

OCTOBER 2023

AJNR

VOLUME 44 • PP 1109-1239

AJNR

AMERICAN JOURNAL OF NEURORADIOLOGY

Official Journal ASNR • ASFNR • ASHNR • ASPNR • ASSR
OCTOBER 2023 | VOLUME 44 | NUMBER 10 | WWW.AJNR.ORG

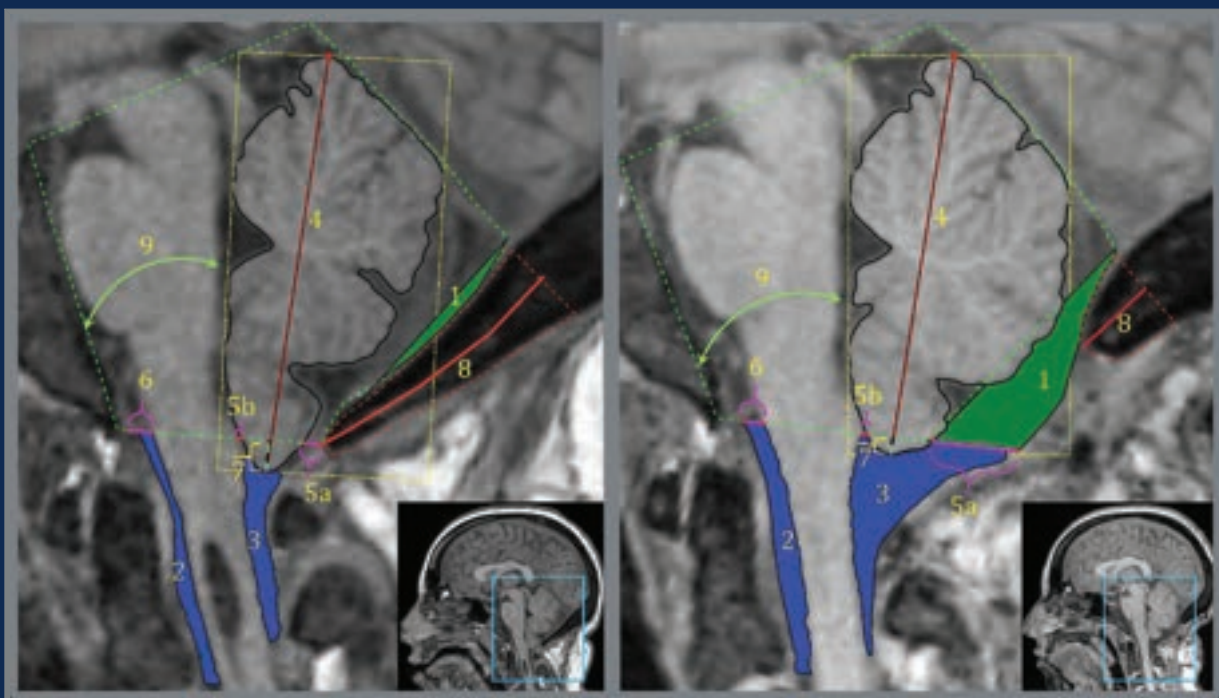
THE JOURNAL OF DIAGNOSTIC AND INTERVENTIONAL NEURORADIOLOGY

Peer learning in neuroradiology

Imaging with histopathologic and genetic correlation of nasal, paranasal, and skull base tumors

U-Net deep learning model for fetal pons and vermis localization

Performance of NI-RADS for recurrent head and neck SCC



FRED™ X™

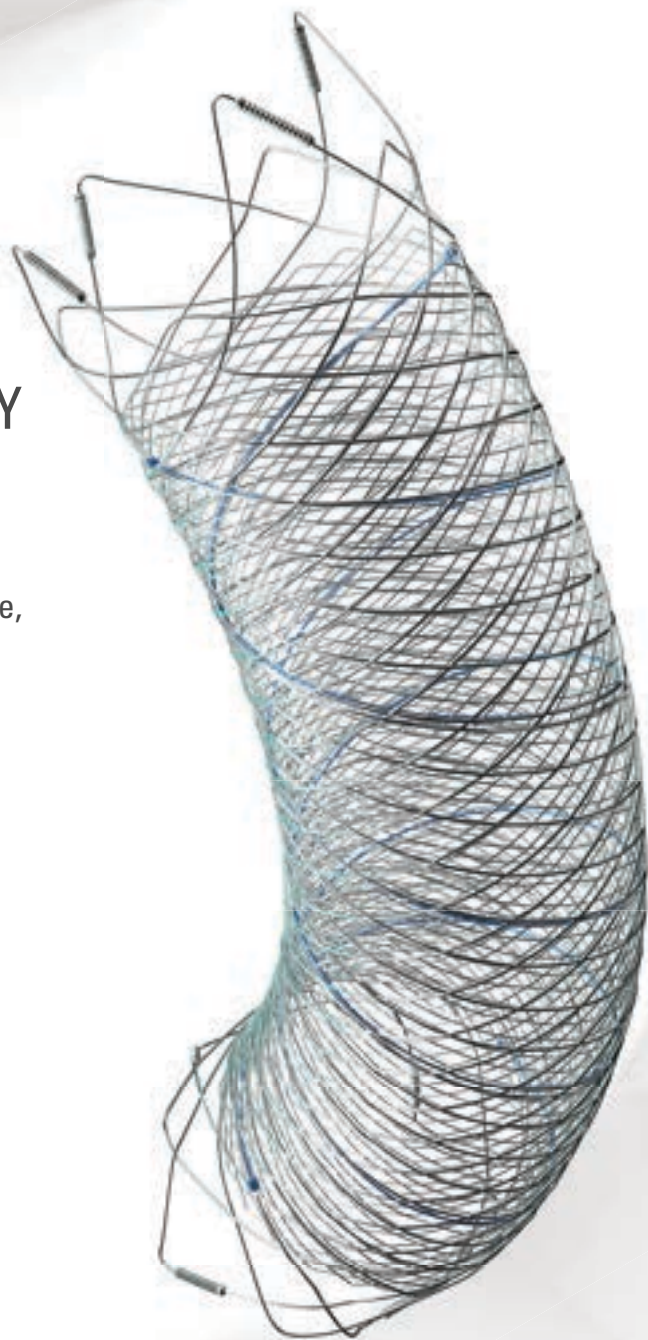
Flow Diverter Stent

THE NEXT ADVANCEMENT IN FLOW DIVERSION TECHNOLOGY

The FRED™ X Flow Diverter features the same precise placement and immediate opening of the FRED™ Device, now with X Technology. X Technology is a covalently bonded, nanoscale surface treatment, designed to:

- **REDUCE MATERIAL THROMBOGENICITY¹**
- **MAINTAIN NATURAL VESSEL HEALING RESPONSE^{2,3,4}**
- **IMPROVE DEVICE DELIVERABILITY AND RESHEATHING¹**

The only FDA PMA approved portfolio with a 0.021" delivery system for smaller device sizes, and no distal lead wire.



For more information, contact your local MicroVention sales representative or visit our website. www.microvention.com



^{*} Data is derived from in vivo and ex vitro testing and may not be representative of clinical performance.

¹ Data on file

² Tanaka M et al. Design of biocompatible and biodegradable polymers based on intermediate water concept. Polymer Journal. 2015;47:114-121.

³ Tanaka M et al. Blood compatible aspects of poly(2-methoxyethylacrylate) (PMEA) – relationship between protein adsorption and platelet adhesion on PMEA surface. Biomaterials. 2000;21:1471-1481.

⁴ Schiel L et al. X Coating™: A new biopassive polymer coating. Canadian Perfusion Canadienne. June 2001;11(2):9.

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

Rx Only: Federal (United States) law restricts this device to sale by or on the order of a physician. For Healthcare professionals intended use only.

MICROVENTION, FRED and HEADWAY are registered trademarks of MicroVention, Inc. in the United States and other jurisdictions. Stylized X is a trademark of MicroVention, Inc. © 2022 MicroVention, Inc. MM1222 US 03/22

WEB™ 17

Aneurysm Embolization System

LOWER PROFILE



NEW SIZES



MORE ACCESS OPTIONS



INDICATIONS FOR USE:

The WEB Aneurysm Embolization System is intended for the endovascular embolization of ruptured and unruptured intracranial aneurysms and other neurovascular abnormalities such as arteriovenous fistulae (AVF). The WEB Aneurysm Embolization System is also intended for vascular occlusion of blood vessels within the neurovascular system to permanently obstruct blood flow to an aneurysm or other vascular malformation.

POTENTIAL COMPLICATIONS:

Potential complications include but are not limited to the following: hematoma at the site of entry, aneurysm rupture, emboli, vessel perforation, parent artery occlusion, hemorrhage, ischemia, vasospasm, clot formation, device migration or misplacement, premature or difficult device detachment, non-detachment, incomplete aneurysm filling, revascularization, post-embolization syndrome, and neurological deficits including stroke and death. For complete indications, potential complications, warnings, precautions, and instructions, see instructions for use (IFU provided with the device).

VIA 21, 27, 33 - The VIA Microcatheter is intended for the introduction of interventional devices (such as the WEB device/stents/flow diverters) and infusion of diagnostic agents (such as contrast media) into the neuro, peripheral, and coronary vasculature.

VIA 17, 17 Preshaped - The VIA Microcatheter is intended for the introduction of interventional devices (such as the WEB device/stents/flow diverters) and infusion of diagnostic agents (such as contrast media) into the neuro, peripheral, and coronary vasculature.

The VIA Microcatheter is contraindicated for use with liquid embolic materials, such as n-butyl 2-cyanoacrylate or ethylene vinyl alcohol & DMSO (dimethyl sulfoxide).

The device should only be used by physicians who have undergone training in all aspects of the WEB Aneurysm Embolization System procedure as prescribed by the manufacturer.

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

For healthcare professional intended use only.



MicroVention Worldwide
Innovator Center PH +1.714.247.8000

35 Enterprise
Aliso Viejo, CA 92656 USA
MicroVention UK Limited
MicroVention Europe, S.A.R.L.
MicroVention Deutschland GmbH
Website
PH +44 (0) 191 258 6777
PH +33 (1) 39 21 77 46
PH +49 211 210 798-0
microvention.com



WEB™ and VIA™ are registered trademarks
of Sequent Medical, Inc. in the United States.

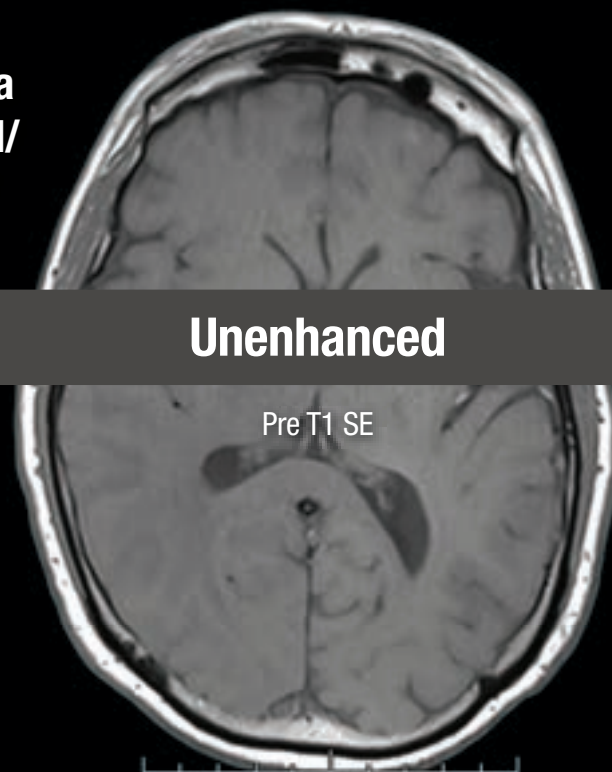
©2021 MicroVention, Inc. MM1184 WW 11/2021

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.^{1-6†}**


Vueway®
[gadopicleinol] 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 (gadopicleinol) 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® (gadopicleinol) 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: NEPHROGENIC SYSTEMIC FIBROSIS (NSF)

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 m²), 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

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 history of a reaction to contrast media, bronchial asthma and/or allergic disorders, administer VUEWAY only in



LIFE FROM INSIDE

65-year-old man – 3.0 T Siemens
Brain metastasis from lung adenocarcinoma⁷

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

Gadopicleenol 0.05 mmol/kg

Post T1 SE

Gadobutrol 0.1 mmol/kg

Post T1 SE

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.

Ensure catheter and venous patency before injecting as **extravasation** may occur, and cause tissue irritation.

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 or call 1-800-FDA-1088.

Please see BRIEF SUMMARY of Prescribing Information for VUEWAY, 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® (gadopicleenol) solution for injection, 485.1 mg/mL Full Prescribing Information and Patient Medication Guide. Monroe Twp., NJ: Bracco Diagnostics Inc.; September 2022. 2. Robic C, Port M, Rousseaux O, et al. Physicochemical and pharmacokinetic profiles of gadopicleenol: 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.; June 2022. 7. Loevner LA, Kolumban B, Hutóczki G, et al. Efficacy and safety of gadopicleenol 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

© 2023 Bracco Diagnostics Inc.
All Rights Reserved. US-VW-2300022 08/23



VISIT
VUEWAY.COM
FOR MORE
INFORMATION

Vueway™ (gadopiclenol) injection, for intravenous use

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

WARNING: NEPHROGENIC SYSTEMIC FIBROSIS (NSF)

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.

- The risk for NSF appears highest among patients with:
 - Chronic, severe kidney disease (GFR < 30 mL/min/1.73 m²), 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 (5.1) 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 gadopicholol.

WARNINGS AND PRECAUTIONS

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 m²) 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 m²) and little, if any, for patients with chronic, mild kidney disease (GFR 60-89 mL/min/1.73 m²). 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).

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 gadopicholol.

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 (5.1) 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 (5.1) in the full Prescribing Information]
- Hypersensitivity Reactions [see Contraindications (4) and Warnings and Precautions (5.2) 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 gadopicholol showed some signs of maternal toxicity in rats 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 gadopicholol 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. Gadopicholol 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 [¹⁵³Gd]-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 gadopicholol 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 gadopicholol 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 (5.1, 5.3, 5.4) and Clinical Pharmacology (12.3) in the full Prescribing Information].

OVERDOSAGE

Among subjects who received a single 0.3 mmol/kg intravenous dose of gadopicholol (6 times the recommended dose of Vueway), headache and nausea were the most frequently reported adverse reactions. Gadopicholol 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 (5.1) 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 (5.3) 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 (5.5) 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 November 2022

ASNR 62nd Annual Meeting Call for Abstracts

Join us May 18-22, 2024, in Las Vegas to present the best scientific research in neuroradiology.

ASNR is looking for abstract submissions on the topics of Adult Brain, Spine, Head and Neck, Pediatrics, Functional/Advanced Imaging, Interventional, Health Policy, and AI/Informatics.

ABSTRACT SUBMISSION DEADLINE

Wednesday, November 1, 2023 (11:59 PM ET)

Submit online at <http://www.aievolution.com/asn2401/>

Acceptance notifications will be sent on or before February 1, 2024, upon conclusion of peer review.

ABSTRACT SUBMISSION CATEGORIES

- **Scientific Abstract Presentations:** *New in 2024 -- Select your preferred format: Oral Presentation with slide summary OR Electronic Poster only*
- **Electronic Excerpta Poster**
- **Electronic Educational Exhibit**

ABSTRACT SUBMISSION INFORMATION AND CRITERIA

- **By submitting an abstract, you agree that the presenting author will register for the meeting at their own expense.** Registration type is dependent on the type of abstract you are presenting (in-person or virtual).
- ASNR requires disclosure from all persons in control of content, including all authors on abstracts. When added to an abstract, all co-authors on your submission will be notified to log-in and complete their own disclosure form. Once all co-authors have submitted disclosure, you will be able to finalize your submission. You will not be able to finalize your submission until all authors have completed their disclosure forms. If any disclosures cannot be submitted by the deadline, remove authors with missing disclosures and submit the abstract. Authors may be added after submission if disclosure is completed. Email education@asnr.org to add authors after the submission deadline.
- Maximum length, not including spaces, title, authors, or image: 2,500 characters (combined from all text fields).
- The submission site allows uploading of a **single JPG file** with each submission. Only one graphic file may be submitted with each abstract. This single JPG can contain multiple images, graphics, and tables.
- At least one reference is required, and up to five references may be included. References should be formatted according to *AJNR* author instructions.
- **DO NOT SUBMIT DUPLICATE ABSTRACTS IN MULTIPLE CATEGORIES OR SUBMISSION TYPES.**
- Changes can be made to submitted abstracts until the deadline.
- ASNR encourages presenters to submit manuscripts based on their work to the *American Journal of Neuroradiology* or *Neurographics* before considering other journals.
- Available ASNR awards include the David M. Yousem Research Fellow Award, Cornelius Dyke Memorial Award, Outstanding Presentation Awards, and Educational Exhibit Awards.

