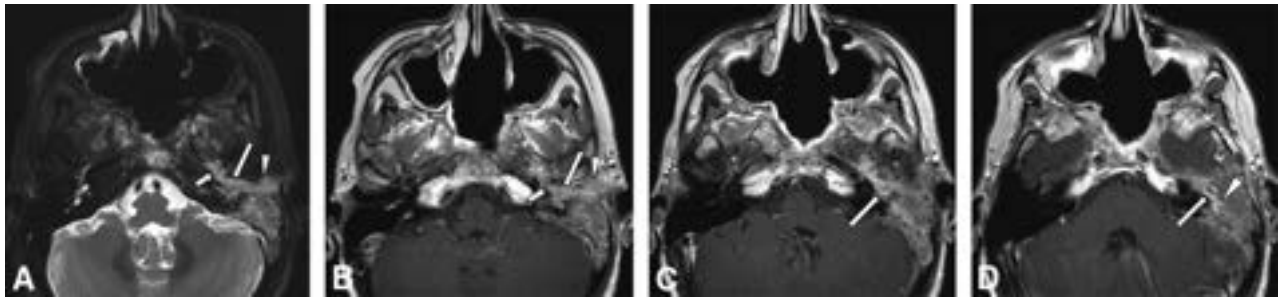
## RESULTS

### Case Series

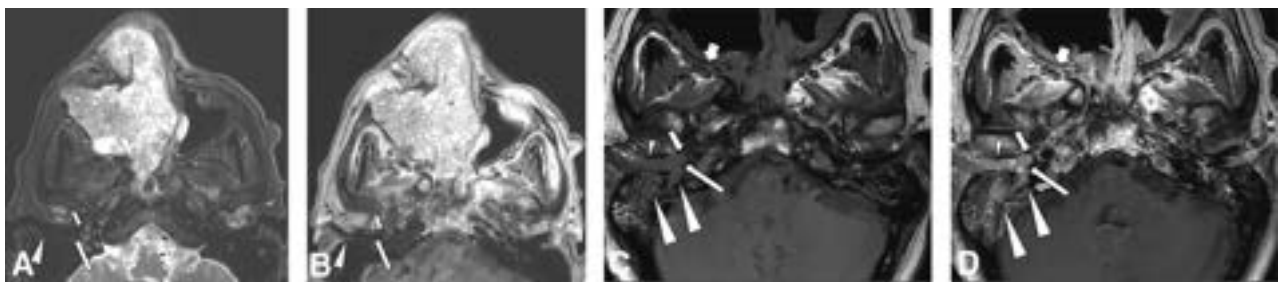
**Patient 1.** A 53-year-old woman presented with left aural fullness. She had a 30-year history of bilateral sinonasal inverted papillomas and left nasopharyngeal papillary squamous cell carcinoma 4 years prior, which was treated with definitive radiation. Imaging revealed abnormal tissue in the left middle ear, sinonasal cavity, and nasopharynx with clival invasion. Biopsy pathology from both sites showed nonkeratinizing invasive squamous cell carcinoma, possibly arising from an inverted papilloma. This case, from more than a decade prior, was a second opinion review with no tissue for additional immunohistochemistry, viral chromogenic in situ hybridization, or molecular testing. Following surgical debulking, postoperative MR imaging demonstrated the residual tumor, including in the left middle ear and mastoid, extending along the Eustachian tube (Fig 1). Despite proton beam radiation and chemotherapy, follow-up imaging showed metastatic lung and mediastinal nodal progression. Local temporal bone progression led to intracranial extension and cerebral herniation.

**Patient 2.** A 65-year-old man presented with right hearing loss, otorrhea, dizziness, and vertigo. He had an 8-year history of multiple, recurrent right sinonasal papillomas, inverted type, with





**FIG 1.** A 53-year-old patient with a history of bilateral sinonasal inverted papillomas and left nasopharyngeal papillary squamous cell carcinoma. Axial T2-weighted MR image (A) shows material of intermediate-to-hyperintense signal in the left posterior Eustachian tube (*short arrow*), middle ear (*long arrow*), and external auditory canal (*arrowhead*). Consecutive axial postcontrast T1-weighted MR images from caudal to cranial (B, C, D) show corresponding enhancement of this abnormal soft tissue in the posterior Eustachian tube (*short arrow* in B), middle ear (*long arrow* in B, C, and D), and external auditory canal (*arrowhead* in B), representing tumor. There is an extension to the mastoid antrum (*arrowhead* in D). Ear biopsy showed squamous cell carcinoma possibly arising from an inverted papilloma.



**FIG 2.** A 65-year-old patient with a history of recurrent, right sinonasal, inverted papilloma and carcinoma. Axial T2-weighted (A) and postcontrast T1-weighted (B) MR images show a large, bulky, lobulated septate enhancing mass in the right sinonasal region. Note the absence of abnormal soft tissue at the right Eustachian tube opening (*short arrow*), middle ear (*long arrow*), and external auditory canal (*arrowhead*). The patient underwent partial resection, chemoradiation, and immunotherapy. Two years later, he presented with otologic symptoms, and MR imaging was performed. Axial precontrast (C) and postcontrast (D) T1-weighted MR images show bulky enhancing soft tissue at the right Eustachian tube opening (*short arrow*), middle ear (*long arrow*), and external auditory canal (*small arrowheads*), with invasion into the mastoid and petrous portions of the temporal bone (*large arrowheads*). Also note tumor involvement in the skull base foramina such as the right pterygopalatine fossa (*short broad arrows*).

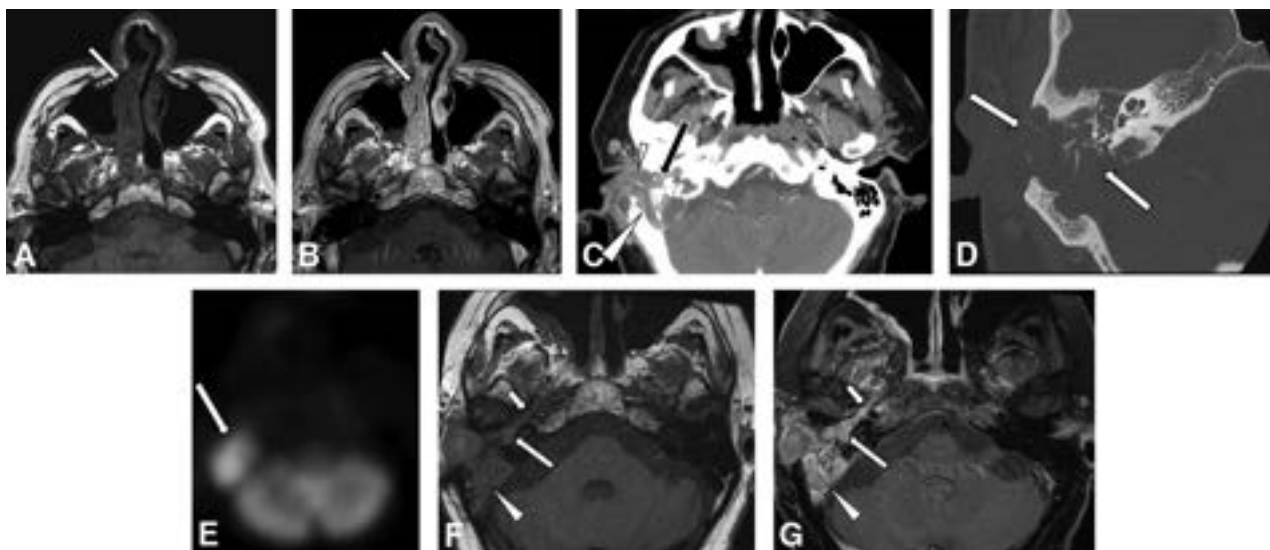
malignant transformation to carcinoma 2 years prior (Fig 2A, -B), which was treated with partial resection, chemoradiation, and immunotherapy but locally progressed. Human papillomavirus (HPV) testing revealed no expression of high-risk HPV types. At the time of otologic presentation, imaging showed bulky enhancing tissue in the right middle ear, mastoid, and external auditory canal, with bone erosion involving the mastoid and petrous apex (Fig 2C, -D). Consistent with tumor spread to the ear, follow-up during 2 years showed tumor progression both in the sinonasal cavity and right temporal bone before the patient was transitioned to palliative care.