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

Visit www.asnr.org/annualmeeting for complete submission details

JOIN US! COMPREHENSIVE NEURORADIOLOGY COURSE

February 1-3, 2024 • Cancun, Mexico / On-Demand

ASNR's Comprehensive Neuroradiology Course offers you three days of image-rich neuroradiology review focused on:

**Adult Brain Imaging | Head and Neck Imaging | Spine Imaging
Pediatric Neuroradiology | Advanced Imaging Techniques**

Join our award-winning faculty and engage in our interactive classroom environment, as well as during planned social and evening activities. Take advantage of a discounted, all-inclusive rate at the host hotel: the Hyatt Ziva Cancun!

Can't attend in person? An on-demand option is also available!

**Get all of the details and register now at
www.asnr.org/cnc24.**

Meet the Faculty

Joshua Nickerson, MD

Division Chief of Neuroradiology, Associate Professor, Oregon Health & Science University

Tabassum Kennedy, MD

Division Chief of Neuroradiology, University of Wisconsin, Madison

Judith Gadde, DO, MBA

Pediatric Neuroradiologist and Director of Academic Innovation, Lurie Children's Hospital of Chicago

Mahmud Mossa-Basha, MD

Professor of Radiology, Vice Chair of Quality & Safety, and Medical Director of MRI, University of North Carolina School of Medicine

Wende Gibbs, MD, MA

Director of Spine Imaging and Intervention, Barrow Neurological Institute

Ashley Aiken, MD,

Director of Head and Neck Imaging and Professor of Radiology and Imaging Sciences, Emory University

AJNR *go green*

***AJNR* urges American Society of Neuroradiology members to reduce their environmental footprint by voluntarily suspending their print subscription.**

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

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

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

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

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

ASNR members who wish to opt out of print can do so by using the *AJNR* Go Green link on the *AJNR* Website (<http://www.ajnr.org/content/subscriber-help-and-services>). Just type your name in the email form to stop print and spare our ecosystem.



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

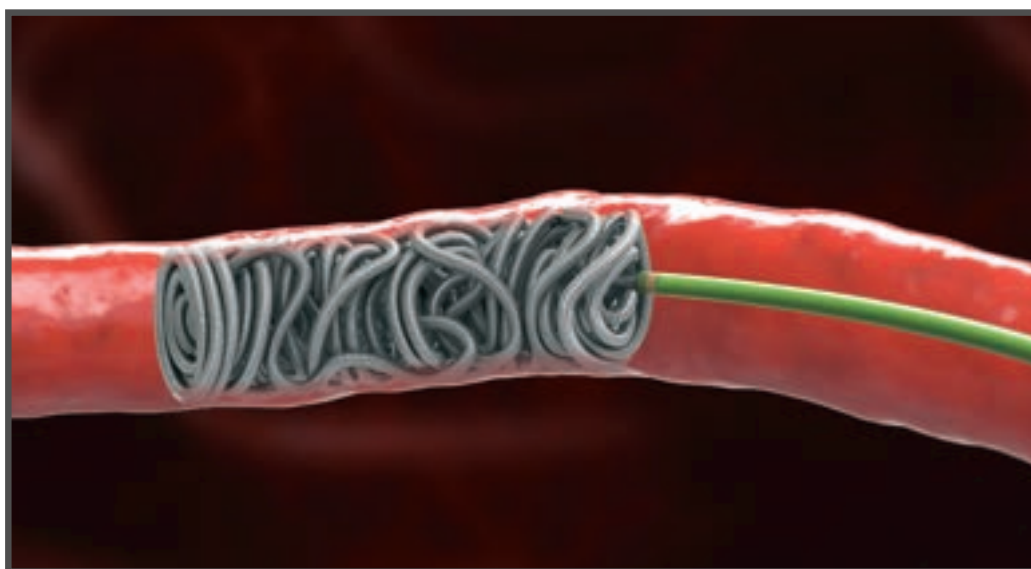
- Employer resources to help you recruit top talent
- Multiple pricing options, including free Fellowship listings
- Resume search

Start here: careers.asnr.org

OPTIBLOCKTM

COIL

PRECISION PLACEMENT



EFFICIENT OCCLUSION

- Precision Anchoring • Concise, Dense Pack • Unique Design and Long Lengths
- $\geq .0165$ Microcatheter Compatible

Visit our products page at baltgroup.com for more information.

Balt USA
29 Parker, Irvine, CA 92618 • tel 949.788.1443 fax 949.788.1444
baltgroup.com
© 2023 BALT USA MKTG-370 Rev. A



AJNR

AMERICAN JOURNAL OF NEURORADIOLOGY

OCTOBER 2023
VOLUME 44
NUMBER 10
WWW.AJNR.ORG

Publication Preview at www.ajnr.org features articles released in advance of print. Visit www.ajnrblog.org to comment on AJNR content and chat with colleagues and AJNR's News Digest at <http://ajnrdigest.org> to read the stories behind the latest research in neuroimaging.

PERSPECTIVES

- 1109 **Peer Learning in Neuroradiology: Not as Easy as It Sounds** *K. Mani, et al.*

REVIEW ARTICLE



- 1116 **Update from the 5th Edition of the WHO Classification of Nasal, Paranasal, and Skull Base Tumors: Imaging Overview with Histopathologic and Genetic Correlation** *A. Agarwal, et al.*

HEAD & NECK

GENERAL CONTENTS



- 1126 **Systematic Literature Review of Machine Learning Algorithms Using Pretherapy Radiologic Imaging for Glioma Molecular Subtype Prediction** *Jan Lost, et al.*

ADULT BRAIN



- 1135 **Segmentation of Brain Metastases Using Background Layer Statistics (BLAST)** *Chris Heyn, et al.*

ADULT BRAIN



- 1144 **Cortical and Subcortical Brain Atrophy Assessment Using Simple Measures on NCCT Compared with MRI in Acute Stroke** *Tanaporn Jaroengarmsamer, et al.*

ADULT BRAIN



- 1150 **Prospective, Longitudinal Study of Clinical Outcome and Morphometric Posterior Fossa Changes after Craniocervical Decompression for Symptomatic Chiari I Malformation** *Alaaddin Ibrahimy, et al.*

ADULT BRAIN



- 1157 **Progressive Changes in Cerebral Apparent Diffusion Values in Fabry Disease: A 5-Year Follow-up MRI Study** *Koen P.A. Baas, et al.*

ADULT BRAIN



- 1165 **Partial (SAVE) versus Complete (Solumbra) Stent Retriever Retraction Technique for Mechanical Thrombectomy: A Randomized In Vitro Study** *Magda Jablonska, et al.*

INTERVENTIONAL

- 1171 **Peritumoral Signal on Postcontrast FLAIR Images: Description and Proposed Biomechanism in Vestibular Schwannomas** *John C. Benson, et al.*

HEAD & NECK



- 1176 **Differentiation between Chondrosarcoma and Synovial Chondromatosis of the Temporomandibular Joint Using CT and MR Imaging** *B.G. Jang, et al.*











HEAD & NECK



- 1184 **Performance of Neck Imaging Reporting and Data System (NI-RADS) for Diagnosis of Recurrence of Head and Neck Squamous Cell Carcinoma: A Systematic Review and Meta-analysis** *Akira Baba, et al.*

HEAD & NECK

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, Denver, NJ 07834, U.S.A. Subscription rates: nonmember \$452 (\$530 foreign) print and online, \$320 online only; institutions \$520 (\$594 foreign) print and basic online, \$1029 (\$1103 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.

- | | | | |
|---|------|---|-------------------|
|  | 1191 | Automatic Localization of the Pons and Vermis on Fetal Brain MR Imaging Using a U-Net Deep Learning Model <i>Farzan Vahedifard, et al.</i> | PEDIATRICS |
|  | 1201 | Dandy-Walker Phenotype with Brainstem Involvement: 2 Distinct Subgroups with Different Prognosis <i>C.A.P.F. Alves, et al.</i> | PEDIATRICS |
|  | 1208 | Utility of Gadolinium-Based Contrast in Initial Evaluation of Seizures in Children Presenting Emergently <i>Denas Andrijauskis, et al.</i> | PEDIATRICS |
|    | 1212 | Dual-Layer Detector Head CT to Maintain Image Quality While Reducing the Radiation Dose in Pediatric Patients <i>Zhengwu Tan, et al.</i> | PEDIATRICS |
|  | 1219 | Arterial Spin-Labeling in the Assessment of Pediatric Nontraumatic Orbital Lesions <i>S. Neumane, et al.</i> | PEDIATRICS |
|  | 1224 | Cerebral Hemodynamic and Metabolic Abnormalities in Neonatal Hypocalcemia: Findings from Advanced MRI <i>Ying Qi, et al.</i> | PEDIATRICS |
|  | 1231 | Neuroimaging Findings in Axenfeld-Rieger Syndrome: A Case Series <i>Samuel White, et al.</i> | PEDIATRICS |
|  | 1236 | Venous Sinus Stenosis with Prominent Emissary Veins: A New Common Cranial MRI Finding of Mucopolysaccharidosis I <i>Shiwei Huang, et al.</i> | PEDIATRICS |

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

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



The study by Ibrahimy, et al, demonstrated changes in the MR imaging morphometrics and quality of life in longitudinal follow-up of patients with Chiari malformation after craniocervical decompression.



Indicates Editor's Choices selection



Indicates Fellows' Journal Club selection



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



Indicates article with supplemental online data



Indicates article with supplemental online video



Evidence-Based Medicine Level 1



Evidence-Based Medicine Level 2

Peer Learning in Neuroradiology: Not as Easy as It Sounds

Peer Learning (PL) is an engaging activity in which practicing radiologists come together to review cases from which they can learn jointly. The major impetus for PL lies in the overarching goal of improving diagnosis in radiology through a team-based culture of viewing mistakes as an opportunity to learn.¹ In its book *Improving Diagnosis in Health Care*, the Institute of Medicine (IOM) found that most people will be affected by at least 1 diagnostic error in health care.²

As a result of the IOM book *Improving Diagnosis in Health Care*, several external drivers are in place to change the practices of health care providers toward improved diagnosis. Radiologists realized that PL, but not random score-based peer review, best meets the IOM goals of establishing effective teamwork, educating practitioners in the diagnostic process, learning from mistakes, creating a culture that improves diagnostic performance, and establishing a reporting mechanism for discrepancies.¹ Following a 2020 Radiology Peer Learning Summit, the American College of Radiology (ACR) developed a new accreditation pathway that replaces score-based random peer review with PL.^{3,4} The American Board of Radiology (ABR) added PL as an alternative participatory activity for meeting Maintenance of Certification (MOC) Part 4 criteria.⁵ The Joint Commission (TJC) serves as another external driver of improved practitioner performance, eg, peer review, through Ongoing Professional Practice Evaluation (OPPE) requirements.⁶

Assessments of agreement among university neuroradiologists showed up to 12.4% disagreement rates,⁷ which are much higher than the reported 2.9% for score-based random peer review.⁸ Among errors in neuroradiology are discrepancies regarding vascular, neoplastic, and congenital disorders, as well as artifacts.⁷ Moreover, neuroradiology errors can also arise from test-selection errors, protocolling errors, technical errors, and failure to communicate results in a timely fashion. Education can decrease errors, given that neuroradiologists with high participation rates in the Tumor Board have lower diagnostic error rates.⁹ Having awareness of “blind spots,” for example with complex head and neck anatomy and pathology, may decrease interpretive errors and could conceivably be improved with PL.^{10,11} In fact, there is currently some evidence that PL creates learning opportunities,^{12,13} but there is still a lack of adoption of PL programs nationwide¹⁴ and a lack of scientific evidence demonstrating its effectiveness.

This Perspectives comes from members of the American Society of Neuroradiology (ASNR) Quality, Safety and Value Committee. Here, we describe several challenges faced by neuroradiologists who are interested in serving as PL champions. Among these challenges is an inability to recruit volunteer champions to drive PL programs, lack of resources for running a PL program, unknown effects on required reporting to TJC, and lack of evidence favoring PL over score-based random peer review.

Challenge 1: Who Wants to Be a PL Champion?

The barriers related to appointing PL champions may be related to implicit expectations regarding clinical rather than noninterpretive performance and to the scope of this role, depending on the existing culture within the neuroradiology practice and the resources available to support a PL program.

The current practice environment in neuroradiology is characterized by rising clinical volumes, tight finances, and the unfolding Great Resignation. As a result, neuroradiologists may cut back on noninterpretive duties.¹⁵ In a pediatric neuroradiology example, it was stated that serving as a PL champion requires, at a minimum, several hours of work in preparation for PL meetings and may require additional time for managing discrepancies as well as for external reporting such as generating and submitting data for the ACR accreditation program, for TJC, or for receiving Continuing Medical Education (CME) credits.¹³ There are currently no established physician roles that would provide PL champions with protected time for these tasks.

PL champions face additional challenges. Some radiologists believe that participating in interesting case conferences is the same as PL. While this belief is literally true, the term PL is to be viewed in the context of meeting goals from the IOM *Improving Diagnosis in Health Care*, which includes several layers of accountability, foremost a process for handling discrepancies.¹ There needs to be a clear process for reporting discrepancies, for consistently notifying the original interpreting radiologist of the discrepancy, and for ensuring optimal patient care. To meet TJC and ACR requirements, PL must include cases with discrepancies, but additionally, the inclusion of great catches, interesting cases, and so forth can foster shared learning. PL champions may need to drive this culture shift, emphasizing the importance of reporting discrepancies for learning purposes.

The ACR acknowledges the importance of separating the performance review of radiologists from learning and professional growth by creating the ACR Accreditation Pathway for PL,⁴ which substitutes agreement/disagreement ratings with measures of participation in a PL program. For neuroradiologists who are used to the randomized score-based peer review, the culture needs to shift away from perceiving discrepancy reporting as punishment or as a performance-assessment tool and toward sharing learning opportunities so the group can learn and grow. This shift in measuring performance is crucial because traditional score-based peer review has been used as a punitive measure in the past.^{14,16} Achieving such a culture shift could be an impossible lift for neuroradiologists. Leadership support is critical to the success of any PL program.

Another culture shift revolves around how discrepant opinions are handled. Absolute certainty in medicine is rare to come by, and often there are differences of opinion regarding a diagnosis in radiology. In the randomized score-based peer review system, a voting system would be used to decide which image interpretation is more “correct.” This culture may be founded in the traditional “learn-what” approach, such as reading an article or taking an online course that may provide a sense of certainty

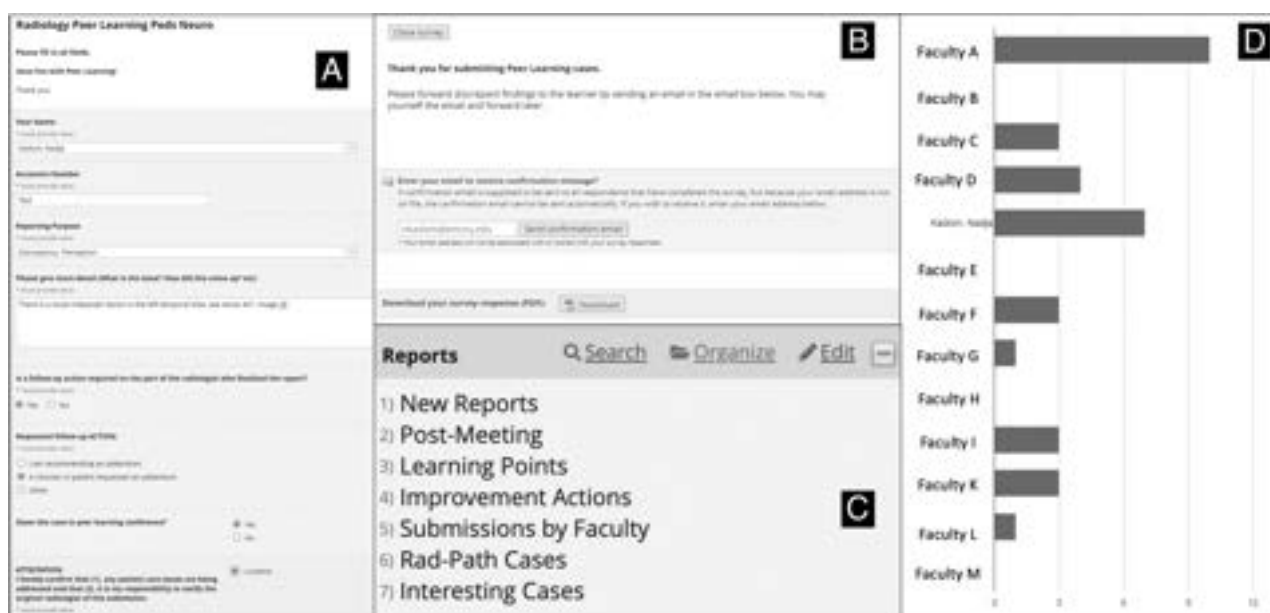


FIG 1. REDCap tool for PL. **A**, The submitter can indicate his or her name to receive credit against the monthly case submission requirement per the ACR Accreditation Pathway for PL. The submitter selects the reason for case submission, which includes discrepancies as well as interesting cases, good catches, and more. We use the PACS accession number as the case identifier. Any additional required actions can be entered, and the submitter attests to being responsible for ensuring optimal patient care. **B**, After submitting the content in the survey, a PDF is created that contains all survey input, except the name of the case submitter. The submitter can input the original reader's email to quickly share the feedback. **C**, The REDCap tool allows the creation of reports that easily summarize information such as learning and improvement actions resulting from the PL program, which can be used for annual reporting on the ACR Accreditation Pathway for PL. **D**, We also have an administrative assistant monitor monthly case submissions and send an email with current submissions to every participating radiologist midway through the monthly reporting period.

of a diagnosis. Instead, PL emphasizes the idea of “learn-how,” by sharing knowledge, offering suggestions, and discussing alternative diagnostic approaches in the setting of discrepant opinions.¹⁷

Thankfully, it is not necessary to design a PL program de novo. The ACR's PL checklist and sample policies provide a great starting point for building a PL program that is founded on the IOM goals for *Improving Diagnosis in Health Care*.^{1,2} The main pillars of PL programs include a mechanism for managing discrepancies, providing a safe learning environment, having a clear separation of learning from performance evaluation, and anonymization. There are ample opportunities for the identification of discrepancies in neuroradiology, for example during any comparison with a prior study, from secondary review for the Neuro-Oncology Tumor Board and other neuroimaging conferences and reading room secondary opinion consultations, during teaching sessions, and from clinical error reporting.