**Patient 3.** A 60-year-old man presented with right hearing loss and otalgia for several months, failing treatments for presumed otitis externa. He had a history of right nasal cavity–based sinonasal inverted papilloma 2 years prior (Fig 3A, -B). An otologic examination found a fleshy mass in the right external auditory canal. CT and PET revealed a hypermetabolic mass centered in the right middle ear and mastoid with extensive bone erosion (Fig 3C, D, and E). Following partial resection, pathology confirmed sinonasal papilloma, inverted type, in both nasal and right ear sites. When he returned to follow-up 1.5 years later, imaging showed local tumor progression in the right temporal bone,

including tumor tracking along the Eustachian tube (Fig 3F, -G). A CT 4 months later showed right level II adenopathy, and biopsy revealed a low-grade sinonasal (Schneiderian) squamous cell carcinoma. The tumor was negative for high-risk HPV, but the patient did have a mutation in *p53* with additional loss of *CDKN2A*. Despite starting chemoradiation, thoracic metastases developed, leading to a focus on comfort care.

**Patient 4.** A 68-year-old man presented with left hearing loss for 2 months with 6 months of left nasal congestion and drainage. CT and MR imaging revealed a mass in the left nasal cavity extending to the left nasopharynx, abutting the Eustachian tube orifice; enhancing tissue was also present within the left middle ear. Surgical pathology review demonstrated a low-grade sinonasal carcinoma with mixed features of sinonasal (Schneiderian), salivary, and odontogenic origin. The tumor was negative for high-risk HPV and Epstein-Barr-encoded messenger RNA (by chromogenic in situ hybridization). The tumor was revealed to have an *NF1* variant.

Five months later, he presented with worsening left ear hearing and otalgia. MR imaging showed persistent enhancing tissue in the left posterior Eustachian tube and middle ear. A nonpulsatile erythematous mass was present on otoscopy. A middle ear



**FIG 3.** A 60-year-old patient with a right nasal cavity inverted papilloma. Axial precontrast (A) and postcontrast (B) T1-weighted images show the initial right sinonasal papilloma (arrow) 2 years prior. On presentation to otology, CT in soft-tissue (C) and bone (D) windows shows an expansile bulky mass in the right middle ear (long arrow in C), external auditory canal (arrowhead in C), and mastoid (large arrowhead in C) with bone destruction (arrows in D). PET (E) shows avid FDG uptake by the mass (arrow). Axial precontrast (F) and postcontrast (G) T1-weighted images show enhancing tumor along the right Eustachian tube (small arrow), in the middle ear (long arrow), and in the mastoid (large arrowhead). Pathology of the ear mass proved to be inverted papilloma.

mass biopsy demonstrated low-grade sinonasal carcinoma with histologic features similar to those in the previous sinonasal tumor. The patient underwent resection and proton beam therapy and showed no recurrence with 2 years of surveillance.

**Patient 5.** A 70-year-old woman presented with imbalance and bilateral hearing loss. Examination revealed a polypoid mass in the left external auditory canal. Biopsy indicated atypical squamous proliferation. CT and then MR imaging showed a soft-tissue mass at both nasal cavities, nasopharynx, and bilateral middle ear and mastoid regions, including near the Eustachian tubes. Additional biopsies of the masses in the right nasal cavity and right ear canal confirmed benign sinonasal papillomas, inverted type. No additional information regarding molecular or viral status is known about this case.

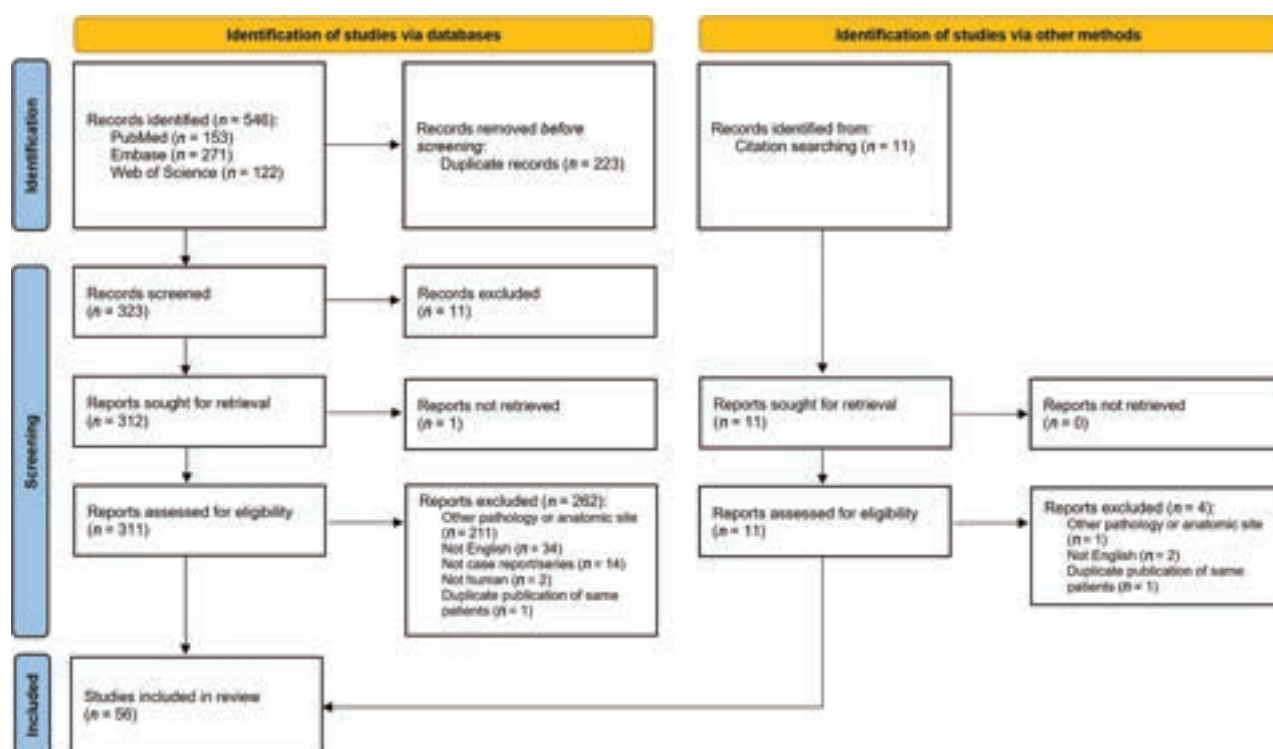
**Patient 6.** A 45-year-old man presented with hearing loss. He had a history of sinonasal inverted papilloma with high-grade dysplasia, resected 3 years prior, with recurrence as sinonasal squamous cell carcinoma, which had been resected 1 year before the current presentation. Review of previously performed MR imaging showed a large sinonasal mass that extended to the nasopharynx, abutting the left Eustachian tube, and abnormal enhancement along the left Eustachian tube and in the middle ear cavity. At the latest presentation, tumor was visible in both external auditory canals; bilateral biopsies demonstrated sinonasal squamous cell carcinoma evolving from sinonasal inverted papilloma. Imaging at this time confirmed destructive enhancing soft-tissue masses in the bilateral mastoid regions invading intracranially. No additional information regarding molecular or viral status is known about this case. He subsequently developed pulmonary metastases and entered hospice care.

**Patient 7.** A 70-year-old man presented with otalgia, otorrhea, trismus, facial swelling, and tooth loss. He reported a history of a “nodule” being surgically removed from his nasal cavity overseas 3 months prior, with unknown pathology. MR imaging demonstrated a destructive mass in the right maxillary sinus and nasal cavity, enhancing soft tissue in bilateral Eustachian tubes and temporal bones, and dural enhancement with frank intracranial extension. Nasal mass biopsy established a diagnosis of sinonasal squamous cell carcinoma with associated benign sinonasal inverted papilloma. Despite chemotherapy, the patient had disease progression and eventually died from the disease complicated by tumoral hemorrhage and respiratory failure.

**Patient 8.** A 36-year-old man with a previous diagnosis of Bell palsy presented with progressive bilateral facial paresis, hearing loss, otalgia, otorrhea, and rhinorrhea. One month prior, tumor was identified in the bilateral ear canals, and a biopsy indicated sinonasal inverted papilloma. At our institution, examinations revealed destructive, multicystic masses in the nasal cavity, nasopharynx, and bilateral mastoid temporal bones. There was intracranial invasion, left sigmoid sinus thrombosis, and temporal lobe edema. Debridement revealed squamous cell carcinoma in the nasopharynx and sinonasal inverted papilloma with severe squamous dysplasia in the skull base. No additional information regarding molecular or viral status is known about this case. Three months later, the patient had left temporal intracerebral hemorrhage, thought to be related to either direct tumor invasion or venous thrombosis.