The PL champion is also tasked with creating a safe learning environment. Groups with higher psychological safety have a “shared belief that the team is safe for interpersonal risk-taking.”¹⁸ As a result, members of such groups practice open communication, are not afraid to voice concerns or ask questions, and seek feedback without fear of being judged. Achieving such a culture requires deliberate effort to flatten authority gradients and eliminate any language that implies blaming, shaming, or judging. To create safe learning environments the PL champions will set clear ground rules, serve as role models, foster nonjudgmental behavior by demonstrating openness to different perspectives, and actively discourage any dismissiveness/hostility.¹⁹

Challenge 2: Are There Resources for Running a PL Program?

Besides a PL champion driving and executing a PL program, there are also resources required for external regulatory reporting, such as the ACR Accreditation Pathway for PL or TJC, and possibly reporting for claiming CME credits. Among required resources, besides the PL champion's time and enthusiasm, are software tools and support staff time.

There are a few commercial tools that help manage various aspects of PL, but there is not currently a comprehensive commercial tool that manages the entire process, starting when a radiologist submits a case and including management of discrepancies, ensuring optimal clinical care, aiding PL champions in preparing for the PL conference, documenting PL performance targets, running a conference with anonymized PACS cases, capturing learning and improvement initiatives, and generating an annual report. Many commercial tools facilitate case submissions and case rating/classification systems (discrepancy, great catch, and so forth) but lack the ability to extract data for monthly and annual tracking of radiologist performance targets, such as the number of monthly cases submitted and participation in PL conferences. Meeting attendance can be tracked separately, and this tracking can be facilitated by choosing virtual meeting platforms that automatically generate attendance forms at the conclusion of the meeting.

At a coauthor's (N.K.) institution in a pediatric neuroradiology program, the REDCap (Research Electronic Data Capture; projectredcap.org) research tool is used to drive a large portion of

Sample approach to defining quantitative data for OPPE use that could replace score-based random review data^a

Competency	Competency Definitions, Radiology Metrics, and Rationale
Patient care	<ul style="list-style-type: none"> • Definition: Able to provide patient care that is compassionate, appropriate, and effective for the treatment of health problems and the promotion of health • Metric: Report TAT • Rationale: Report TAT is an important process metric in radiology; providing timely result reporting contributes to timely patient management, may avoid complications related to imaging findings, and supports the patient's shared decision-making process
Medical knowledge	<ul style="list-style-type: none"> • Definition: Must demonstrate knowledge about established and evolving biomedical, clinical, and cognate (eg, epidemiologic, social behavioral) sciences and the application of this knowledge to patient care • Metric: Maintenance of state licensure and radiology board certification • Rationale: Both licensure and maintenance of board certification require ongoing education and skills assessments (CME, MOC)
Practice-based learning and improvement	<ul style="list-style-type: none"> • Definition: Able to investigate and evaluate patient care practices, appraise and assimilate scientific evidence, and improve patient care practices • Metric: PL conferences • Rationale: It has been demonstrated that participation in a PL program has measurable educational value, which can conceivably result in improvement of a radiologist's practice
Interpersonal and communication skills	<ul style="list-style-type: none"> • Definition: Able to demonstrate interpersonal and communication skills that result in effective information exchange and teaming with patients, patients' families, and professional associates • Metric: PL feedback • Rationale: In providing feedback regarding discrepancies that require a follow-up action, ie, an addendum, radiologists can demonstrate their interpersonal and communication skills to ensure effective patient care
Professionalism	<ul style="list-style-type: none"> • Definition: Must demonstrate a commitment to carrying out professional responsibilities, adherence to ethical principles, and sensitivity to a diverse patient population • Metric: PL meeting targets • Rationale: Program participation and meeting targets demonstrate professional commitment by radiology practitioners

Note:—ACGME, indicates the Accreditation Council for Graduate Medical Education.

^a The first column lists the 6 ACGME core competencies for physicians; the second column provides first the ACGME definition for each competency, then suggests how either TAT or various peer learning metrics can meet this definition. This table could serve as a sample approach when entering discussions with local OPPE representatives and can be modified to reflect any metrics that the local radiology practice considers meaningful and feasible.

the PL process. Specifically, the tool serves as a case-reporting tool, it can be used to notify original readers of their reported cases and reporting reasons, it indicates whether any actions for clinical care need to be taken by the original reader, and it allows data-tracking and data summary required by the ACR Accreditation Pathway for PL (Fig 1). Having to create tools and processes and then implement them can be time-consuming and may require collaboration with other subject matter experts, which can delay their implementation.

Being able to present cases in an anonymized fashion can represent another challenge. Preserving anonymity in PL is important because it can positively impact learners' perceptions of the value of PL, can foster the provision of more critical peer feedback, and can lead to increased performance.²⁰ The extent of anonymity required may depend on the maturity level of the safety culture within a group of neuroradiologists but generally involves anonymity of the notification of a discrepancy as well as anonymous case presentation during PL meetings. Anonymity facilitates a nonpunitive atmosphere during review of cases among a group of attendings and trainees. The cases should be prepared by the PL champion to minimize the number of people who can identify cases and readers. Most interesting, to meet the ACR PL

Accreditation Pathway criteria, the identity of anyone submitting cases needs to be captured to meet performance targets, but the identity of the original reader whose report was flagged as a discrepancy is not required to be captured. Inclusion of interesting cases or great catches, however, can increase PL participation and transparency²¹ and represents an opportunity to celebrate individuals by name. Of note, many PACS systems do not completely anonymize patient identification, and PL champions need to be cautious when screensharing the entire PACS window. Certain virtual meeting applications allow sharing only a portion of the screen, which may be better suited to preserving anonymity. Another way to preserve anonymity is to create slide presentations, which can be very time-consuming.¹³

A key outcome of a PL program from the perspective of the ACR Accreditation Pathway for PL is the documentation of quality improvements that arose from the PL program. PL meetings and case discussions can lead to the discovery of process and system issues that can be addressed. Many issues may be addressed directly by the neuroradiologists in the PL group, such as changing CT and MR imaging protocols or reporting templates, but larger issues, such as a broken system for providing feedback from neuroradiologists to technologists, may require escalation

I. Citizenship Measures:
a. Total number of suspensions due to incomplete medical records: 0

II. Practitioner Peer Review:

Cases <i>Cases can fit into multiple categories</i>		Breakdown of Significant Dispositions	
3	Total Referrals	1	Letter of Information
1	Total Significant Dispositions	0	Trend
0	Behavior Concerns	0	Letter of Education
0	Reviewed at SPPEC*/MEC	0	Verbal Intervention
0	Other (PIP, Behavior Contract, etc.)	0	Administrative CI
0	From OPPE	0	Bundled

*SPPEC- System Professional Practice Evaluation Committee / MEC- Medical Executive Committee

~Contact Manager, Professional Practice Evaluation at [REDACTED] for further information~

III. Clinical Privileges Subject to focused evaluation for over 12 months:

Only applicable if checked

☒ I have reviewed the above Practitioner Peer Review information.

☐ I have reviewed the above information and have contacted [REDACTED] Manager of Professional Practice Evaluation to obtain further information.

Signature _____

Signature _____

FIG 2. A sample OPPE form allowing peers to evaluate their peers. This qualitative evaluation aligns with the 6 Accreditation Council for Graduate Medical Education (ACGME) core competencies and can serve to identify any practice concerns.

to a dedicated improvement team.²² It may be challenging to set up a process for handing off such projects to a dedicated quality team if the practice even has access to one.²³

Challenge 3: How Does Peer Learning Meet TJC Requirements?

Another barrier affecting the transition from random score-based peer review to PL relates to external reporting of radiologists' performance. Specifically, it is still unknown whether TJC will accept metrics derived from PL to replace the widely accepted random score-based review performance evaluation metrics of agreement/disagreement rates between radiologists.

TJC requires health care entities to provide both qualitative and quantitative data for OPPE and has traditionally accepted data from random score-based peer review to meet this requirement in radiology. It is technically possible to maintain random score-based peer review for external reporting purposes while also participating in PL, but this practice could cause confusion and mistrust among practicing radiologists, which would counteract the basic principles of a safe PL environment.

It may be better to replace score-based random peer review data reporting for OPPE with a different set of performance data. For example, 1 coauthor (N.K.) is proposing the use of report turnaround times (TAT) in conjunction with PL metrics (the number of cases submitted per radiologist per month, PL meeting

participation, and so forth) for reporting only quantitative data for OPPE²¹ (Table). Additional qualitative data that TJC may require for OPPE could be collected through other pathways. For example, annual peer evaluations could be collected like those commonly used in the credentialing process (Fig 2). In addition, data from reporting systems for issues of physician practices could be used to reflect a qualitative assessment of radiologists' performance (Fig 3).

Neuroradiologists who are interested in discontinuing random score-based peer review will have to consider external reporting requirements and work with representatives from those agencies to ensure that any new metrics meet existing requirements.

Challenge 4: What Is the Scientific Evidence Favoring PL?

It may be difficult for PL champions to convince leadership of abandoning traditional peer review in favor of PL. There is some evidence in the scientific literature that PL is a better approach to improving diagnosis than randomized score-based peer review but not necessarily that it is a performance-evaluation tool for radiologists. There is, however, ample evidence that score-based peer review is a flawed performance-evaluation tool²⁴ that failed to demonstrate learning²⁵ and failed to engage radiologists.^{14,26}

If one is considering addendum rates as a surrogate marker of improved patient care, then PL by far exceeds the effects of score-

Ongoing Professional Practice Evaluation (OPPE) - FORM A

All Allied Health Professionals & Advanced Practice Providers

Practitioner's Name/Title: _____ Employee ID: _____

Evaluator's Name/Title: _____ Date: _____

Competency Measures 1= Does not Perform 2= Performs Below Expectations 3= Meets Standard 4= Intermittently Exceeds Expectations 5= Consistently Exceeds Expectations										
	Competencies					Method				
	1	2	3	4	5	DO	PR	A	S	
Patient Care • Provides patient care that is compassionate, appropriate and effective for the promotion of health, prevention of illness, treatment of disease and care at the end of life										
Medical/Clinical Knowledge • Demonstrates knowledge of established and evolving biomedical, clinical and social sciences in the application of knowledge to patient care and the education of others										
Systems Based Practice • Demonstrates both an understanding of the contexts and systems in which health care is provided, and the ability to apply this knowledge to improve and optimize health care within the network.										
Practice Based Learning and Improvement • Able to use scientific evidence and methods to investigate, evaluate and improve patient care practices										
Interpersonal Communications • Demonstrates interpersonal and communication skills that establish and maintain professional relationships with patients, families, coworkers, and other members of the healthcare team.										
Professionalism • Demonstrates behaviors that reflect a commitment to continuous professional development, ethical practice, an understanding and sensitivity to diversity and a responsible attitude toward patients, profession and society.										

Self-report any complicated cases: _____

Collaborating/Supervising Physician: Please Check ☐ below and sign.

☐ This provider meets expectations of performance

☐ This provider meets expectations of performance based on the privileges granted with an opportunity for improvement in the area of _____
 A summary of the performance improvement plan is described below.

☐ This provider meets most expectations of performance based on the privileges granted except in the area of _____
 resulting in an FPPE. See FPPE for further details.

Summary of comments for the Competency Measures (May add additional comments here. Must provide a plan of action for competency rates below a 2)

Signature of Collaborating/Supervising Physician _____ Date _____

FIG 3. A sample report describing Focused Professional Practice Evaluation (FPPE) events to division directors, which represents and qualitative assessment that OPPE can use for TJC reporting. Division directors will not know the nature of the events that have been investigated, but they can still easily glean from this type of reporting across time whether a radiologist's practice raises concern in terms of a higher-than-usual number of relevant issues and dispositions, such as behavior concerns, contract violations, and verbal or administrative interventions.

based random peer review,²⁷ but data directly linking peer learning to improved patient outcomes are still missing.

There is some evidence that PL may lead to greater radiologist engagement.^{12,13} Physician burnout poses an increased risk of patient safety incidents as well as poor quality of care and low patient satisfaction.²⁸ A recent report stated that reported burnout among US neuroradiologists ranged from 49% to 79%.²⁹ There remains an opportunity to generate additional scientific evidence linking PL to radiologist engagement metrics and linking improved engagement to improved patient outcomes.

Overall, the field of PL offers an opportunity for neuroradiologists to apply a scholarly angle. For example, we need evidence that PL leads to a neuroradiologist's improved ability to reliably make an accurate diagnosis, that PL improves the cohesiveness of neuroradiology teams, and that PL could reduce burnout. There is a traditional view that high clinical volumes lead to lower academic output in neuroradiology, as measured by peer-reviewed articles, presentations, and abstracts.³⁰ However, this simplistic linkage, which disregards factors like seniority and work schedules, has been criticized by other neuroradiologists³¹ and should not deter neuroradiologists from engaging in roles that do not contribute to clinical output.

The field of PL in radiology is still evolving. Neuroradiologists have an opportunity to become leaders in this field. Meeting the challenges presented in this article can result in professional and personal growth, improved job satisfaction, and reduced feelings of burnout. These are important possible gains to consider when weighing the commitment required to fill a PL champion role.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

- Larson DB, Donnelly LF, Podberesky DJ, et al. **Peer feedback, learning, and improvement: answering the call of the Institute of Medicine report on diagnostic error.** *Radiology* 2017;283:231–41 CrossRef Medline
- Balogh EP, Miller BT, Ball JR, et al. **Committee on diagnostic error in health care.** In: Balogh EP, Miller BT, Ball JR, et al. *Improving Diagnosis in Health Care*. National Academies Press; December 29, 2015
- Larson DB, Broder JC, Bhargavan-Chatfield M, et al. **Transitioning from peer review to peer learning: report of the 2020 Peer Learning Summit.** *J Am Coll Radiol* 2020;17:1499–508 CrossRef Medline
- American College of Radiology. **Peer Learning Resources.** www.acr.org/Practice-Management-Quality-Informatics/Peer-Learning-Resources. Accessed May 11, 2023
- American Board of Radiology. **Participatory Activities.** March 17, 2022. <https://www.theabr.org/diagnostic-radiology/maintenance-of-certification/improvement-medical-practice/participatory-activities>. Accessed May 18, 2023
- The Joint Commission. **What are the key elements needed to meet the Ongoing Professional Practice Evaluation (OPPE) requirements?** <https://www.jointcommission.org/standards/standard-faqs/hospital-and-hospital-clinics/medical-staff-ms/000001500>. Accessed May 21, 2023
- 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
- 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 Medline
- Ivanovic V, Assadsangabi R, Haccin-Bey L, et al. **Neuroradiology diagnostic errors at a tertiary academic centre: effect of participation in tumour boards and physician experience.** *Clin Radiol* 2022;77:607–12 CrossRef Medline
- Assadsangabi R, Maralani P, Chen AF, et al. **Common blind spots and interpretive errors of neck imaging.** *Clin Imaging* 2022;82:29–37 CrossRef Medline
- Vong S, Chang J, Assadsangabi R, et al. **Analysis of perceptual errors in skull-base pathology.** *Neuroradiol J* 2022 Jun 18 [Epub ahead of print] CrossRef Medline
- Sharpe RE Jr, Huffman RI, Congdon RG, et al. **Implementation of a peer learning program replacing score-based peer review in a multispecialty integrated practice.** *AJR Am J Roentgenol* 2018;211:949–56 CrossRef Medline
- Kadom N, Reddy KM, Khanna G, et al. **Peer learning program metrics: a pediatric neuroradiology example.** *AJNR Am J Neuroradiol* 2022;43:1680–84 CrossRef Medline
- Lee CS, Neumann C, Jha P, et al. **Current status and future wish list of peer review: A National Questionnaire of U.S. Radiologists.** *AJR Am J Roentgenol* 2020;214:493–97 CrossRef Medline
- Chen JY, Vedantham S, Lexa FJ. **Burnout and work-work imbalance in radiology—wicked problems on a global scale: a baseline pre-COVID-19 survey of US neuroradiologists compared to international radiologists and adjacent staff.** *Eur J Radiol* 2022;155:110153 CrossRef Medline
- Ngo AV, Stanescu AL, Swenson DW, et al. **Practical considerations when implementing peer learning conferences.** *Pediatr Radiol* 2019;49:526–30 CrossRef Medline
- Edmondson AC. *The Fearless Organization*. John Wiley & Sons 2018;175–77
- Edmondson A. **Psychological safety and learning behavior in work teams.** *Adm Sci Q* 1999;44:350–83 CrossRef
- Holley LC, Steiner S. **Safe space: student perspectives on classroom environment.** *J Soc Work Educ* 2005;41:49–64 CrossRef
- Panadero E, Alqassab M. **An empirical review of anonymity effects in peer assessment, peer feedback, peer review, peer evaluation and peer grading.** *Assessment & Evaluation in Higher Education* 2019;44:1253–78 CrossRef
- Khader A, Ali S, Wald C, et al. **Abdominal peer learning: advantages and lessons learned.** *Abdom Radiol (NY)* 2023;48:1526–35 CrossRef Medline
- Donnelly LF, Larson DB, Heller RE III, et al. **Practical suggestions on how to move from peer review to peer learning.** *AJR Am J Roentgenol* 2018;210:578–82 CrossRef Medline
- Broder JC, Scheirey CD, Wald C. **Step by step: a structured approach for proposing, developing and implementing a radiology peer learning program.** *Curr Probl Diagn Radiol* 2021;50:457–60 CrossRef Medline
- Bender LC, Linnau KF, Meier EN, et al. **Interrater agreement in the evaluation of discrepant imaging findings with the Radpeer system.** *AJR Am J Roentgenol* 2012;199:1320–27 CrossRef Medline
- Eisenberg RL, Cunningham ML, Siewert B, et al. **Survey of faculty perceptions regarding a peer review system.** *J Am Coll Radiol* 2014;11:397–401 CrossRef Medline
- Chaudhry H, Del Gaizo AJ, Frigini LA, et al. **Forty-one million RADPEER reviews later: what we have learned and are still learning.** *J Am Coll Radiol* 2020;17:779–85 CrossRef Medline
- Trinh TW, Boland GW, Khorasani R. **Improving radiology peer learning: comparing a novel electronic peer learning tool and a traditional score-based peer review system.** *AJR Am J Roentgenol* 2019;212:135–41 CrossRef Medline
- Panagioti M, Geraghty K, Johnson J, et al. **Association between physician burnout and patient safety, professionalism, and patient satisfaction: a systematic review and meta-analysis.** *JAMA Intern Med* 2018;178:1317–31 (Retraction published in *JAMA Intern Med* 2020;180:931) CrossRef Medline
- Weissman IA, Van Geertruyden P, Prabhakar AM, et al. **Practice resources to address radiologist burnout.** *J Am Coll Radiol* 2023;20:494–99 CrossRef Medline
- Eschelman DJ, Sullivan KL, Parker L, et al. **The relationship of clinical and academic productivity in a university hospital radiology department.** *AJR Am J Roentgenol* 2000;174:27–31 CrossRef Medline