### Systematic Review

A PRISMA flow diagram is shown in Fig 4, detailing the systematic identification, screening, and determination of inclusion or



**FIG 4.** PRISMA flow diagram showing search algorithm used for systematic review.

exclusion of publications. A total of 56 articles containing 68 cases were included (Online Supplemental Data).

**Bias.** All included published studies were case reports or case series up to a maximum of 4 patients. According to the Oxford Center for Evidence-Based Medicine Levels of Evidence of etiology, case series are level 4 out of 5. Case series are at risk of selection/publication bias, so unusual features or associations may be overrepresented compared with typical practice.

**Patient and Tumor Site Characteristics.** The systematic review and our case series of 8 patients sums the current analysis to a total of 76 cases. There were 45 males and 31 females (ratio 1.5:1). The median age at presentation of the middle ear tumor was 53 years (range, 11–82 years; interquartile range [IQR], 44–63 years).

Of these, 37 (49%) patients had tumor involving the middle ear primarily without sinonasal involvement, while 39 (51%) patients had tumor involving the middle ear, secondary to sinonasal involvement. In no case did Schneiderian-type tumor in the middle ear precede sinonasal occurrence of tumor. Of the secondary tumor cases, the median reported time interval between the sinonasal and ear tumor diagnosis was 1 year (range, 0–40 years; IQR, 0.5–4 years); cases were considered synchronous (within 6 months) with the sinonasal tumor in 9/39 (23%) and metachronous (at least 6 months) in 30/39 (77%).

Of the cases that reported the laterality of tympanomastoid involvement, most cases were unilateral (65/72; 90%). Of the secondary tumor cases that reported a laterality of sinonasal involvement, most cases were unilateral (31/38; 82%). All 7 cases of sinonasal-type (Schneiderian) tumors presenting in the bilateral

tympanomastoid cavities were secondary to a sinonasal tumor, ie, “trilateral.” Unilateral secondary sinonasal-type tumors of the ear had ipsilateral (25/31; 81%) or bilateral (6/31; 19%) sinonasal counterparts; there were no cases of unilateral tympanomastoid and contralateral unilateral sinonasal involvement.

**Clinical Presentation.** The 76 patients with sinonasal-type tumors in the middle ear most commonly presented with the following: hearing loss (67%), otorrhea (47%), aural fullness (18%), otalgia (17%), and tinnitus or pulsatile tinnitus (16%), among others (Table 1). When imaging was reported, most patients underwent both CT and MR imaging (34/61; 56%; including all 8 cases in our series) or CT alone (25/61; 41%).

**Histopathologic Characteristics.** The histopathology of the middle ear tumors on initial diagnosis was benign in 65% (48/74) and malignant in 35% (26/74); the ear component of 2 cases from our series was diagnosed as malignant on clinicoradiologic grounds. Primary (isolated) tumors were predominantly benign (31/37; 84%), while secondary sinonasal-type tumors of the ear were often malignant (22/39; 54%) (Table 2). The association of secondary tumors with malignancy was statistically significant (OR, 6.7; 95% CI, 2.3–19.1;  $\chi^2$ ,  $P < .001$ ).

Of the secondary tumor cases, the sinonasal tumor histopathology closest in time to the ear tumor diagnosis was benign in 62% (24/39) and malignant in 38% (15/39). Tumors in the middle ear and the sinonasal region were always of the same histologic lineage (sinonasal papilloma/carcinoma).

Testing for HPV and other viral or molecular markers was often not performed or reported. In the 15 cases in which findings

**Table 1: Summary of patient characteristics in systematic review and case series of temporal bone Schneiderian neoplasms<sup>a</sup>**

Characteristic	
Demographics (n = 76)	
Age at presentation (median) (IQR)	53 (44–63)
Sex ratio (male/female)	1.5 (45:31)
Main presenting otologic signs/symptoms (n = 76) <sup>b</sup>	
Hearing loss	51 (67%)
Otorrhea or otorrhagia	36 (47%)
Aural fullness	14 (18%)
Otalgia	13 (17%)
Tinnitus or pulsatile tinnitus	12 (16%)
Acute or chronic otitis media and/or mastoiditis	9 (12%)
Visible ear mass	9 (12%)
Facial paralysis	7 (9%)
Dizziness or vertigo	2 (3%)
Association of temporal bone tumor with sinonasal tumor (n = 76)	
Primary (isolated)	37 (49%)
Secondary	39 (51%)
Temporal association of secondary tumor diagnosis with sinonasal tumor (n = 39)	
Synchronous	9 (23%)
Metachronous	30 (77%)
Temporal bone tumor pathology (n = 74)	
Benign	48 (65%)
Malignant	26 (35%)
Sinonasal tumor pathology (n = 39)	
Benign	24 (62%)
Malignant/high-risk	15 (38%)

<sup>a</sup> Missing data are excluded from the denominators of each section.

<sup>b</sup> Cells add up to >100% as cases may feature multiple presenting signs/symptoms.

**Table 2: Contingency table on the association of temporal bone Schneiderian tumors that are primary (isolated) or secondary (synchronous or metachronous to a similar sinonasal tumor) and pathology that is malignant/high-risk (carcinoma, carcinoma in situ, or papilloma with severe/high-grade dysplasia) or benign (papilloma without severe/high-grade dysplasia)**

	Primary	Secondary
Malignant	6	22
Benign	31	17

of HPV were negative, most tumors were benign (11/15; 73%). In the 12 cases in which findings of HPV were positive, most were malignant (7/12; 58%).

We examined the role of the Eustachian tube in disease involvement and spread. Among the 39 patients with secondary tumors, 9 (23%) were reported to be positive for Eustachian tube tumor involvement by biopsy/surgical resection histology or by intraoperative visual examination, 16 (41%) had imaging studies showing enhancing soft tissue consistent with tumor along the Eustachian tube, and 3 (8%) were presumed to be involved due to clinical presentation or proximity of lesions. In 6 case reports (15%), the Eustachian tube was thought not to be involved either by CT, visual inspection during surgery, biopsies taken along the Eustachian tube, or not otherwise specified. The Eustachian tube was not mentioned in 5 cases.

## DISCUSSION

We present a comprehensive systematic review that incorporates all reported cases in the English literature on sinonasal-type

(Schneiderian) tumors involving the middle ear, with or without antecedent or concurrent tumors of the same histopathologic lineage in the sinonasal region. We also present a robust cohort of previously unreported cases in patients with concomitant middle ear and sinonasal tumors.

Because the sinonasal (Schneiderian) epithelium is not native to the temporal bone, hypothesized mechanisms for tumors of sinonasal (Schneiderian) origin developing in the middle ear include the following: 1) direct spread via the Eustachian tube, akin to the manner in which nasopharyngeal carcinoma has been shown to spread to the middle ear;<sup>4–6</sup> 2) metaplasia of the middle ear mucosa to Schneiderian mucosa, perhaps as a result of chronic inflammation (though this may be of doubtful validity because inflammation-induced metaplasia does not convert the simple cuboidal epithelium of the middle ear canal into ciliated respiratory [Schneiderian] mucosa); and 3) anomalous embryonic migration of Schneiderian-type ciliated respiratory mucosa. The Eustachian tube is a physical connection between the 2 sites, and if nasopharyngeal carcinoma can extend via the Eustachian tube to the middle ear,<sup>4–6</sup> then sinonasal tumor spreading along the sinonasal and pharyngeal mucosa could well continue along the Eustachian tube in the same manner.

Direct tumor spread from the sinonasal region to the middle ear is a well-supported mechanism.<sup>9</sup> More than one-half of the cases reviewed here had a synchronous or metachronous prior diagnosis of sinonasal tumor of similar histologic lineage (sinonasal mucosa origin). All those tumors presented concurrently with or subsequent to the sinonasal tumor diagnosis, never preceding it, and were more likely malignant. The involved ear was never contralateral to the side of the sinonasal tumor, only ipsilateral or bilateral. All 8 patients in our series and a number of previously published cases showed involvement of the Eustachian tube on imaging (characterized by soft tissue isodense/isointense to tumor along the Eustachian tube course). These observations support the postulation that sinonasal tumors may spread through the Eustachian tube.

Squamous tumors of different sites may have variable histologic morphologies. Some squamous epithelium is keratinizing or nonkeratinizing, some lesions have viral cytopathic changes, and some locations may boast squamous epithelium of mixed morphology. When tumors spread to unusual sites and have multiple possible primary sites, identifying the tumor origin may require more in-depth tissue evaluation, particularly involving site/tissue-specific genetic signatures, such as *epidermal growth factor receptor* or HPV-associated anomalies in sinonasal (Schneiderian) tumors.<sup>10</sup>

We propose that in patients with sinonasal-type tumors, clinical and imaging surveillance should include otologic examination. Temporal bone involvement may currently be underdetected. Radiologists should be aware of the need to scrutinize the Eustachian tube and middle ear, whether on CT, MR imaging, or potentially PET. While middle ear effusions are common after treatment of sinonasal tumors,<sup>11</sup> the alternative possibility of tumor spread to the middle ear should be considered when there is solid, enhancing tissue in the middle ear similar to the original sinonasal tumor. Given the aggressive nature of malignant sinonasal tumors and the high rate of malignancy when



found in the middle ear, it behooves clinicians to scrutinize the Eustachian tube region and have a low threshold for otology consultation and consideration of biopsy. Furthermore, the concept of the primary origin of these tumor types in the middle ear, without a thorough examination of prior history and imaging, should warrant revisitation, given that all the cases in this series showed concurrent or prior sinonasal tumor. Most important, prompt detection of tumor spread may influence management such as radiation therapy target delineation.

This review relies on case series and case reports. The reporting of features may be selective and inconsistent across time because additional testing modalities and genetic signatures have evolved with their increased availabilities. This inconsistency limits the generalizability of observed associations. Also, given that these are historical cases that have accumulated during longer than the past decade, some testing modalities were not available for some cases, particularly some of the molecular testing for consult cases, and molecular data are incomplete.

## CONCLUSIONS





The observed tumor patterns in this systematic review and our case series support the hypothesis that the Eustachian tube may facilitate the spread of sinonasal (formerly Schneiderian) neoplasms from the sinonasal cavity to the temporal bone. Transtubal spread of sinonasal-origin neoplasms should be considered among the rare causes of middle ear masses. For the subset of this tumor type reported by others to appear primarily in the middle ear, a thorough review of the patient's history, particularly with respect to any prior sinonasal surgery, for any reason, is essential. Clinicians and radiologists should be diligent when considering middle ear disease in patients with sinonasal neoplasms to avoid neglecting tumor extension to the middle ear from a sinonasal source, ensuring prompt and appropriate treatment considerations of all involved sites. Recent progress in molecular classification of these tumors may also help in discerning tumor type/origin.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

1. Bishop JA. **OSPs and ESPs and ISPs, Oh my! An update on sinonasal (Schneiderian) papillomas.** *Head Neck Pathol* 2017;11:269–77 CrossRef Medline
2. El-Naggar A, Chan J, Grandis J, et al, eds. **WHO Classification of Head and Neck Tumours.** Vol 9. 4th ed. World Health Organization; 2017
3. International Agency for Research on Cancer. **WHO Classification of Tumours online.** 2022 <https://tumourclassification.iarc.who.int>. Accessed November 13, 2023
4. Chin SC, Fatterpekar G, Chen CY, et al. **MR imaging of diverse manifestations of nasopharyngeal carcinomas.** *AJR Am J Roentgenol* 2003;180:1715–22 CrossRef Medline
5. Yang MS, Chen CC, Cheng YY, et al. **Nasopharyngeal carcinoma spreading along the eustachian tube: the imaging appearance.** *J Chin Med Assoc* 2004;67:200–03 Medline
6. Lai JW, Cheng AC, Tong M, et al. **Sneaking its way up: external auditory canal involvement by an otherwise inconspicuous nasopharyngeal carcinoma via spread through the eustachian tube.** *Hong Kong J Radiol* 2015;18:e1–e6 CrossRef
7. Lin HW, Richmon JD, Emerick KS, et al. **Malignant transformation of a highly aggressive human papillomavirus type 11-associated recurrent respiratory papillomatosis.** *Am J Otolaryngol* 2010;31:291–96 CrossRef Medline
8. Centre for Evidence-Based Medicine. **Oxford Centre for Evidence-Based Medicine: Levels of Evidence (March 2009).** <https://www.cebm.ox.ac.uk/resources/levels-of-evidence/oxford-centre-for-evidence-based-medicine-levels-of-evidence-march-2009>. Accessed May 5, 2023
9. Kawaguchi M, Kato H, Tomita H, et al. **Imaging characteristics of malignant sinonasal tumors.** *J Clin Med* 2017;6:116 CrossRef Medline
10. Brown NA, Plouffe KR, Yilmaz O, et al. **TP53 mutations and CDKN2A mutations/deletions are highly recurrent molecular alterations in the malignant progression of sinonasal papillomas.** *Mod Pathol* 2021;34:1133–42 CrossRef Medline
11. Redaelli de Zinis LO, Parrinello G, Schreiber A, et al. **Middle ear effusion in patients with sinonasal cancer treated by surgery with or without radiotherapy.** *Otolaryngol Head Neck Surg* 2013;148:619–24 CrossRef Medline

# Utility of MR Neurography for the Evaluation of Peripheral Trigeminal Neuropathies in the Postoperative Period

 Tyler Thornton,  Shuda Xia,  John R. Zuniga, and  Avneesh Chhabra

## ABSTRACT

**SUMMARY:** Peripheral trigeminal neuropathies are assessed by MR neurography for presurgical mapping. In this clinical report, we aimed to understand the utility of MR neurography following nerve-repair procedures. We hypothesized that postoperative MR neurography assists in determining nerve integrity, and worsening MR neurography findings will corroborate poor patient outcomes. Ten patients with peripheral trigeminal neuropathy were retrospectively identified after nerve-repair procedures, with postsurgical MR neurography performed from July 2015 to September 2023. Postsurgical MR neurography findings were graded as per postintervention category and subcategories of the Neuropathy Score Reporting and Data System (NS-RADS). Descriptive statistics of demographics, inciting injury, injury severity, NS-RADS scoring, and clinical outcomes were obtained. There were 6 women and 4 men (age range, 25–73 years). Most injuries resulted from third molar removals (8/10), with an average time from the inciting event to nerve-repair surgery of 6.1 (SD, 4.6) months. In Neuropathy Score Reporting and Data System-Injury (NS-RADS I), NS-RADS I-4 injuries (neuroma in continuity) were found in 8/10 patients, and NS-RADS I-5 injuries were found in the remaining patients, all confirmed at surgery. Surgeries performed included microdissection with neurolysis, neuroma excision, and nerve allograft with Axoguard protection. Three patients with expected postsurgical MR neurography findings experienced either partial improvement or complete symptom resolution, while among 7 patient with persistent or recurrent neuropathy on postsurgical MR neurography, one demonstrated partial improvement of sensation, pain, and taste and one experienced only pain improvement; the remaining 5 patients demonstrated no improvement. Postsurgical MR neurography consistently coincided with clinical outcomes related to pain, sensation, and lip biting and speech challenges. Lip biting and speech challenges were most amenable to recovery, even with evidence of persistent nerve pathology on postsurgical MR neurography.

**ABBREVIATIONS:** MRN = MR neurography; NS-RADS I = Neuropathy Score Reporting and Data System Injury; NS-RADS PI = Neuropathy Score Reporting and Data System Postintervention; NST = neurosensory testing; PI = postintervention; PTN = peripheral trigeminal neuropathy

The trigeminal nerve is the largest of the cranial nerves that provide sensory innervation to the face in addition to motor innervation to the muscles of mastication.<sup>1</sup> It arises from the anterolateral aspect of the midpons with a dominant sensory and a smaller motor branch before dividing at the trigeminal ganglion into the 3 branches: ophthalmic, maxillary, and mandibular. These subdivisions further branch into smaller nerves that provide innervation to their respective maxillofacial distributions. For example, the inferior alveolar nerve innervates the lower jaw, and the lingual nerve innervates the tongue.<sup>2</sup>

Received September 11, 2023; accepted after revision December 20.

From the University of North Texas Health Science Center (T.T.), Fort Worth, Texas; University of Texas Southwestern Medical Center (S.X.), Dallas, Texas; and Departments of Oral and Maxillofacial Surgery (J.R.Z.), and Radiology (A.C.), University of Texas Southwestern Medical Center, Dallas, Texas.