31. Yousem DM. **Academic and clinical productivity: relative value units do not tell the whole story.** *AJR Am J Roentgenol* 2001;176:1598–600
CrossRef Medline

● **K. Mani**

University Radiology Group
Rutgers University School of Medicine
Newark, New Jersey

● **K. Shah**

MD Anderson Cancer Center
Houston, Texas

● **N. Kadom**

Emory University School of Medicine
Children's Healthcare of Atlanta
Atlanta, Georgia

● **D. Seidenwurm**

Sutter Health
Sacramento, California

● **A.J. Nemeth**

Northwestern University, Feinberg School of Medicine
Northwestern Memorial Hospital
Chicago, Illinois

Update from the 5th Edition of the WHO Classification of Nasal, Paranasal, and Skull Base Tumors: Imaging Overview with Histopathologic and Genetic Correlation

A. Agarwal, A.A. Bhatt, G. Bathla, S. Kanekar, N. Soni, J. Murray, K. Vijay, P. Vibhute, and P.H. Rhyner



ABSTRACT

SUMMARY: Sinonasal and skull base tumors are a heterogeneous group of neoplasms with considerable histologic variation and overlapping imaging features. In 2022, the World Health Organization updated the head and neck tumor classification, further emphasizing the importance of molecular data and genetic alterations in sinonasal neoplasms. The changes include the addition of new entities and discussion of emerging entities, as well as changes to the taxonomy and characterization of tumors. The new classification focuses on entities that develop in these sites either exclusively (eg, olfactory neuroblastoma) or most frequently. Another change includes reduction in the number of categories by creating separate category-specific chapters for soft-tissue, hematolymphoid, and neuroectodermal lesions. In this review, we briefly discuss the various categories in the new classification with a more detailed description of the 2 new entities (SWI/SNF-related complex-deficient sinonasal carcinomas and human papillomavirus-related multiphenotypic sinonasal carcinoma). We also highlight the emerging entities including *IDH*-mutant sinonasal malignancies and *DEK-AFF2* carcinoma, presently classified as sinonasal undifferentiated carcinoma and nonkeratinizing squamous cell carcinoma, respectively.

ABBREVIATIONS: HPV = human papillomavirus; *NUT* = nuclear protein in testis; REAH = respiratory epithelial adenomatoid hamartoma; SCCa = squamous cell carcinoma; *SMARCB1* = SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily B member 1; SNT = sinonasal tract; SNUC = sinonasal undifferentiated carcinomas; SWI/SNF = SWI/SNF-related complex-deficient; WHO = World Health Organization

The 5th edition of the World Health Organization (WHO) classification of head and neck tumors was released online in 2022 with a focus on distinctive molecular and genetic characteristics of tumors including in the sinonasal tract (SNT).¹ The increasing incorporation of molecular and genetic information reflects our continuously evolving understanding of the underlying biologic underpinnings and is broadly true for most updated WHO tumor classifications in other specialties. Apart from increasing the diagnostic accuracy, these genetic alterations also have prognostic and predictive value, allowing a more personalized approach. The 5th edition of the WHO Classification of Head and Neck Tumors has several changes from the 4th edition (2017), which consider the multidimensional nature of tumor

classification with a better consensus definition and provide greater insight into the pathogenesis of these tumors.^{1,2}

The SNT (including the nasal cavity, paranasal sinuses, and the skull base) is an anatomic region that has high morphologic, phenotypic, and genotypic diversity of neoplasms with considerable overlap of histologic features. The new classification focusses on either entities that develop in these sites exclusively (eg, olfactory neuroblastoma) or those with a preponderance of disease in this location. Another major change was exclusion of tumor subcategories that do not occur exclusively or predominantly in the SNT region, to avoid redundancy and duplication. These were moved to their own dedicated chapters, such as soft-tissue tumors, melanocytic tumors, neuroendocrine neoplasms (such as paraganglioma), and so forth, with few exceptions.³ Adamantinomatous craniopharyngioma is the only entity that has been moved to the SNT (from the nasopharynx in the 4th edition), because the SNT is the exclusive ectopic location for these tumors. A new chapter with genetic tumor syndromes was added with 15 entities with predominant head and neck manifestations. The Table outlines the changes to the major categories of SNTs. The new classification allows a more logical and stratified approach, with successive entities showing progression from benign to malignant and higher-grade tumors. There is a new category for mesenchymal tumors of the SNT and a category for “other” sinonasal tumors that includes

Received May 26, 2023; accepted after revision June 22.

From the Departments of Radiology (A.A., J.M., P.V., P.H.R.), and Radiology & Otolaryngology (A.A.B.), Mayo Clinic, Jacksonville, Florida; Department of Radiology (G.B.), Mayo Clinic, Rochester, Minnesota; Penn State University Health System (S.K.), Hershey, Pennsylvania; Department of Radiology (N.S.), University of Rochester Medical Center, Rochester, New York; and Department of Radiology (K.V.), University of Texas Southwestern Medical Center, Dallas, Texas.

Please address correspondence to Amit Agarwal, MD, Department of Radiology, 4500 San Pablo Rd, Mayo Clinic, Jacksonville, FL 32224; e-mail: amitmamc@gmail.com; @AmitAgarwalMD; @GBathlaMD

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

Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A7960>

Summary of changes including deletion, addition, and reclassification of SNT tumors in the 5th edition of the WHO Classification 2022

WHO Classification 4th Edition 2017 (Chapter 1)		WHO Classification 5th Edition 2022 (Chapter 2)
Respiratory epithelial lesions REAH Seromucinous hamartoma (new entity in 4th edition)		Hamartomas (new category) REAH Seromucinous hamartoma Nasal chondromesenchymal hamartoma (reclassification from “other” category)
Sinonasal papillomas	Subcategory under respiratory epithelial lesions in 5th edition (2022)	Respiratory epithelial lesions (includes papillomas, carcinomas, adenocarcinomas with removal of hamartomatous lesions) Sinonasal papillomas
Carcinomas Keratinizing SCCa Nonkeratinizing SCCa Spindle cell (sarcomatoid) SCCa Lymphoepithelial carcinoma Sinonasal undifferentiated carcinoma <i>NUT</i> carcinoma (new entity in 4th edition) Neuroendocrine carcinoma Adenocarcinoma Intestinal-type adenocarcinoma Non–intestinal-type adenocarcinoma	Subcategory under respiratory epithelial lesions in 5th edition (2022) Adenocarcinomas separated as a new subcategory	Carcinomas Keratinizing SCCa Nonkeratinizing SCCa (includes emerging entity <i>DEK-AFF2</i> carcinoma) <i>NUT</i> carcinoma SWI/SNF complex-deficient sinonasal carcinoma (new entity) Sinonasal lymphoepithelial carcinoma Sinonasal undifferentiated carcinoma (includes emerging entity <i>IDH</i> -mutated carcinoma) Teratocarcinosarcoma HPV-associated multiphenotypic sinonasal carcinoma (new entity) Adenocarcinomas (new subcategory, was previously under carcinomas) Intestinal-type sinonasal adenocarcinoma Non-intestinal-type sinonasal adenocarcinoma
Malignant soft-tissue tumors Borderline/low-grade malignant soft-tissue tumors Benign soft-tissue tumors Neuroectodermal/melanocytic tumors Hematolymphoid tumors	Categories removed with dedicated chapters on soft-tissue tumors in the new edition	Mesenchymal tumors of the SNT (new category with few retained entities from soft-tissue tumors, occurring exclusively or primarily in the SNT) SNT angiofibroma Sinonasal glomangiopericytoma Biphenotypic sinonasal sarcoma Chordoma
Other tumors Meningioma Sinonasal ameloblastoma Chondromesenchymal hamartoma (moved to hamartomas)		Other sinonasal tumors Sinonasal ameloblastoma Adamantinomatous craniopharyngioma (reclassified from chapter on “nasopharynx”) Meningioma of the SNT, ear Olfactory neuroblastoma

olfactory neuroblastoma.¹⁻³ Imaging characterization of these new entities is still at an early stage with significant overlap in the radiographic presentation. Nevertheless, it is important for the radiologist to be aware of the clinical course of these tumors, including the epidemiologic features and how they respond to treatment. A great example would be the *DEK-AFF2* subtype of nonkeratinizing squamous cell carcinoma (SCCa), which classically occurs along the posterior part of middle turbinate, with excellent response to immunotherapy.

Molecular and Genetic Markers

Histopathologic analysis remains a vital part of the pathologic work-up of SNTs, despite the challenging morphologic overlap. Molecular testing including immunohistochemistry has now become routine for pathologic evaluation of the SNTs in developed countries. Genetic profiling has caused an explosion in the subclassification of sinonasal malignancies, particularly focused

on subtyping for improved prognostication and treatment, though profiling is still lagging for head and neck tumors, compared with other systems like CNS tumors. Although genetic panels for head and neck tumors are still at an early stage, aberrations are now widely reported across different tumor types with varying degrees of sensitivity. A case in point would be the identification of *IDH 1/2* mutation in 50–80% of tumors currently classified as sinonasal undifferentiated carcinoma (SNUC), and a smaller proportion of cases classified as large cell neuroendocrine carcinoma or olfactory neuroblastoma. *IDH 1/2* mutation in these tumors is associated with less aggressive clinical behavior and is now noted to be a distinct emerging entity (Fig 1).^{4,5} Similarly, identification of the *DICER* mutation in nasal chondromesenchymal hamartoma would signify an underlying tumor-predisposition disorder with recognition of several seemingly unrelated neoplasms, including pulmonary blastoma, multinodular goiter, and thyroid carcinomas

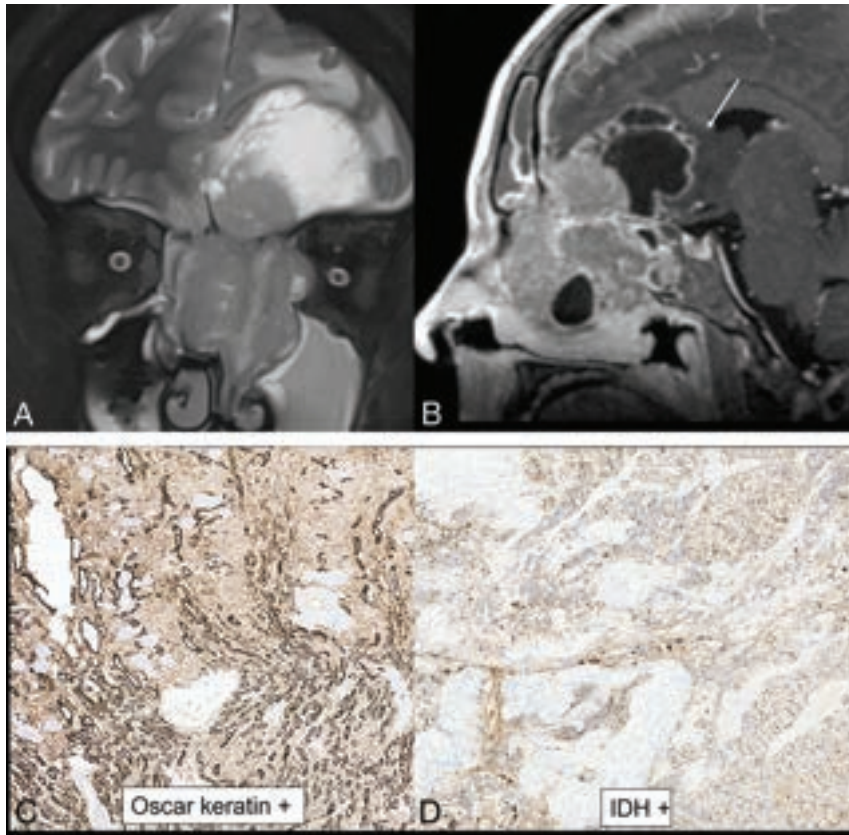


FIG 1. *IDH2*-mutated SNUC. A large, infiltrative sinonasal mass with homogeneous, mild T2 hyperintensity (A), avid contrast enhancement, and erosion of the skull base with intracranial extension (B). A large intratumoral cyst (B, arrow) with enhancement of the cyst walls. Radiographic differentials included SNUC and olfactory neuroblastoma. Histopathology shows sheet-like and nested growth of cells with hyperchromatic nuclei, with diffuse positivity on cytokeratin (Oscar) stain (C) distinguishing carcinomas (like SNUC) from nonepithelial malignancies. Immunohistochemistry for *IDH1/2* (D) shows strong and diffuse granular cytoplasmic staining, confirming an *IDH*-mutation.

(Fig 2).^{1,6} Newly recognized entities like SWItch/Sucrose Non-Fermentable (SWI/SNF) complex-deficient sinonasal carcinoma (provisionally included as SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily B member 1 [*SMARCB1*]-deficient sinonasal carcinoma in the 4th edition) also rely on a lack of *SMARCB1*/INI1 protein expression on immunohistochemistry. The 5th edition adds and expands on the genetic arrangements included in the definition of multiple entities, including that of sinonasal papillomas and hamartomas. The Online Supplemental Data outline the common molecular and genetic mutations in the newly recognized sinonasal tumors along with their pathogenetic pathway. The histologic grading of SNTs is an important independent predictor of tumor behavior. However, the propensity for grade transformation, progression, and recurrence is also dependent on the molecular subtypes, underlining their importance not just for the surgeons and oncologists but also for radiologists. It is almost certain that as the field evolves, additional neoplastic entities will be identified in the near future.