Please address correspondence to Avneesh Chhabra, MD, MBA, FACR, 5959 Harry Hines Blvd., 11th Floor, Suite 300, Dallas, Texas 75390; e-mail: avneesh.chhabra@utsouthwestern.edu

<https://dx.doi.org/10.3174/ajnr.A8152>

Peripheral trigeminal neuropathies (PTNs) result in loss of sensation and/or the development of neuropathic pain, and these can be caused by both traumatic (third molar removal, oral implants) and nontraumatic (neoplastic, atypical facial pain) etiologies. The most common cause is third molar extractions with up to 10 million being performed each year, accounting for 60% of nerve injuries around the jaw. Among the trigeminal nerve branches, the lingual nerve, including the special sensory branch of the chorda tympani (cranial nerve VII), and the inferior alveolar nerve divisions of the mandibular nerve are most affected. Although the rate of permanent paresthesia in third molar extraction is relatively low at 0.33%, with ~3.5 million lower third molar extractions performed annually in the United States, the incidence of permanent paresthesia can be 11,500–35,000 persons per year.<sup>3–8</sup> The relatively high incidence rate of these injuries, in association with their notable pain and loss of function, leads to substantial negative effects on the quality of life.<sup>9</sup>

**Table 1: MRN protocol on 3T scanner (Ingenia, Achieva)**

Sequence	TR/TE (ms)	Section Thickness (mm)	Matrix	FOV (cm)	Comments	Acquisition Time (min:sec)
Axial T2-weighted SPAIR	2000/60	3.0	268 × 248	16	Corpus callosum to chin	5:20
Axial T1-weighted	580/9	3.0	320 × 310	16	Corpus callosum to chin	5:10
Axial 3D balanced FFE	5.32/2.66	0.65	270 × 270	16	Corpus callosum to chin	6:00
Axial DTI	14,000/70	5.0	196 × 192	18	Skull base to chin; b-values = 0 and 600 s/mm <sup>2</sup> ; 12 directions	7:00
Coronal 3D STIR (optional)	1500/78	1.5 (Isotropic Voxel)	...	20	Corpus callosum to chin	7:15
Coronal 3D PSIF	12/2.5	0.9 (Isotropic Voxel)	...	20	Corpus callosum to chin	7:30

**Note:**—FFE indicates fast-field echo; SPAIR, spectral attenuated inversion recovery; STIR, short tau inversion recovery; PSIF, diffusion-weighted reversed fast imaging with steady-state precession.

The current reference standard for the diagnosis and monitoring of PTNs is centered around neurosensory testing (NST) in combination with the clinical findings. With NST, 3 nerve-function domains are evaluated, including the spatiotemporal sensory domain, monofilament contact detection, and pain tolerance/thresholds.<sup>8</sup> NST results are used to determine the severity of nerve injury based on the Sunderland classification criteria, which categorize nerve injuries in classes I–V, depending on the damage to the components of the nerve structure (myelin loss, axonal damage, and damage to the endoneurium, perineurium, and epineurium in that order).<sup>10,11</sup> While NST has been useful in the diagnosis and monitoring of PTNs, various limitations exist, such as operator dependence, patient subjectivity, minimal anatomic specificity, and the inability to determine patient recovery after nerve surgery.<sup>8,12</sup> MR neurography (MRN), an imaging technique dedicated to the evaluation of peripheral nerves, provides a 3D map of the neural anatomy like MRA. It has been validated for determining the Sunderland classification of nerve injuries preoperatively, with moderate-to-good correlation with NST and surgical findings for both the inferior alveolar nerve and lingual nerve injuries.<sup>13–15</sup> MRN results are often achieved earlier in the posttraumatic setting than NST results, aiding in the prompt diagnosis of more severe injuries and improved patient outcomes.<sup>16</sup>

Current research on the utility of MRN in the management of nerve injuries including PTN has largely focused on its uses in the preoperative diagnostic setting. NST is extremely limited and not useful in the postoperative setting due to variable pain and sensory responses. In this clinical report, we aimed to determine the utility of MRN in monitoring of structural changes following nerve-repair procedures. We hypothesized that postoperative MRN can assist in determining nerve integrity, and worsening MRN findings will corroborate poor patient outcomes.

## CASE SERIES

This report was retrospective and performed under institutional review board approval. Informed consent was waived, and all Health Insurance Portability and Accountability Act of 1996 regulations were followed.

## MATERIALS AND METHODS

### Patient Population and Study Sample

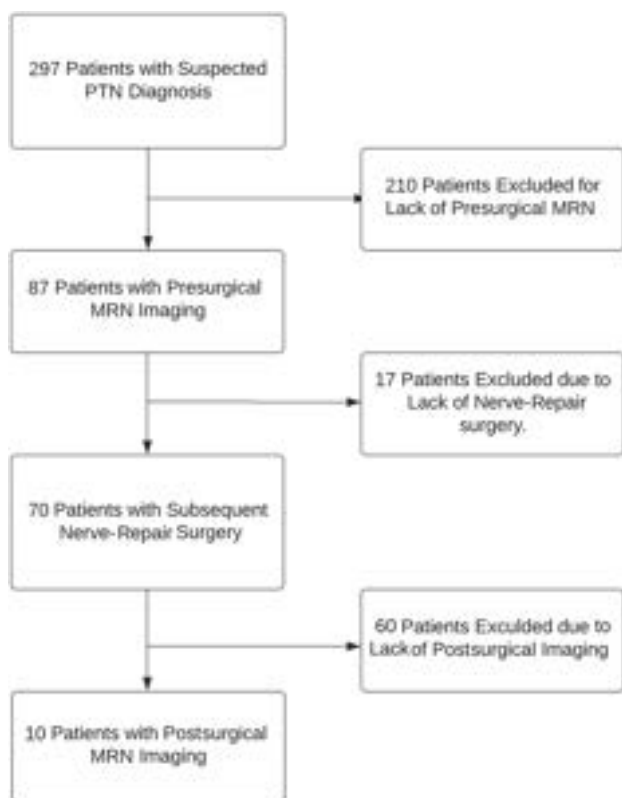
Two hundred ninety-seven consecutively sampled adult patients of both sexes with suspected PTNs were retrospectively queried from the institution database during an 8-year period (July 2015

to September 2023). All these patients were referred from the institutional Oral and Maxillofacial Surgery Clinic. They were evaluated by an experienced maxillofacial surgeon (J.R.Z.) and underwent trigeminal NST with subsequent trigeminal MRN imaging. From this patient population with PTN, 210 patients were excluded because they did not undergo presurgical MRN, and an additional 17 were excluded because they did not have a corresponding nerve-repair procedure. Of the 70 remaining patients, 60 patients were excluded due to lack of postsurgical MRN. The 10 remaining patients had postoperative MRN, all interpreted by a single fellowship-trained radiologist (A.C.) with >15 years of experience in musculoskeletal radiology and extensive expertise in MRN.

Patient charts were evaluated for the inciting event and pre-surgical clinical and NST findings with their corresponding Sunderland classification (classes I–V). The clinical findings captured in this review included pain, sensory abnormalities, taste changes, and associated functional changes (ie, lip biting and speech challenges). All patients underwent presurgical MRN at the same institution using a uniform protocol and a 32-channel head coil, with interpretation performed blinded to clinical NST findings and before surgical intervention (see Table 1 for MRN protocol details). The scans were obtained on 3T and 1.5T scanners (Achieva and Ingenia; Philips Healthcare), and MIP reconstructions were performed using a slab thickness of 10 mm in the oblique coronal and sagittal planes along the respective nerve course for its long-axis depiction. A review of presurgical MRN imaging identified the severity of nerve injury by the Neuropathy Score Reporting and Data System Injury (NS-RADS I 1–5) criteria.<sup>17</sup> Surgical notes were then reviewed for procedures performed, surgical findings, time elapsed from inciting event to surgery date, and injury severity according to the Sunderland classification.

When we reviewed Sunderland classifications, findings of patients not specifically stratified into 1 class (for example, class III/IV or class IV/V) at the time of the initial presentation were recorded as inconclusive, and patients who were not clinically tested for various reasons (severe pain or inability to open the mouth in acute injury) were considered unclassified for the purpose of this clinical report. Inconclusive MRN classifications were re-evaluated and assigned a specific nerve injury classification on re-evaluation by an experienced radiologist.

A corresponding review of postoperative MRN findings for each patient was performed and compared with the postoperative clinical findings. Postsurgical MRN findings were



**FIG 1.** Patient population in this study.

categorized by the Neuropathy Score Reporting and Data System Postintervention (NS-RADS PI) score, with NS-RADS PI-1 showing expected postsurgical findings, NS-RADS PI-2 demonstrating possible persistent neuropathy, and NS-RADS PI-3 demonstrating definitive recurrent or persistent neuropathy, including the formation of new neuromas, in accordance with the MR imaging reporting guidelines of peripheral neuropathy.<sup>17</sup> Postsurgical MRN was interpreted by the same expert radiologist blinded to postsurgical clinical information and outcomes.

### Clinical and Surgical Classifications

All NSTs and surgical procedures were performed by the same experienced maxillofacial surgeon using the same allograft technique for nerve reconstructions, apart from neurolysis. Postsurgical clinical findings evaluated were identical to their presurgical counterparts and were categorized on the basis of their status (no change, partial improvement, complete resolution).

### Statistical Analysis

Descriptive statistics were used for demographic information (age, sex, and so forth), inciting injury, injury severity, and comparing postsurgical neuropathy recovery with findings on postsurgical MRN using Neuropathy Score Reporting and Data System scoring and clinical outcomes. We only obtained descriptive statistics due to having a small sample with 3 different subcategories of postintervention (PI) states. Concordance was defined as resolution or partial improvement of symptoms in an MRN-improved state (PI-1) and persistence or worsening of symptoms in PI-2/3 states.

## RESULTS

### Patient Population and Study Sample

Of the continuously sampled 297 patients with PTN initially reviewed, 227 were excluded due to lack of surgical intervention, and 60 of the remaining 70 patients were excluded because they had no postsurgical MRN (See Fig 1 for details of exclusion and inclusion criteria of the patient population). Among the final sample of 10 patients reviewed, 6/10 were women with mean ages of 44.8 (SD, 16.8) years with an age range of 25–73 years for all patients. Iatrogenic trauma was the inciting event for all patients, with third molar removal being the most common cause, occurring in 8/10 patients. The remaining 2 patients underwent multiple surgeries: one having various procedures and revisions involving the temporomandibular joint and the other having prior nerve-repair procedures related to the inferior alveolar nerve and mental nerve. The lingual nerve was injured in 8/10 patients, while injury to the mental nerve and infraorbital nerve was present in 1 patient each. The presurgical neuropathy symptoms of hypoesthesia/anesthesia and burning pain were each found in 7 patients, while hypogeusia/ageusia and dysgeusia were found in 6 and 4 patients, respectively. Presurgical neuropathy functional symptoms of lip biting were found in 7 patients. Speech difficulties were noted in 3 patients, and chewing challenges were seen in 2 patients.

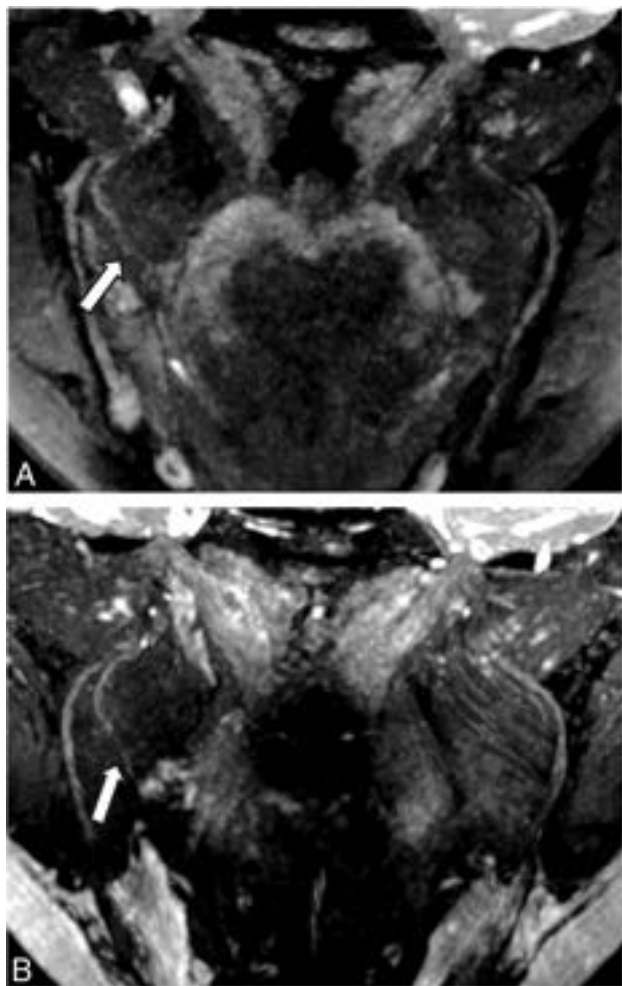
### Clinical and Surgical Classifications

Of the patients with presurgical NST, 8/10 patients were classified by the Sunderland classification criteria, with 6/10 being assigned a conclusive score. Of these 6, one patient had a class V injury, 3 patients were diagnosed with class IV injuries, and 1 patient each was diagnosed with class III and II injuries, respectively. A total of 13 presurgical MRN reports were evaluated for these 10 patients. Eight patients were imaged on 3T scanners, and 2 were imaged on 1.5T scanners, using the same protocols. On presurgical MRN, an NS-RADS Injury (I) score was assigned to all reports except 2, which were initially inconclusive and were subsequently revised by the same expert radiologist following re-assessment blinded to the surgical findings (NS-RADS I-3/4 changed to NS-RADS I-4, NS-RADS I-2/3 changed to NS-RADS I-4). Nerve injury on presurgical MRN ranged from NS-RADS I-4 to NS-RADS I-5 injuries, with NS-RADS I-4 injuries being the most common and presenting in 8/10 patients (Figs 2–6). No patients demonstrated tongue muscle atrophy on imaging, but 1 patient was noted to have salivary gland atrophy on the ipsilateral side of the injury on MRN.

All 10 patients underwent surgery following their presurgical clinical and MRN evaluations; these procedures included microdissection with neurolysis and neuroma excision in 2 patients, neuroma excision in 2 patients, and neurorrhaphy with allograft and Axoguard Nerve Protector (Axogen) in the 6 remaining patients. A conclusive nerve injury classification was described in each surgical case, with 8 demonstrating class IV injuries and 2 showing class V injuries. The mean time from the inciting event to surgery for patients with a single inciting event was 6.1 (SD, 4.6) months. The remaining 2 patients with multiple surgeries between the inciting event and the evaluated nerve reconstruction surgery in this report had a



duration of 54 and 242 months, respectively. The mean time from presurgical MRN to surgery was 46.4 (SD, 28.0) days, with a range of 11–86 days.



**FIG 2.** NS-RADS PI-1. A 54-year-old woman, status post third molar removal with decreased sensation, burning pain, and dysgeusia. A, Presurgical coronal MRN MIP image 89 days status post inciting event shows a neuroma in continuity of the right lingual nerve (NS-RADS I-4, arrow). B, Postsurgical coronal corresponding MRN 351 days following right lingual nerve neuroma excision and neurorrhaphy with allograft and Axoguard placement demonstrates the expected postsurgical appearance of the nerve (NS-RADS PI-1) with no loss of continuity, neuroma reformation, or substantial nerve-caliber changes (arrow).

There were 13 postsurgical MRNs (all performed on 3T scanners) for 10 patients. These were assessed by the same expert radiologist blinded to patient outcomes. Of these, 3 patients had postsurgical MRN findings consistent with NS-RADS PI-1 (Fig 2), 3 patients had findings consistent with NS-RADS PI-2 (Figs 4 and 6), and findings of 4 were consistent with NS-RADS PI-3 (Figs 3 and 5). The mean time from surgery to postsurgical MRN was 186.5 (SD, 103.0) days, with a range of 68–353 days.

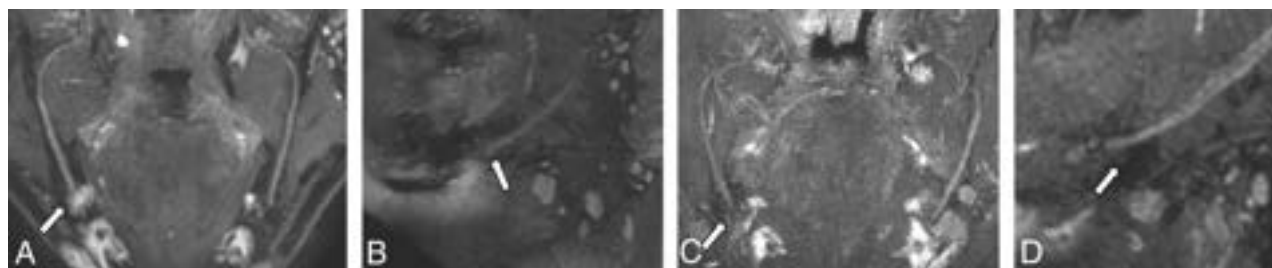
The mean postsurgical follow-up of patients was 360.2 (SD, 249.3) days, ranging from 97 to 832 days. The postsurgical clinical status of each patient is described in Table 2.