Hamartomas

Sinonasal hamartomas include respiratory epithelial adenomatoid hamartoma (REAH), seromucinous hamartoma, and

nasal chondromesenchymal hamartoma. REAH is a benign glandular neoplasm of the sinonasal cavities, which presents in isolation or secondary to allergy and inflammatory processes such as sinonasal polyps. This is the most common hamartoma of the SNT, seen in old adults with a strong male predilection, classically located along the posterior septum and olfactory cells. It has also been described in association with inverted papilloma and low-grade sinonasal adenocarcinoma, with reports suggesting that it might be the precursor for the latter. Imaging reveals exophytic polypoid homogeneous lesions with smooth expansion of the olfactory cleft without bony erosion, though identification can be challenging when associated with sinonasal polyposis. Seromucinous hamartomas are polypoid masses seen typically in the posterior nasal septum and nasopharynx, ranging from a few millimeters to a few centimeters. On imaging, these lesions cannot be differentiated from REAH, with both exhibiting a characteristic crescentic “half-moon” appearance on the sagittal images when situated along the olfactory cleft (Fig 3). The differential diagnosis for seromucinous hamartomas and REAH includes inflammatory polyp, encephalocele, olfactory neuroblastoma, and low-grade nonintestinal type adenocarcinoma.⁷⁻⁹

Chondromesenchymal hamartoma is the most distinctive lesion in this category, with a strong association with the *DICER1* (tumor-suppressor gene) mutation, usually presenting in infants. This is most common within the ethmoid sinuses and is frequently (25%) bilateral, with imaging revealing a complex, solid, and cystic heterogeneous appearance with calcification and bony erosions (Fig 2). *DICER1* syndrome is a highly pleiotropic tumor-predisposition entity that has been increasingly recognized during the past decade. It is no longer limited to pulmonary blastoma, with other associations including multinodular goiter, thyroid carcinoma, ovarian sex cord stromal tumors, and pituitary blastoma. The most common symptoms for all these hamartomatous lesions include nasal obstruction, rhinorrhea, and epistaxis.^{8,10,11}

Respiratory Epithelial Lesions

Sinonasal Papillomas. Sinonasal papillomas, also known as Schneiderian papillomas, are benign sinonasal neoplasms that arise from the Schneiderian epithelium of the nasal cavity and paranasal sinuses. They have 3 different histologic subtypes, including inverted papilloma (most common, 50%–80%), exophytic papilloma (20%–50%), and oncocytic papilloma (least

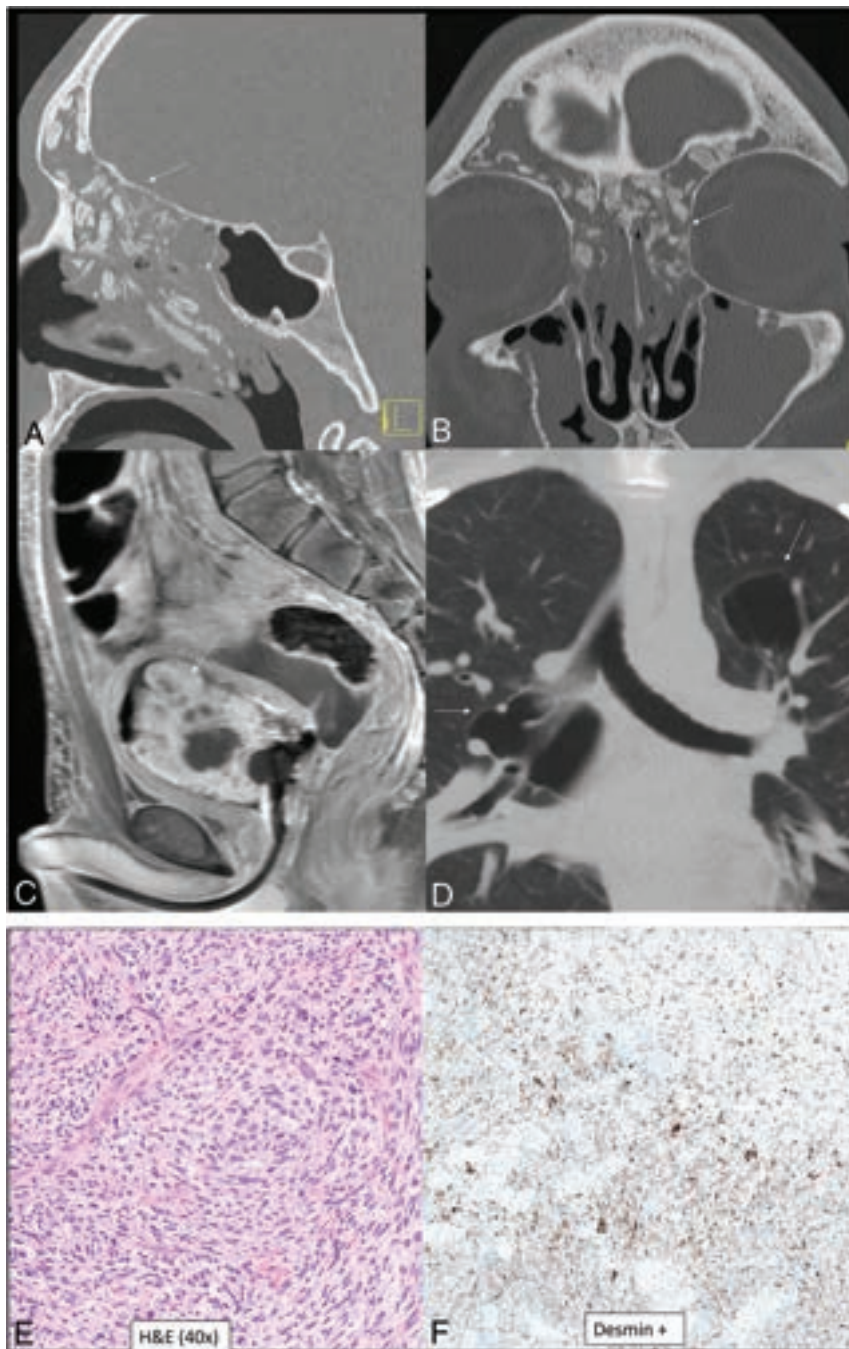


FIG 2. *DICER1*-mutant tumors. Nasal chondromesenchymal hamartomas with pathogenic germline variation in *DICER1* presenting with extensive soft-tissue “masses” in the frontal and ethmoid sinuses involving the anterior skull base with calcific/ossific bodies within the matrix (A and B, arrows). Embryonal bladder rhabdomyosarcoma in the same patient, which is also one of the hallmark tumors associated with a *DICER1* mutation. Bladder rhabdomyosarcoma is seen as a heterogeneous, solid, avidly-enhancing mass within the bladder lumen (C, arrow). Multiple thin-walled pulmonary cysts (D, arrow) with tiny septal nodules are indeterminate but may represent type 1r (regressed) pleuropulmonary blastomas in the setting of the *DICER1* mutation. The histopathology of a bladder lesion reveals embryonal rhabdomyosarcoma with diffuse anaplasia (E) and extensive cartilaginous differentiation and diffusely positive staining for desmin (muscle marker) (F).

common, 5%).¹² These account for 0.5%–4% of all nasal tumors and are usually seen in the fourth-to-fifth decades of life. There have been important updates to the etiology and pathogenesis of these lesions with somatic epidermal growth factor receptor

(*EGFR*) mutations reported in 90% of inverted papillomas and 80% of carcinomas developing from these lesions. Inverted papillomas generally occur along the lateral wall of the nasal cavity and the paranasal sinuses (most commonly the ethmoid or maxillary sinus). Imaging shows a soft-tissue mass centered along the middle meatus with frequent calcification and a classic convoluted cerebriform pattern, seen on both T2-weighted and contrast-enhanced T1-weighted images. These, however, cannot be distinguished from sinonasal carcinomas and are treated by en bloc resection, given the high risk of malignant transformation. Oncocytic papillomas also involve the lateral wall of the nasal cavity and appear similar to inverted sinonasal papilloma on imaging but have a lower risk of malignant transformation. They, however, may show T1-shortening (Fig 4) with multiple mucinous cysts along with lack of focal osteitis, all of which may help differentiate them from inverted papillomas.^{13,14} Exophytic sinonasal papilloma is a benign SNT epithelial neoplasm usually arising from the lower anterior nasal septum with a broad base. It is around 2 cm but lacks any specific imaging appearance. They are invariably unilateral and can involve the lateral nasal walls generally without involvement of the paranasal sinuses.^{15,16}

Carcinomas. Carcinomas of the SNT account for 3% of malignant head and neck neoplasms. Two new entities have been added to the category of sinonasal carcinomas and include human papillomavirus (HPV)-related multiphenotypic sinonasal carcinoma (provisionally included as HPV-related carcinoma with adenoid cystic-like features in the 4th edition) and SWI/SNF complex-deficient sinonasal carcinoma (provisionally included as *SMARCB1*-deficient sinonasal carcinoma in the 4th edition).¹ Neuroendocrine carcinoma has been removed from this section in the current edition and is discussed separately. Other carcinomas

have been retained from the prior edition and include SCCa (keratinizing and nonkeratinizing), *NUT* carcinoma, and lymphoepithelial and undifferentiated carcinoma. Although these lesions have different anatomic preferences, imaging findings are

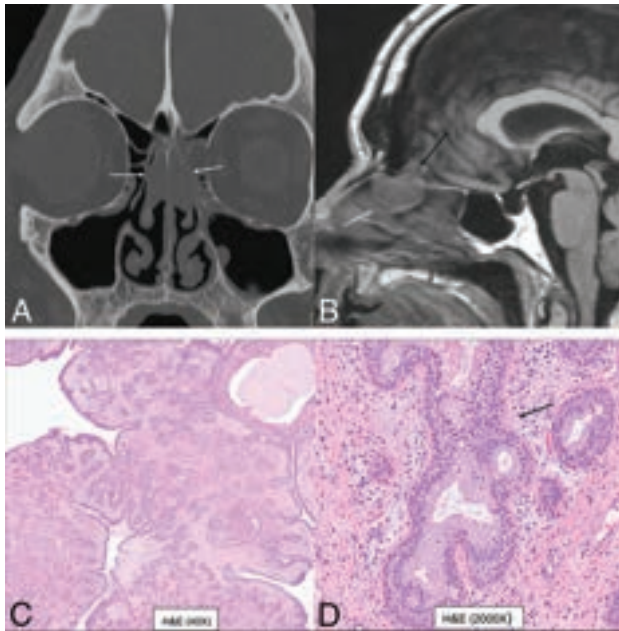


FIG 3. Bilateral REAH. Smooth polypoid lesions noted within the nasal cavity bilaterally (A and B, white arrows). The lesion expands the olfactory cleft (B, black arrow) with smooth remodeling, without erosive changes. This cannot be differentiated from seromucinous hamartoma with both exhibiting a crescentic (half-moon) shape on sagittal images (B). Histopathology revealed prominent glandular structures of varying sizes (C). Glands are lined with ciliated respiratory epithelium with bland nuclei and mucinous gland metaplasia (D, black arrow).

overall similar, with aggressive features, presenting as a poorly circumscribed soft-tissue mass with irregular margins often with bone destruction and frequent multifocal and metastatic disease.¹⁷ Keratinizing SCCa is rarely associated with HPV, may be associated with sinonasal papillomas, and is histologically identical to any other affected site. This is the most common malignancy of the nasal vestibule, with the maxillary antrum with the lateral nasal wall being the other common locations. Nonkeratinizing SCCa is a distinctive sinonasal tumor that most commonly arises in the nasal cavity or maxillary sinuses. It is characterized by minimal-to-no keratinization and is strongly positive for immunohistochemical markers such as p40 and CK5/6, which also help in differentiating them from morphologic mimickers like *NUT* carcinoma, SNUC, and neuroendocrine tumors.^{18,19}

NUT carcinoma was a new entity in the previous (4th) edition and is better defined in the current 5th edition.¹ It was previously called *NUT* midline carcinoma due to proclivity for midline sites. *NUT* carcinoma has been recognized in up to 18% of upper aerodigestive tract poorly differentiated carcinomas, with recognition aided by improved testing techniques and increased availability of commercial antibodies. The defining genetic feature of *NUT* carcinomas includes fusion of the *NUTM1* gene, most commonly to *BRD4*.²⁰ There are <100 cases described in the literature, limiting detailed description of imaging features. Limited series and published case reports have described aggressive imaging findings, like those in the mediastinum, with bony hyperostosis, internal mineralization, avid

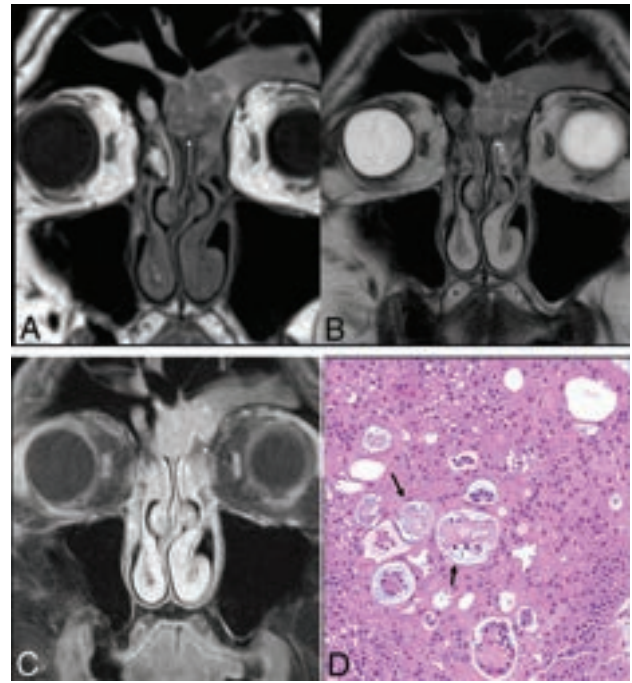


FIG 4. Sinonasal papilloma, oncocytic type, in the left frontal sinus in a 78-year-old man. Smooth polypoid lesions are noted within the nasal cavity with intrinsic T1 hyperintensity (A, arrow). The lesions obstruct the frontoethmoidal recess with trapped proteinaceous content in the frontal sinus. Lesions show a "cerebriform" pattern on the T2-weighted image (B, arrow) with avid enhancement (C, arrow). Histopathology reveals characteristic oncocytic cuboidal-to-columnar cells and intraepithelial microcysts with mucin and/or neutrophilic microabscesses (D, black arrows).

contrast enhancement, and high FDG uptake (Fig 5).^{8,21} Other tumors in this category include sinonasal lymphoepithelial carcinoma, teratocarcinosarcoma, and SNUC. However, with advancement in molecular markers and identification of newly defined entities (eg, SWI/SNF complex-deficient sinonasal carcinomas), the diagnosis of poorly differentiated or undifferentiated entities like SNUC is becoming less common.¹

SWI/SNF complex-deficient sinonasal carcinomas were provisionally included as *SMARCB1*-deficient sinonasal carcinoma in the 4th edition. These are poorly differentiated, highly-aggressive, and infiltrative epithelial malignancies defined by loss of 1 SWI/SNF complex subunit (most commonly *SMARCB1* or less frequently *SMARCA4*), usually presenting at an advanced stage.¹ These comprise 1%–3% of sinonasal carcinomas and 3%–20% of tumors diagnosed as SNUC. These tumors are seen in young and old adults (age range, 33–78 years) with no sex predilection. The *SMARCB1*-deficient carcinomas involve the paranasal sinuses (most commonly the ethmoid), whereas the *SMARCA4*-deficient carcinoma predominantly involves the nasal cavity. These are destructive lesions with common involvement of the skull base and orbits, frequent (50%) calcification, a "hair-on-end" pattern periosteal reaction, and increased uptake on PET. On MR imaging, these are variably hyperintense on T2WI and show avid enhancement with low ADC values. Given the recent identification of this tumor, detailed studies on the imaging pattern are limited. The differentiation is based on immunohistochemical markers with

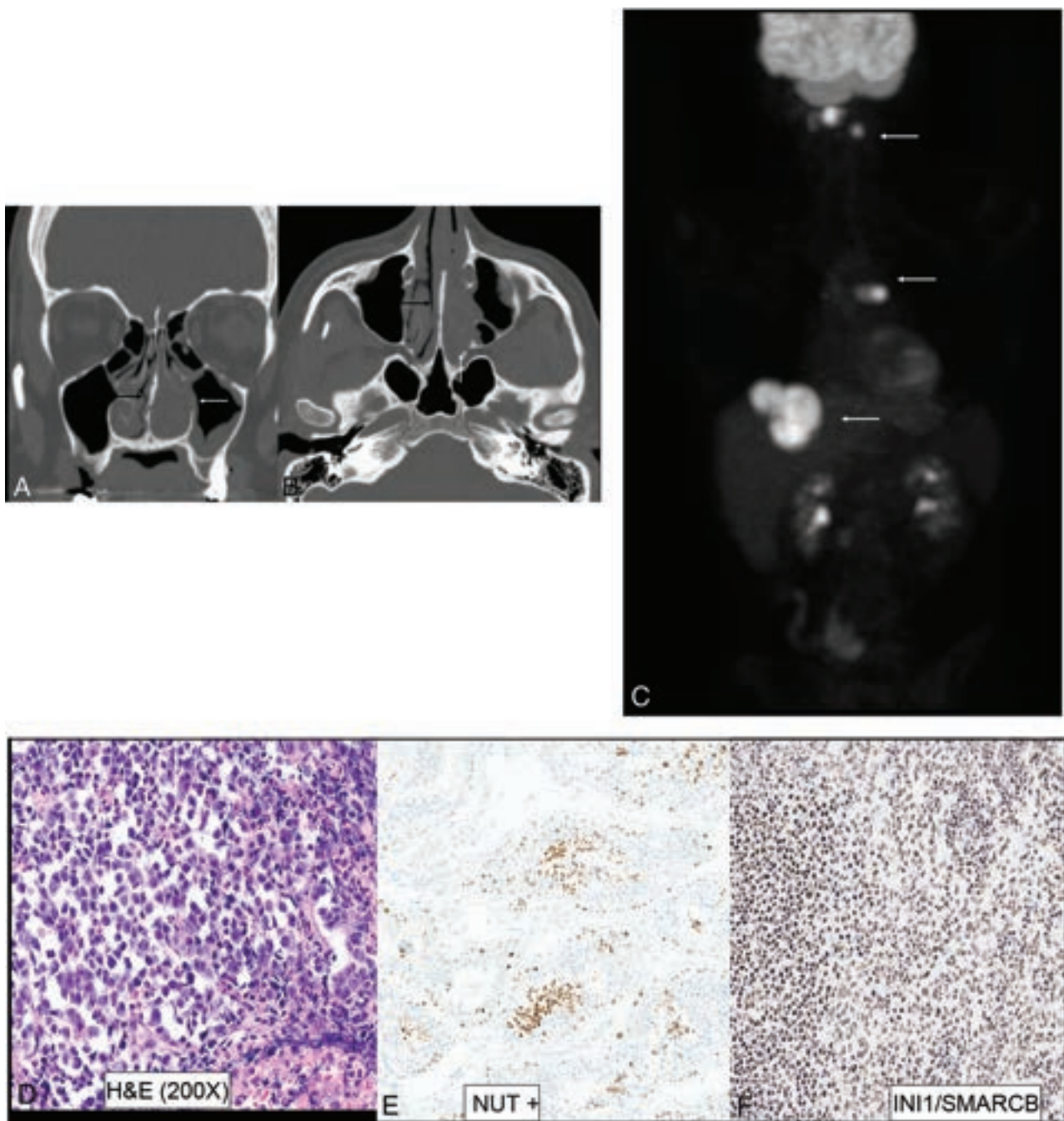


FIG 5. *NUT* sinonasal carcinoma in a 22-year-old man. Left nasal cavity mass (A and B, white arrows) with hyperostosis of the septum (black arrows). Metastatic disease with lesions in the lymph nodes, liver, and lung (C, white arrows). H&E stain (D) shows sheets and nests of high-grade tumor cells with zones of tumor necrosis. The tumor exhibits positive staining on *NUT* immunohistochemistry (E) with preserved INI1 staining (F) (ruling out *SMARCB1* deficiency). Genetic analysis revealed *NUT-BRD4* fusion, the defining feature of *NUT* carcinomas.

complete loss of expression of *SMARCB1* (INI1-stain negative) or *SMARCB4* (BRG1-stain negative) in the SWI/SNF complex-deficient sinonasal carcinomas (Fig 6).^{21,22} The common differential for an aggressive sinonasal tumor along the anterior skull base thus includes SWI/SNF complex-deficient sinonasal carcinomas, SNUC, neuroendocrine tumor, and olfactory neuroblastomas.