Of the 3 patients who had postsurgical MRN findings consistent with NS-RADS PI-1, two experienced partial improvement and 1 experienced complete resolution of the clinical symptoms. Of the 3 patients with NS-RADS PI-2 findings, one experienced no postsurgical change in clinical symptoms and 2 demonstrated partial improvement. One patient's improvement was limited to lip biting and speech, with no changes in pain, sensation, or taste, while the other had no pain on initial presurgical presentation and demonstrated no improvement in tongue or taste sensation. No postsurgical change was noted in 3 of 4 patients with NS-RADS PI-3 findings, with the final patient demonstrating partial improvement in lip biting and speech challenges with, however, no sensation improvement or taste changes. Specific improvement for each clinical finding is further described in Table 2.

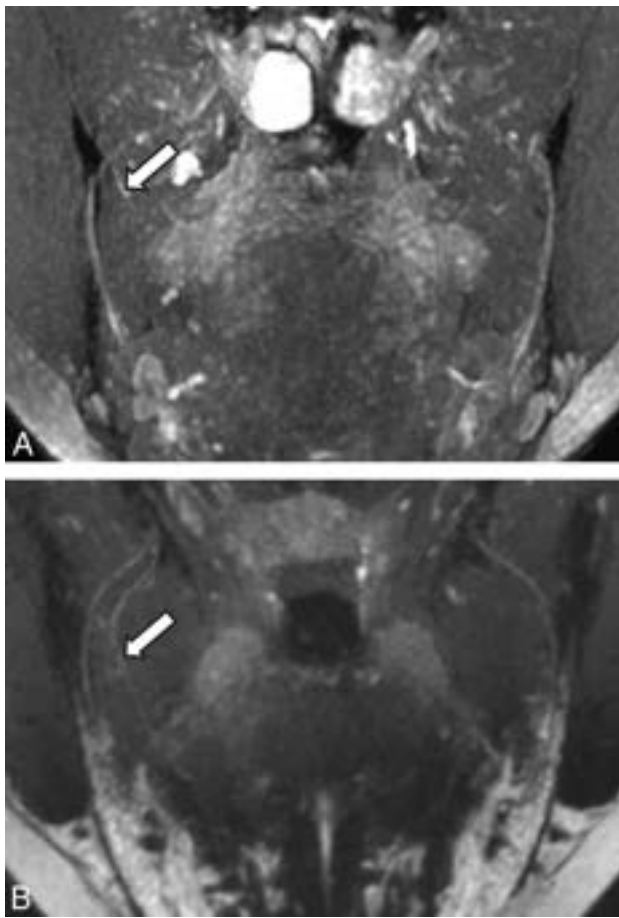
## DISCUSSION

This is the initial report of its kind to evaluate the utility of MRN in the postsurgical patient population following nerve-repair procedures on small nerves of the jaw. With the current clinical management of such patients in the postsurgical setting being restricted to mostly clinical and NST findings, this report demonstrates a meaningful precedence for its utility and further exploration in larger or multicenter studies.

Similar to how MRN has demonstrated utility in aiding in an early diagnosis of PTNs, this clinical report demonstrates its promising potential in the setting of nerve-repair recovery. Postsurgical MRN examinations were not degraded by surgical changes and consistently demonstrated both healing and worsening structural changes following nerve-repair surgery. These pertinent MRN findings included the identification of expected

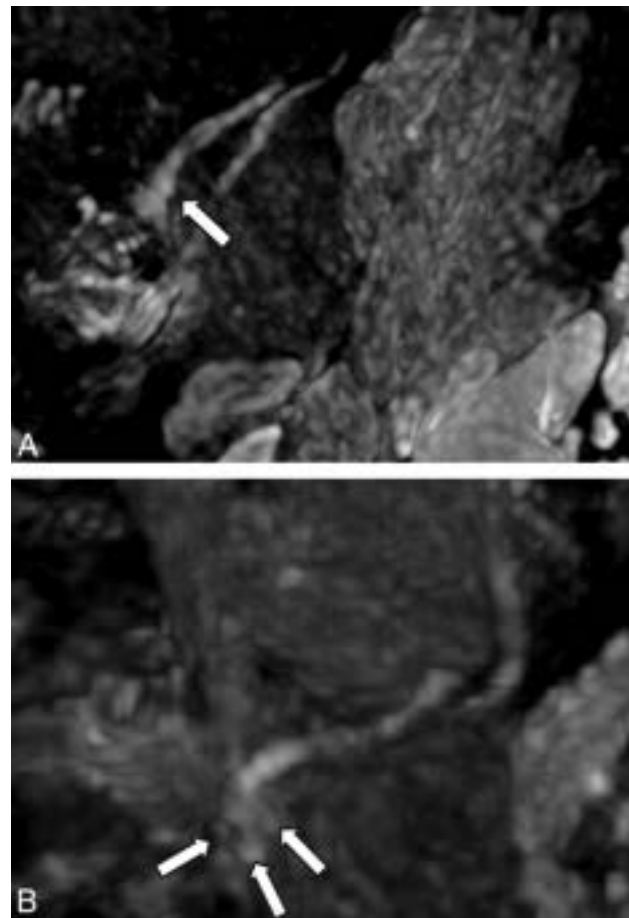


**FIG 3.** NS-RADS PI-3. A 64-year-old man with a history of multiple nerve-repair procedures of the inferior alveolar nerve and mental nerve with a history of burning pain, lip biting, and speech difficulties. A and B, Presurgical coronal and sagittal MRN MIP images demonstrate a right mental nerve lateral neuroma in continuity (NS-RADS I-4, arrows). C and D, Postsurgical coronal and sagittal MRN MIP images 353 days following neuroma excision and neurorrhaphy with allograft and Axoguard placement show a recurrent right mental nerve neuroma in continuity (arrows).



**FIG 4.** NS-RADS PI-2. A 27-year-old woman 214 days status post third molar removal with decreased sensation, burning pain, and hypogeusia. *A*, Presurgical coronal MRN MIP image shows the right lingual nerve demonstrating nerve-caliber focal thickening and increased signal instead of uniformly distally decreasing nerve caliber, consistent with a neuroma in continuity (NS-RADS I-4, *arrow*). *B*, Postsurgical coronal MRN MIP image 98 days status post right lingual nerve neuroma excision and neurorrhaphy with allograft and Axoguard placement demonstrates incomplete regeneration, ie, minimal increased residual signal of the nerve without a new neuroma with minor caliber change compared with preoperative MRN, consistent with NS-RADS PI-2 findings (*arrow*).

postsurgical findings and pathology recurrence in the form of neuroma reformation, persistent nerve discontinuity, or perineural fibrosis. These initial MRN results additionally demonstrated a good concordance with the clinical outcomes experienced by patients. This result was because all 3 patients with NS-RADS PI-1 findings experienced partial improvement or complete resolution, whereas only 1 of 3 patients with NS-RADS PI-2 and 1 of 4 patients with NS-RADS PI-3 findings had similar improvement. Of the 3 patients who demonstrated NS-RADS PI-1 findings, all 3 experienced a concordant improvement in sensation and pain and 2 demonstrated improvement in taste sensation as well. Of interest, lip biting and speech difficulties appeared to improve regardless of the postsurgical MRN findings. This result may be because mechanical issues could improve with behavioral and speech therapies, though we did not have the full details of such treatments in the patients’

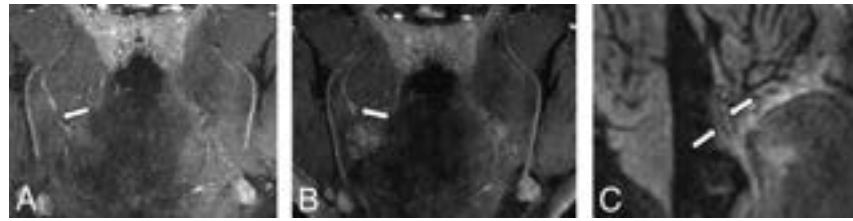


**FIG 5.** NS-RADS PI-3. A 21-year-old woman status post third molar removal experiencing decreased sensation, burning pain, and hypogeusia. *A*, Presurgical sagittal MRN MIP 55 days status post third molar removal shows right lingual nerve neuroma (NS-RADS I-4, *arrow*). *B*, Postsurgical sagittal MRN MIP 166 days status post neuroma excision and neurorrhaphy with allograft and Axoguard placement demonstrates the re-formation of multiple neuromas (*arrows*).

charts. The somatic and special sensory responses seem to be more resistant to optimal improvement.<sup>18</sup>

The need for follow-up MRN analysis is encouraged by our findings that showed good corroboration with the clinical outcomes. Just as initial nerve injuries have a spectrum of outcomes, nerve-repair surgical outcomes often vary due to the severity of initial injury, the complexity/invasiveness of the procedure, and the subtle complexities of healing inherent in each patient. These challenges have been described in multiple nerve-repair clinical outcome reviews and further substantiate the need for improved visualization of the nerve-healing processes.<sup>19–22</sup> This initial report demonstrates the promising ability of MRN to fill this need and aid in the clinical decision-making of PTN management.