HPV-associated multiphenotypic sinonasal carcinoma is the other new tumor in the 5th edition that was provisionally included as HPV-related carcinoma with adenoid cystic-like

features in the 4th edition.^{1,2} This unique neoplasm exhibits histologic features of both surface-derived and salivary gland carcinomas, hence the name “multiphenotypic.” This is associated with high-risk HPV (usually type 33) and, unlike adenoid cystic carcinoma, is restricted to the SNT. Most tumors affect the nasal cavity (89%), with a predilection for the turbinate, with or without involvement of the paranasal sinuses. There is a wide age range of presentation from the second-to-ninth decades, with slightly higher predilection for women (1.5:1). Tumors show

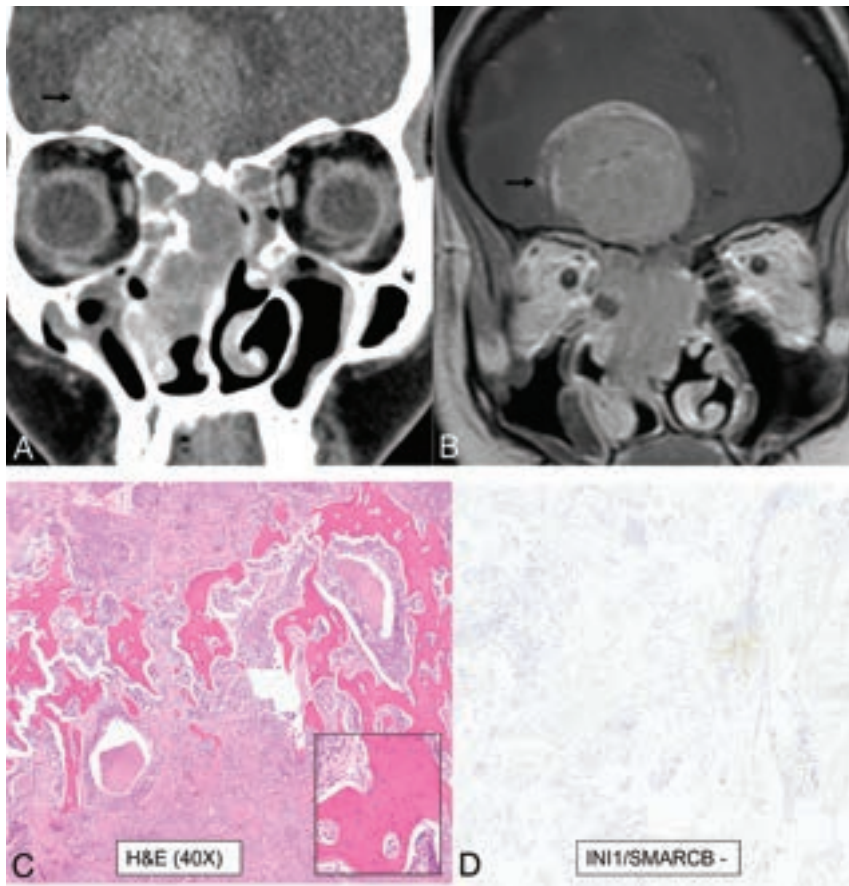


FIG 6. SWI/SNF complex-deficient sinonasal carcinoma in a 30-year-old woman. A large infiltrative sinonasal mass with intracranial extension (A and B, black arrows) and marked peritumoral edema in the frontal lobes. Peritumoral cysts were noted in the intracranial component. Radiologic differentials included olfactory neuroblastoma, SNUC, SWI/SNF complex-deficient (SMARCB1-deficient) tumor, *NUT* carcinoma, and high-grade neuroendocrine tumor. H&E stain shows a basaloid low-power appearance (C) with hyaline-appearing cytoplasm showing a “pink” cell appearance, in which the tumor cells are somewhat plasmacytoid (C, inset). Complete loss of INI1/SMARCB1 (D) expression by immunohistochemistry defines this tumor, as seen here. Immunostaining for *NUT* and synaptophysin (a neuroendocrine marker) were negative.

myoepithelial differentiation on histology, with strong and diffuse nuclear and cytoplasmic p16 immunoreactivity (a surrogate marker for HPV) and do not have the *MYB/MYBL1* fusions that characterize most adenoid cystic carcinomas.^{22,23} Descriptions of the imaging appearance are limited and include nasal cavity origin, infiltrative appearance with heterogeneous T2 hyperintensity, periosteal reaction, and hyperostotic changes (Fig 7). Despite having adenoid cystic features, perineural tumor spread is uncommon. No cystic nodal metastases have yet been described, despite the HPV origin of these tumors.^{8,24} This is different from HPV-induced SCCa, which originates along the tonsillar crypts, is most commonly caused by type 16 HPV, has nonkeratinizing SCCa morphology, and in which regional/nodal metastases are common.

Adenocarcinomas. Adenocarcinomas of the SNT arise from the respiratory epithelium or the underlying seromucinous glands. These malignancies were previously (3rd edition) divided into salivary type and nonsalivary type. The salivary type (eg, adenoid cystic carcinomas) since then has been removed from the SNT

chapter and placed under the salivary gland chapter. The nonsalivary type is further divided into intestinal-type and nonintestinal-type adenocarcinomas. Intestinal-type adenocarcinoma is the second most common type of sinonasal adenocarcinoma, most often localized in the ethmoid sinus (40%), followed by the nasal cavity (25%). These tumors are aggressive and frequently involve adjacent structures, including the orbit, pterygopalatine fossa, and infratemporal fossa, along with intracranial extension. Nonintestinal-type adenocarcinomas are usually low-grade, often arising in the nasal cavity (along nasal turbinates) with an imaging appearance similar to that of inverted papillomas. The low-grade subtypes present as solid masses, filling the nasal cavity and/or paranasal sinuses, with no osseous destruction. Emerging molecular studies suggest that these low-grade adenocarcinomas have distinctive mutations (eg, *CTNNB1*) or fusions (eg, *ETV6-NTRK3*). Imaging shows a poorly circumscribed, enhancing SNT mass and is used to assess the disease burden and local extension.^{25,26}

Mesenchymal Tumors and Other Tumors of the SNT

Multiple categories of tumor have been separately described in the new edition, including benign soft-tissue tumors and borderline, low-grade malignant soft-tissue tumors. However, some soft-tissue tumors that occur exclusively or predominantly in the sinonasal region

were retained in the newly created category of mesenchymal tumors. These include angiofibroma, glomangiopericytoma, biphenotypic sinonasal sarcoma, and chordoma. Among these, biphenotypic sinonasal sarcoma is the relatively recent entity, introduced only in the 4th edition (2017).^{1,2} These tumors have neural and myogenic features and are histologically similar to cellular schwannomas or malignant peripheral nerve sheath tumors. Rearrangement of the *PAX3* gene is required for the diagnosis of these tumors. They are characteristically seen in middle-aged women, arising from the nasal cavity or ethmoid sinus. Imaging reveals locally aggressive nasoethmoid-enhancing masses that can erode through the orbits and skull base with frequent hyperostotic bony changes (Fig 8).

Neuroectodermal tumor and the hematolymphoid tumor categories have also been removed and described in dedicated chapters. The only exception is olfactory neuroblastoma, which has been retained in the “other” category of tumors because these tumors occur predominantly within the SNT. These are malignant, neuroectodermal neoplasms with a bimodal age distribution and

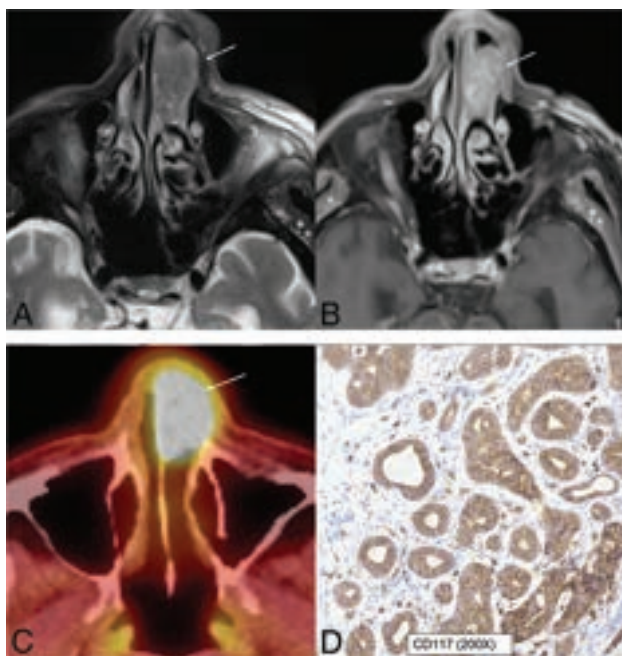


FIG 7. HPV-associated multiphenotypic sinonasal carcinoma in an 85-year-old man. A well-circumscribed, low T2-signal mass (A, arrow) with avid contrast enhancement (B, arrow) seen within the left nasal cavity with smooth expansion and high FDG uptake on PET/CT (C, arrow). CD117 (c-kit) (D) highlights the tumor ducts in a staining pattern that is the inverse of the myoepithelial cell markers. Findings of in situ hybridization for HPV (E6/E7) RNA was positive.

are confined to the cribriform plate, superior turbinate, and superior half of the nasal septum. With intracranial extension, peritumoral cysts between the mass and underlying brain are often present. Tumor cysts and speckled calcifications are characteristic of the intracranial component of the tumor, though not exclusive. These tumors show diffuse staining for conventional neuroendocrine markers and S-100 protein, with high avidity on radionuclide studies including indium-111 (^{111}In) pentetate (OctreoScan; https://www.accessdata.fda.gov/drugsatfda_docs/label/2022/020314s012bl.pdf) and gallium-67 (^{67}Ga) DOTATATE PET/CT studies.^{27,28}

Adamantinomatous craniopharyngioma is also a new entity in the “other” category. This is, however, only for taxonomic clarity, with the entity being shifted from the nasopharynx (4th edition) to the SNT. Adamantinomatous craniopharyngioma arises from the cellular elements related to the Rathke pouch and can occur anywhere along the craniopharyngeal canal, most commonly in the suprasellar region and very rarely in the nasopharynx and SNT. Although the latter is exceedingly rare, the SNTs are nevertheless the only ectopic location of this tumor. These are seen as cystic masses with peripheral enhancement and prominent calcifications.^{1,29,30}

Emerging Entities

Advancements in genetic studies have led to better understanding and more accurate classification of undifferentiated carcinomas. Studies have shown that a significant proportion of what remains in the SNUC group has underlying *IDH2* hotspot mutations with better prognosis. *IDH2*-mutated sinonasal carcinoma will likely

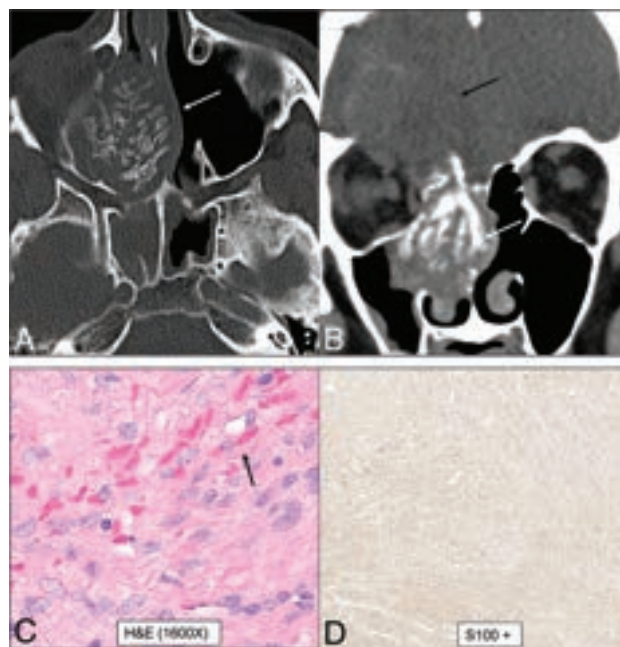


FIG 8. Biphenotypic sinonasal sarcoma (low-grade sinonasal sarcoma with neural and myogenic features) in a 68-year-old woman. Marked intralesional hyperostotic changes are noted within the nasal component of the tumor (A and B, white arrows). The tumor is well-margined but locally aggressive with intracranial extension (B, black arrows). H&E stain (C) shows the tumor growing as fascicles of uniform spindle cells with minimal mitotic activity or atypia and overt rhabdomyoblastic differentiation in the form of strap cells (C, arrow). Diffusely positive staining of S-100 (nerve sheath tumor marker) seen as diffuse brown staining (D). Fluorescence in situ hybridization testing was positive for *PAX3* gene rearrangement, supporting the above diagnosis.

be a separate tumor entity in future editions. The diagnosis of *IDH2*-mutated SNT is increasingly being made, given the availability of inexpensive immunohistochemical studies, and it is now recognized as the most common variant present in SNUC (accounting for 50%–88% of cases). The median age of patients with *IDH2*-mutated sinonasal carcinomas is 57 years, compared with 71 years in *IDH2* wild-type, with better prognosis. These range from 2.5 to 7 cm, most commonly occurring in the nasal cavity and ethmoid sinus and usually present with advanced disease. Histologically, these mimic undifferentiated or poorly differentiated carcinomas but show *IDH2*-positivity (Fig 1).^{4,31} The radiographic differential for an aggressive sinonasal mass with orbital and intracranial extension would include SNUC (including the *IDH*-mutant subset) neuroendocrinal tumors, neuroectodermal tumors (like olfactory neuroblastoma), and tumors with defined molecular alterations (*NUT*, *SMARCB1*, and *SMARCA4/A2*), with significant overlap in imaging features (Fig 9).

Along similar lines, *DEK-AFF2* carcinoma is an emerging entity, currently included as a subtype of nonkeratinizing SCCa, most commonly occurring within the nasal cavity with a peculiar propensity for the posterior part of the middle turbinate (Fig 10). Epidemiologic data on this entity are limited, with the largest series of 13 patients (Rooper et al³²) having a median age at presentation of 56 years (range, 18–79 years). Limited published data on this entity have shown that despite bland histologic features,

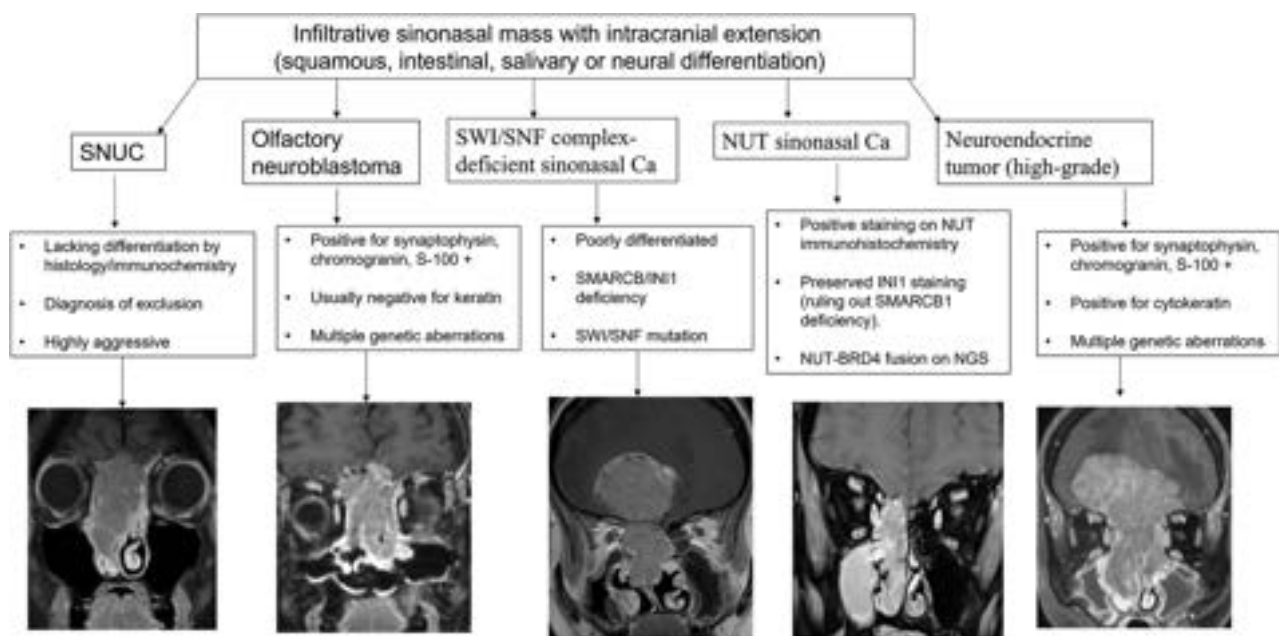


FIG 9. Radiographic differentials of an invasive sinonasal mass with intracranial extension are broad with significant overlap in imaging features, including previously described characteristic findings like intratumoral cyst in olfactory neuroblastoma. The final diagnosis is now based primarily on immunohistochemistry and genetics rather than histology. The prognosis of these tumors is also widely variable depending on the underlying molecular and genetic expression. Ca indicates carcinoma.

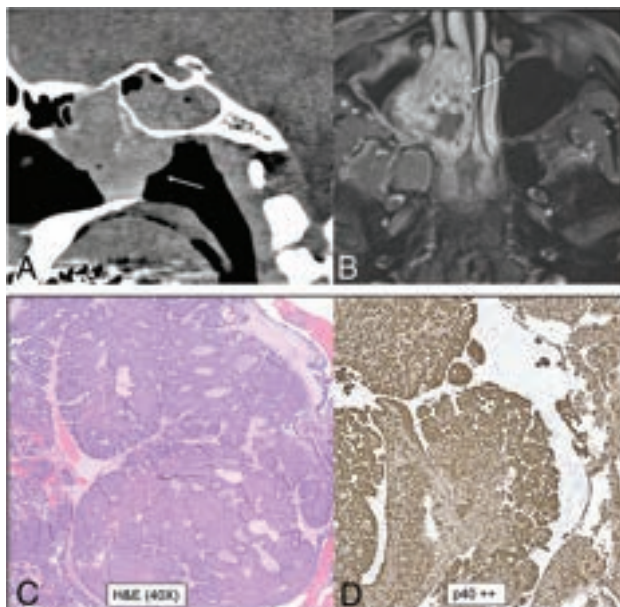


FIG 10. *DEK-AFF2* carcinoma. A smooth homogeneous density mass (A, arrow) noted within the posterior ethmoid and nasal cavity with heterogeneous enhancement on MRI (B, arrow). The mass blocks the sphenothmoidal recess with trapped secretions in the sphenoid sinus. The carcinoma demonstrates classic features of nonkeratinizing SCCa, with blunted papillae and invasion into the stroma as interconnecting ribbons on the H&E stains (C) with infiltrate of inflammatory cells, including lymphocytes and neutrophils. The tumor is diffusely positive for the squamous immunohistochemical marker p40 (D). To evaluate fusion-driven tumor, we performed targeted RNA sequencing, and *DEK-AFF2* fusion was found, with breakpoints of *DEK* (exon 7) and *AFF2* (exon 6).

these carcinomas are clinically aggressive tumors, with a higher risk of recurrence, nodal spread, and metastasis compared with the parent category. *DEK-AFF2* carcinomas should be considered in the differential diagnosis of high-grade sinonasal malignancies. This subcategory, however, shows an excellent response to immunotherapy. There is increasing evidence that nonkeratinizing SCCa with *DEK-AFF2* is a distinctive tumor entity and serves as a good example of how “poorly differentiated” and “undifferentiated” tumors are being continuously refined.^{32,33} RNA sequencing is increasingly being performed on multiple sinonasal neoplasms because immunohistochemical assays are not available, especially for newer entities. Finally, there is a newly created section dedicated to genetic tumor syndromes involving the head and neck. The initiative was undertaken to facilitate better understanding of the tumors and associated syndromes and to establish recommendations for monitoring and treating these patients. Many tumors that were thought to be sporadic earlier are now known to be a part of syndrome complex, with identification of specific mutations. A great example of this would be the *PTEN* germline mutation found in about 85% of patients with Cowden syndrome, characterized by multiple hamartomas involving the oral cavity.^{1,34}

CONCLUSIONS

In the past decade, molecular markers and genetics have revolutionized the taxonomy of sinonasal and skull base tumors, leading to recognition of new entities and more accurate understanding of the underlying pathogenesis. The most striking example of this change would be the SNUC tumors in which multiple new entities have been described on the basis of the underlying mutation. Radiologists should be abreast of these changes because

discussion about these molecular markers and genetic changes is common during multidisciplinary tumor board meetings. This field is evolving, with continuous improvement in the classification system, recognition of new entities, and standardization in diagnosis, paving the way for targeted treatments.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

- WHO Classification of Tumours Editorial Board. *Head and neck tumours*. Lyon (France): International Agency for Research on Cancer; 2022. WHO classification of tumours series, 5th ed. vol. 9. Available from: <https://tumourclassification.iarc.who.int/chapters/52>
- El-Naggar AK, Chan JKC, Grandis JR, et al. *WHO classification of head and neck tumours*. Lyon (France): International Agency for Research on Cancer; 2017
- Thompson LDR, Bishop JA. **Update from the 5th Edition of the World Health Organization Classification of Head and Neck Tumors: Nasal Cavity, Paranasal Sinuses and Skull Base.** *Head Neck Pathol* 2022;16:1–18 CrossRef Medline
- Glöss S, Jurmeister P, Thieme A, et al. **IDH2 R172 mutations across poorly differentiated sinonasal tract malignancies: forty molecularly homogenous and histologically variable cases with favorable outcome.** *Am J Surg Pathol* 2021;45:1190–204 CrossRef Medline
- Alzumaili B, Sadow PM. **IDH2-mutated sinonasal tumors: a review.** *Adv Anat Pathol* 2023;30:104–11 CrossRef Medline
- Schultz KA, Williams GM, Kamihara J, et al. **DICER1 and associated conditions: identification of at-risk individuals and recommended surveillance strategies.** *Clin Cancer Res* 2018;24:2251–61 CrossRef Medline
- Lee JT, Garg R, Brunworth J, et al. **Sinonasal respiratory epithelial adenomatoid hamartomas: series of 51 cases and literature review.** *Am J Rhinol Allergy* 2013;27:322–28 CrossRef Medline
- Dean KE, Shatzkes D, Phillips CD. **Imaging review of new and emerging sinonasal tumors and tumor-like entities from the Fourth Edition of the World Health Organization Classification of Head and Neck Tumors.** *AJNR Am J Neuroradiol* 2019;40:584–90 CrossRef Medline
- Fleming KE, Perez-Ordoñez B, Nasser JG, et al. **Sinonasal seromucinous hamartoma: a review of the literature and a case report with focal myoepithelial cells.** *Head Neck Pathol* 2012;6:395–99 CrossRef Medline
- Guillerman RP, Foulkes WD, Priest JR. **Imaging of DICER1 syndrome.** *Pediatr Radiol* 2019;49:1488–505 CrossRef Medline
- Mason K, Navaratnam A, Theodorakopoulou E, et al. **Nasal chondromesenchymal hamartoma (NCMH): a systematic review of the literature with a new case report.** *J Otolaryngol Head Neck Surg* 2015;44:28 CrossRef Medline
- Vorasubin N, Vira D, Suh JD, et al. **Schneiderian papillomas: comparative review of exophytic, oncocytic, and inverted types.** *Am J Rhinol Allergy* 2013;27:287–92 CrossRef Medline
- Jeon TY, Kim HJ, Chung SK, et al. **Sinonasal inverted papilloma: value of convoluted cerebriform pattern on MR imaging.** *AJNR Am J Neuroradiol* 2008;29:1556–60 CrossRef Medline
- Yang B, Li J, Dong J. **MR imaging and CT features of oncocytic papilloma of the sinonasal tract with comparison to inverted papilloma.** *Br J Radiol* 2018;91:20170957 CrossRef Medline
- Buchwald C, Franzmann MB, Tos M. **Sinonasal papillomas: a report of 82 cases in Copenhagen County, including a longitudinal epidemiological and clinical study.** *Laryngoscope* 1995;105:72–79 CrossRef Medline
- Glâtre R, De Kermadec H, Alsamad IA, et al. **Exophytic sinonasal papillomas and nasal florid papillomatosis: a retrospective study.** *Head Neck* 2018;40:740–46 CrossRef Medline
- Raghavan P, Phillips CD. **Magnetic resonance imaging of sinonasal malignancies.** *Top Magn Reson Imaging* 2007;18:259–67 CrossRef Medline
- Agger A, von Buchwald C, Madsen AR, et al. **Squamous cell carcinoma of the nasal vestibule 1993-2002: a nationwide retrospective study from DAHANCA.** *Head Neck* 2009;31:1593–99 CrossRef Medline
- Ansa B, Goodman M, Ward K, et al. **Paranasal sinus squamous cell carcinoma incidence and survival based on surveillance, epidemiology, and end results data, 1973 to 2009.** *Cancer* 2013;119:2602–10 CrossRef Medline
- Lee T, Cho J, Baek CH, et al. **Prevalence of NUT carcinoma in head and neck: analysis of 362 cases with literature review.** *Head Neck* 2020;42:924–38 CrossRef Medline
- Edgar M, Caruso AM, Kim E, et al. **NUT midline carcinoma of the nasal cavity.** *Head Neck Pathol* 2017;11:389–92 CrossRef Medline
- Shatzkes DR, Ginsberg LE, Wong M, et al. **Imaging appearance of SMARCB1 (INI1)-deficient sinonasal carcinoma: a newly described sinonasal malignancy.** *AJNR Am J Neuroradiol* 2016;37:1925–29 CrossRef Medline
- Bishop JA, Ogawa T, Stelow EB, et al. **Human papillomavirus-related carcinoma with adenoid cystic-like features: a peculiar variant of head and neck cancer restricted to the sinonasal tract.** *Am J Surg Pathol* 2013;37:836–44 CrossRef Medline
- Bishop JA, Andreasen S, Hang JF, et al. **HPV-related multiphenotypic sinonasal carcinoma: an expanded series of 49 cases of the tumor formerly known as HPV-related carcinoma with adenoid cystic carcinoma-like features.** *Am J Surg Pathol* 2017;41:1690–701 CrossRef Medline
- Leivo I. **Sinonasal adenocarcinoma: update on classification, immunophenotype and molecular features.** *Head Neck Pathol* 2016;10:68–74 CrossRef Medline
- Eggesbø HB. **Imaging of sinonasal tumours.** *Cancer Imaging* 2012;12:136–52 CrossRef Medline
- Dublin AB, Bobinski M. **Imaging characteristics of olfactory neuroblastoma (esthesioneuroblastoma).** *J Neurol Surg B Skull Base* 2016;77:1–5 CrossRef Medline
- Rostomily RC, Elias M, Deng M, et al. **Clinical utility of somatostatin receptor scintigraphic imaging (octreoscan) in esthesioneuroblastoma: a case study and survey of somatostatin receptor subtype expression.** *Head Neck* 2006;28:305–12 CrossRef Medline
- Magill JC, Ferguson MS, Sandison A, et al. **Nasal craniopharyngioma: case report and literature review.** *J Laryngol Otol* 2011;125:517–19 CrossRef Medline
- Yu X, Liu R, Wang Y, et al. **Infraselar craniopharyngioma.** *Clin Neurol Neurosurg* 2012;114:112–19 CrossRef Medline
- Riobello C, López-Hernández A, Cabal VN, et al. **IDH2 mutation analysis in undifferentiated and poorly differentiated sinonasal carcinomas for diagnosis and clinical management.** *Am J Surg Pathol* 2020;44:396–405 CrossRef Medline
- Rooper LM, Agaimy A, Dickson BC, et al. **DEK-AFF2 carcinoma of the sinonasal region and skull base: detailed clinicopathologic characterization of a distinctive entity.** *Am J Surg Pathol* 2021;45:1682–93 CrossRef Medline
- Bishop JA, Gagan J, Paterson C, et al. **Nonkeratinizing squamous cell carcinoma of the sinonasal tract with DEK-AFF2: further solidifying an emerging entity.** *Am J Surg Pathol* 2021;45:718–20 CrossRef Medline
- Nosé V. **Genodermatosis affecting the skin and mucosa of the head and neck: clinicopathologic, genetic, and molecular aspect-PTEN-hamartoma tumor syndrome/Cowden syndrome.** *Head Neck Pathol* 2016;10:131–38 CrossRef Medline

Systematic Literature Review of Machine Learning Algorithms Using Pretherapy Radiologic Imaging for Glioma Molecular Subtype Prediction

Jan Lost, Tej Verma, Leon Jekel, Marc von Reppert, Niklas Tillmanns, Sara Merkaj, Gabriel Cassinelli Petersen, Ryan Bahar, Ayyüce Gordem, Muhammad A. Haider, Harry Subramanian, Waverly Brim, Ichiro Ikuta, Antonio Omuro, Gian Marco Conte, Bernadette V. Marquez-Nostra, Arman Avesta, Khaled Bousabarah, Ali Nabavizadeh, Anahita Fathi Kazerooni, Sanjay Aneja, Spyridon Bakas, MingDe Lin, Michael Sabel, and Mariam Aboian



ABSTRACT

BACKGROUND: The molecular profile of gliomas is a prognostic indicator for survival, driving clinical decision-making for treatment. Pathology-based molecular diagnosis is challenging because of the invasiveness of the procedure, exclusion from neoadjuvant therapy options, and the heterogeneous nature of the tumor.

PURPOSE: We performed a systematic review of algorithms that predict molecular subtypes of gliomas from MR Imaging.

DATA SOURCES: Data sources were Ovid Embase, Ovid MEDLINE, Cochrane Central Register of Controlled Trials, Web of Science.

STUDY SELECTION: Per the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, 12,318 abstracts were screened and 1323 underwent full-text review, with 85 articles meeting the inclusion criteria.

DATA ANALYSIS: We compared prediction results from different machine learning approaches for predicting molecular subtypes of gliomas. Bias analysis was conducted for each study, following the Prediction model Risk Of Bias Assessment Tool (PROBAST) guidelines.

DATA SYNTHESIS: Isocitrate dehydrogenase mutation status was reported with an area under the curve and accuracy of 0.88 and 85% in internal validation and 0.86 and 87% in limited external validation data sets, respectively. For the prediction of *O6-methylguanine-DNA methyltransferase* promoter methylation, the area under the curve and accuracy in internal validation data sets were 0.79 and 77%, and in limited external validation, 0.89 and 83%, respectively. PROBAST scoring demonstrated high bias in all articles.

LIMITATIONS: The low number of external validation and studies with incomplete data resulted in unequal data analysis. Comparing the best prediction pipelines of each study may introduce bias.

CONCLUSIONS: While the high area under the curve and accuracy for the prediction of molecular subtypes of gliomas are reported in internal and external validation data sets, limited use of external validation and the increased risk of bias in all articles may present obstacles for clinical translation of these techniques.

ABBREVIATIONS: AUC = area under the curve; DL = deep learning; MGMT = *O6-methylguanine-DNA methyltransferase*; ML = machine learning; SVM = support vector machine; WHO = World Health Organization; IDH = *isocitrate dehydrogenase*

Gliomas account for approximately 33% of brain tumor diagnoses; among adults, >50% of these cases are high-grade

gliomas.^{1,2} The 2021 World Health Organization (WHO) classification identified different forms of gliomas based on pathologic characteristics of biopsied or resected tumor and the molecular subtype.³ This method of diagnosis requires invasive sampling of the tumor, which risks surgery and the potential for tumor misclassification due to tumor heterogeneity with sampling bias.^{4,5}

Received February 27, 2023; accepted after revision August 1.