In this early analysis of the utility of MRN in the postsurgical setting, an obvious limitation is the small case sample size. This is, in part, due to MRN being a relatively novel imaging technique, requiring unique radiologic training, which is not widespread among practicing radiologists. Although presurgical MRN has been increasingly used and shown to correlate well with both clinical NST and surgical findings,<sup>1</sup> postsurgical MRN



**FIG 6.** NS-RADS PI-2. A 31-year-old man status post third molar removal experiencing decreased sensation and ageusia. A and B, Presurgical coronal MRN MIP 54 days status post third molar removal shows a right lingual nerve end-bulb neuroma with complete transection with no distal continuity (NS-RADS I-5, arrows). C, Postsurgical coronal MRN MIP 238 days status post neuroma excision and neurorrhaphy with allograft and Axoguard placement demonstrates partial regeneration (arrows).

**Table 2: Postsurgical MRN NS-RADS PI distribution among 10 patients with the number of patients experiencing clinical improvement in specified neuropathic symptoms and overall clinical outcome**

NS-RADS PI	Patient Count	Patients with Postsurgical Clinical Improvement				Nerve-Repair PTN Outcome		
		Pain	Sensation	Taste	Lip Biting-Speech	Complete Resolution	Partial Improvement	No Change
NS-RADS PI-1	3	3	3	2	3	1	2	0
NS-RADS PI-2	3	0	0	0	2	0	1	2
NS-RADS PI-3	4	0	0	0	3	0	1	3

has not been routinely implemented. When used, postsurgical MRN is mostly used when the desired postsurgical results are not achieved.

Although nerve injuries of varying Sunderland classifications were analyzed in this report, there is a demonstrated shift toward more severe injuries (classes IV and V). This shift is expected because these injuries are more likely to require surgical intervention as opposed to milder injuries.

This report is an initial description of postoperative MRN results. Further analysis is needed to evaluate the generalizability of these findings. Additionally, more analysis is needed to assess postsurgical MRN utility in less-severe nerve injuries, though this is likely to be a persistent challenge given current recommendations for surgical interventions in the domain of peripheral nerve injuries. With a larger sample size, additional evaluation of the timeline should also be evaluated to better determine the capacity for MRN to elicit early postsurgical structural changes and furthermore determine a suggested timeframe for imaging and related management recommendations.

## ACKNOWLEDGMENTS

This work was partly supported by Stephen B. Milan Award from the Oral and Maxillofacial Surgery Foundation.

Disclosure forms provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

- Romano N, Federici M, Castaldi A. **Imaging of cranial nerves: a pictorial overview.** *Insights Imaging* 2019;10:33 CrossRef Medline
- Casselmann J, Mermuys K, Delanote J, et al. **MRI of the cranial nerves—more than meets the eye: technical considerations and advanced anatomy.** *Neuroimaging Clin N Am* 2008;18:197–231 CrossRef Medline
- Jerjes W, Upile T, Shah P, et al. **Risk factors associated with injury to the inferior alveolar and lingual nerves following third molar**

**surgery-revisited.** *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2010;109:335–45 CrossRef Medline

- Pogrel MA, Kaban LB. **Injuries to the inferior alveolar and lingual nerves.** *J Calif Dent Assoc* 1993;21:50–54 Medline
- Friedman JW. **The prophylactic extraction of third molars: a public health hazard.** *Am J Public Health* 2007;97:1554–59 CrossRef Medline
- Osborn TP, Frederickson G Jr, Small IA, et al. **A prospective study of complications related to mandibular third molar surgery.** *J Oral Maxillofac Surg* 1985;43:767–69 CrossRef Medline
- Valmaseda-Castellón E, Berini-Aytés L, Gay-Escoda C. **Inferior alveolar nerve damage after lower third molar surgical extraction: a prospective study of 1117 surgical extractions.** *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2001;92:377–83 CrossRef Medline
- Zuniga JR, Meyer RA, Gregg JM, et al. **The accuracy of clinical neurosensory testing for nerve injury diagnosis.** *J Oral Maxillofac Surg* 1998;56:2–8 CrossRef Medline
- Smith JG, Elias LA, Yilmaz Z, et al. **The psychosocial and affective burden of posttraumatic neuropathy following injuries to the trigeminal nerve.** *J Orofac Pain* 2013;27:293–303 CrossRef Medline
- Chhabra A, Ahlawat S, Belzberg A, et al. **Peripheral nerve injury grading simplified on MR neurography: as referenced to Seddon and Sunderland classifications.** *Indian J Radiol Imaging* 2014;24:217–24 CrossRef Medline
- Sunderland S. **A classification of peripheral nerve injuries producing loss of function.** *Brain* 1951;74:491–516 CrossRef Medline
- Robinson PP. **Observations on the recovery of sensation following inferior alveolar nerve injuries.** *Br J Oral Maxillofac Surg* 1988;26:177–89 CrossRef Medline
- Manoliu A, Ho M, Nanz D, et al. **MR neurographic orthopantomogram: ultrashort echo-time imaging of mandibular bone and teeth complemented with high-resolution morphological and functional MR neurography.** *J Magn Reson Imaging* 2016;44:393–400 CrossRef Medline
- Zuniga JR, AbdelBaky O, Alian A, et al. **Does pre-surgical magnetic resonance neurography predict surgical gap size in trigeminal class IV and V injuries?** *J Oral Maxillofac Surg* 2021;79:2574–81 CrossRef Medline
- Cox B, Zuniga JR, Panchal N, et al. **Magnetic resonance neurography in the management of peripheral trigeminal neuropathy: experience in a tertiary care centre.** *Eur Radiol* 2016;26:3392–400 CrossRef Medline

16. Dessouky R, Xi Y, Zuniga J, et al. **Role of MR neurography for the diagnosis of peripheral trigeminal nerve injuries in patients with prior molar tooth extraction.** *AJNR Am J Neuroradiol* 2018;39:162–69 CrossRef Medline
17. Chhabra A, Deshmukh SD, Lutz AM, et al. **Neuropathy Score Reporting and Data System: a reporting guideline for MRI of peripheral neuropathy with a multicenter validation study.** *AJR Am J Roentgenol* 2022;219:279–91 CrossRef Medline
18. Atkins S, Kyriakidou E. **Clinical outcomes of lingual nerve-repair.** *Br J Oral Maxillofac Surg* 2021;59:39–45 CrossRef Medline
19. Suhaym O, Miloro M. **Does early repair of trigeminal nerve injuries influence neurosensory recovery? A systematic review and meta-analysis.** *Int J Oral Maxillofac Surg* 2021;50:820–29 CrossRef Medline
20. Gregg JM. **Studies on traumatic neuralgias in the maxillofacial region: symptom complexes and responses to microsurgery.** *J Oral Maxillofac Surg* 1990;48:135–40; discussion 141 CrossRef Medline
21. Kogan M, Lee KC, Chuang SK, et al. **Outcomes of direct lingual nerve-repair after an injury: a systematic review.** *J Oral Maxillofac Surg* 2021;79:697–703 CrossRef Medline
22. Scholz J, Finnerup NB, Attal N, et al; Classification Committee of the Neuropathic Pain Special Interest Group (NeuPSIGT). **The IASP classification of chronic pain for ICD-11: chronic neuropathic pain.** *Pain* 2019;160:53–59 CrossRef Medline