From the Departments of Radiology and Biomedical Imaging (J.L., T.V., L.J., M.v.R., N.T., S.M., G.C.P., R.B., A.G., M.A.H., H.S., W.B., B.V.M.-N., A.A., M.L., M.A.), Neurology and Yale Cancer Center (A.O.), and Therapeutic Radiology (S.A.), Yale School of Medicine, New Haven, Connecticut; Department of Neurosurgery (J.L., M.S.), Heinrich-Heine-University, Duesseldorf, Germany; Department of Radiology (I.I.), Mayo Clinic Arizona, Phoenix, Arizona; Department of Radiology (G.M.C.), Mayo Clinic, Rochester, Minnesota; Visage Imaging Inc (K.B., M.L.), San Diego, California; Department of Radiology (A.N.), Perelman School of Medicine, Hospital of University of Pennsylvania, University of Pennsylvania, Philadelphia, Pennsylvania; Department of Neurosurgery (A.F.K.), Center for Biomedical Image Computing and Analytics (S.B.), and Department of Radiology (S.B.), Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; Division of Neurosurgery (A.F.K.), and Center for Data-Driven Discovery (A.F.K.), Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; and Richards Medical Research Laboratories (S.B.), Philadelphia, Pennsylvania.

Please address correspondence to Mariam Aboian, MD, Department of Radiology and Biomedical Imaging, Yale School of Medicine, 333 Cedar St, New Haven, CT 06510; e-mail: mariam.aboian@yale.edu

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

Indicates article with online supplemental data.
<http://dx.doi.org/10.3174/ajnr.A8000>

In addition, the inability to accurately predict tumor subtype before surgery can result in limited access to neoadjuvant therapies. Of note, the class of tumor affects a patient's predicted survival as well as medical and surgical treatment options.^{3,6}

Recent advancements in machine learning (ML) applications in neuro-oncology have shown promise in tumor segmentation,^{7,8} differentiating gliomas from other intracranial malignancies such as brain metastases⁹ and lymphomas,¹⁰ predicting glioma grade,¹¹ and predicting the patient's overall survival.^{12,13} Most of the published literature on applying ML to neuro-oncology demonstrates a high area under the curve (AUC) and accuracy in the internal testing data sets, but validation of algorithms on external data sets is limited to a very few studies. As a result, there is a limited translation of algorithms among different hospitals and study settings. In addition, previous systematic reviews have demonstrated that most of the literature is focused on small well-curated data sets with an over-representation of commonly used data sets, such as RSNA-ASNR-MICCAI Brain Tumor Segmentation (BraTS; <https://www.kaggle.com/datasets/dschettler8845/brats-2021-task1>) and The Cancer Imaging Archive (TCIA; <https://wiki.cancerimagingarchive.net/display/Public/Collections>).^{8-12,14}

This feature further limits translation of algorithms between different institutions. In addition, ML literature in neuro-oncology often scores low in established reporting guidelines such as Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis Or Diagnosis (TRIPOD; <https://www.equator-network.org/reporting-guidelines/tripod-statement/>) and Checklist for Artificial Intelligence in Medical Imaging (CLAIM; <https://pubs.rsna.org/doi/10.1148/ryai.2020200029>) with high bias reported by the Prediction model Risk Of Bias Assessment Tool (PROBAST).¹⁵⁻¹⁷ These findings suggest that increased use of standardized reporting criteria in publications and the establishment of large databases of annotated images in individual hospitals are critically needed to translate useful algorithms into patient care and to bridge the gap between the diagnosis and targeted therapy. ML can potentially identify patterns that remain invisible to a radiologist's clinical interpretation and can extract information from pretreatment imaging that can influence the implementation of neoadjuvant therapy and the extent of surgical resection.

Our study evaluated the literature that uses ML to predict molecular subtypes of gliomas, such as *isocitrate dehydrogenase (IDH)* mutation status and *O6-methylguanine-DNA methyltransferase (MGMT)* promoter methylation status, which can change treatment options for patients. As an example, *IDH*-mutant astrocytomas have better survival and different treatment options compared with glioblastomas that are uniformly *IDH* wild-type.³ Glioblastomas have dismal survival and require maximal resection to improve survival.¹⁸ Therefore, predicting *IDH* mutation on preoperative imaging can change the patient's treatment strategy and outcomes. In addition, the prediction of *MGMT* methylation status can predict which patients will respond to temozolomide therapy and, therefore, would influence neoadjuvant options before resection.¹⁹ To guide medical care, the algorithm must have high precision and sensitivity and be translatable to multiple hospitals with different imaging protocols. The goal of our study was to critically evaluate the literature that reports different

algorithms for predicting molecular subtypes of gliomas and to assess the potential obstacles to translating these algorithms into clinical practice.

MATERIALS AND METHODS

Study Selection

Literature screening for this systematic review conformed to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, using Ovid EMBASE, Ovid MEDLINE, Cochrane Central Register of Controlled Trials, and the Web of Science Core Collection as databases. The study was registered with the Prospective Register of Systematic Reviews (PROSPERO, CRD42020209938). A clinical librarian collected data in September 2020, January 2021, and September 2021, respectively. The literature search strategy involved the use of keywords such as "artificial intelligence," "machine learning," "deep learning," "radiomics," "MR imaging," "glioma," "glioblastoma," and related terms. We identified 12,470 studies (Fig 1A) and added them for further screening to the Covidence Software (Veritas Health Innovation).

After removing 152 duplicates, a neuroradiology attending physician, a neuroradiology resident, and 3 graduate students screened 12,318 studies. Conflicting assessments were resolved by the board-certified neuroradiology attending physician after discussion with screeners. After abstract reviews, 10,995 studies were excluded due to a lack of ML and neuro-oncology applicability, resulting in 1323 full-text reviews. A secondary full-text review was conducted on 886 studies, by either a radiology resident or a graduate student, followed by a second review by a board-certified neuroradiologist attending physician.

We predefined 8 uniform exclusion criteria: 1) abstract-only, 2) no application of ML reported, 3) not an original article, 4) not published in English, 5) no investigation of glioma/glioblastoma, 6) unrelated to MR imaging, MR spectroscopy, or PET imaging, 7) no human research subjects, and 8) duplicates. Overall, 437 studies met ≥ 1 of the predefined exclusion criteria and were subsequently excluded (Fig 1A). In the end, for this systematic review, 85 studies were included that specifically analyzed molecular subtype prediction in glioma, using ML techniques based on pretherapy imaging. This resulted in 801 studies being excluded from further review.

Data Extraction

Data extraction was performed by 3 graduate students and 1 undergraduate in Excel (Microsoft 2022). Studies were reviewed twice, and disagreements were resolved in regular meetings with a supervising neuroradiology attending physician until a consensus was reached. Extracted information included article characteristics (title, authors, publication year), patient data (patient number, cohort sizes, data sources), tumor classification (tumor type, analyzed molecular subtypes), model characteristics (imaging data and best-performing ML algorithms), as well as validation techniques. Generally, accuracy results from the best-performing prediction pipeline were reported. If patient cohorts were split up within a study, different prediction results were reported. We accounted for the use of external validation if patient cohorts were geographically split. Data from the same

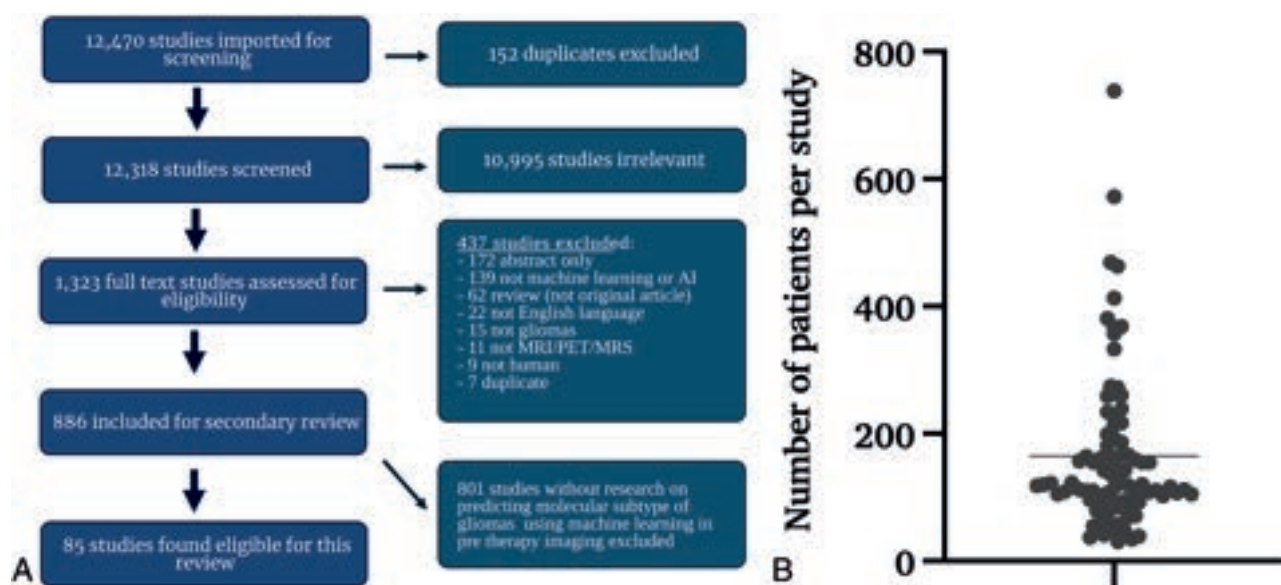


FIG 1. A, Inclusion/exclusion criteria and resultant study data. Flow chart represents screening workflow and exclusion criteria to visualize the eligibility of studies. The search strategy included keywords “artificial intelligence,” “machine learning,” “deep learning,” “radiomics,” “MR imaging,” “glioma,” “glioblastoma,” and related terms. An independent librarian reviewed the data. We predefined 8 uniform exclusion criteria: 1) abstract-only, 2) no application of ML reported, 3) not an original article, 4) not published in English, 5) no investigation of glioma/glioblastoma, 6) unrelated to MR imaging, MR spectroscopy, or PET imaging, 7) no human research subjects, and 8) duplicates. B, Distribution of all patients per study included in training or validation of a predictive model. We excluded patients whose data were strictly used for models other than the prediction of molecular subtypes. The line indicates the mean number of patients. AI indicates artificial intelligence.

source, solely split up by the time of inclusion, were classified as internal validation. To give more insight into the different validation types used, we split internal validation techniques into cross-validation and holdout-validation. For an overview of extracted data, see the Online Supplemental Data.

Meta-analysis

All studies reporting AUC and 95% confidence intervals were subjected to a meta-analysis using MedCalc for Windows, Version 20.009 (MedCalc software). Heterogeneity was examined using the Higgins I^2 test, publication bias was evaluated using the Egger test, and the results of the quantitative analysis were illustrated with a forest plot (Online Supplemental Data).

Risk of Bias Assessment

Risk of bias was assessed using PROBAST.¹⁷ Per PROBAST guidelines, studies were classified into development, validation, or development and validation studies. All included studies established predictive models and, therefore, were at least considered development studies; those that additionally tested their developed model on an independent cohort were considered to have both development and validation.

PROBAST uses signaling questions across 4 distinct domains to evaluate potential biases in each study. The first domain (participants) pertains to the data sources and participant enrollment. The second domain (predictors) evaluates the definition and measurement of predictors and their association with the outcome. Domain 3 (outcome) addresses potential biases in defining and terminating the outcome in each study. Last, domain 4 (analysis) evaluates whether inappropriate analysis methods were used or

important statistical considerations were overlooked.¹⁷ Additional information on the application of PROBAST can be found in the work of Moons et al,¹⁷ which is beyond the scope of this article. Signaling questions from each domain were assessed separately for development and validation cohorts. The average item scores from each study and the risk of bias for all 4 domains were evaluated using Excel (Microsoft).

RESULTS

After a full-text review of 886 studies, 85 studies published between 2016 and 2021 met the eligibility criteria of our analysis.

Patient Data

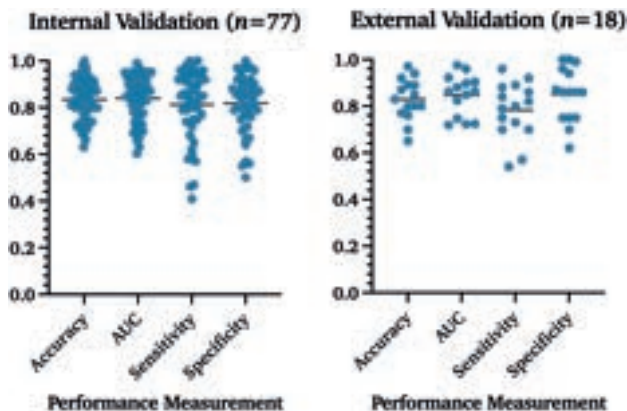
The mean patient number in all 85 studies was 165.32 (Fig 1B). Data were taken from single-center hospital data, public databases, or multicenter hospital data. The overall number of data sources included is higher than the total number of studies because 17 studies included patients from 2 sources, resulting in a total of 102 sources among the 85 studies. While 52% ($n = 53/102$) of patient data was obtained from single-center hospitals, 33% ($n = 34/102$) was from publicly available databases and 15% ($n = 15/102$) was from multiple institutions (Online Supplemental Data).

Description of Internal and External Validation Data Sets and Techniques

Of 85 studies, 7 reported >1 distinct model pipeline to predict glioma molecular subtypes, resulting in 95 analyzed pipelines.²⁰⁻²⁶ The best performing model results for each study are reported for molecular alterations in gliomas, including *IDH*, *MGMT*, *1p/19q* codeletion, histone *H3 K27M*, *ATRX*, *TERT*, and others. The results from external validation were reported. For studies with

Table 1: Mean performance measurements of all studies

	Accuracy	AUC	Sensitivity	Specificity
Internal validation	83% ($n = 66/77$)	0.84 ($n = 60/77$)	81% ($n = 50/77$)	82% ($n = 49/77$)
External validation	83% ($n = 15/18$)	0.85 ($n = 14/18$)	78% ($n = 15/18$)	85% ($n = 15/18$)
<i>P</i> value (Mann-Whitney <i>U</i> test)	.83	.79	.25	.45

**FIG 2.** Algorithm performance measurement in internal and external validation data sets. Percentages are reported as fractions to visualize measurements in 1 graph.

internal validation only, the holdout validation was prioritized over cross-validation. Techniques varied among leave-one-out-cross-validation, *n*-fold cross-validation to holdout, and external validation. In 5% ($n = 5/95$) of the articles, the validation cohorts were separated from the initial cohort by time only and were, therefore, re-classified as internal validation in our analysis. Overall, 81% ($n = 77/95$) used internal validation of any kind, while only 37% ($n = 35/95$) used forms of cross-validation and 39% ($n = 37/95$) used holdout validation. Only 19% ($n = 18/95$) reported accuracy from an external validation cohort.

We performed a Mann-Whitney *U* test on the studies that reported accuracy, AUC, sensitivity, or specificity to analyze whether statistically significant internal and external validation differences were present. No significant differences were found between both groups for all 4 categories (Table 1 and Fig 2).

Performance of Molecular Subtype Prediction

If studies had multiple predictions with the same model, the highest AUC and accuracy are reported. In internal validation studies, *IDH* mutation status was the most frequently evaluated molecular subtype, with the overall highest mean AUC and accuracy values of 0.88 and 85%, respectively. *MGMT* promoter methylation was reported in 12 studies as the best predicted subtype, with a mean AUC and accuracy of 0.82 and 80%. The *1p/19q* codeletion was predicted in 9, with a mean AUC and accuracy of 0.84 and 85%. Prediction models for histone *H3 K27M* were identified in 6 studies with a mean AUC and accuracy of 0.80 and 81%. *ATRX* status was predicted in 1 study with an AUC and accuracy of 0.93 and 92%. *TERT* promoter mutation was predicted in 2 studies, resulting in a mean AUC and accuracy of 0.77 and 77%. Other subtypes included common subgroups like *EGFR*,^{27,28} *p53*,²² *RBI*,²⁹ and *VEGF*,³⁰ as well as a pooled accumulation of genetic

information that was predicted as groups (Online Supplemental Data).³¹ Buda et al³¹ analyzed the prediction of pathologic biomarkers Ki-67, S-100, and glial fibrillary acidic protein,²³ which were also included in this category. Overall, these studies achieved a mean AUC and accuracy of 0.77 and 82% (Fig 3A).

Among studies that used external validation, 7 studies predicted *IDH* status with a mean AUC = 0.89 and accuracy = 86%; 3 studies predicted *MGMT* promoter methylation with a mean AUC = 0.89 and accuracy = 83%; 2 studies predicted *1p/19q* codeletion status with a mean AUC = 0.82 and accuracy = 75%; 2 studies predicted *ATRX* status with a mean AUC = 0.72 and accuracy = 77%; and 1 study predicted *EGFR* status with a mean AUC = 0.82 and accuracy = 85%. Other molecular subtypes, such as *PTEN* (accuracy = 82.5%)³² and *BRAF* (AUC = 0.85)³³ mutation status, were each evaluated in 1 study. No study with external validation predicted histone *H3 K27M* status (Fig 3A).

Algorithm-Based Prediction Models

Overall, of internally validated models, 35% ($n = 27/77$) included tree-based; 27% ($n = 21/77$), support vector machine (SVM); 32% ($n = 25/77$), neural networks; and 5% ($n = 4/77$), other classifiers. In externally validated models, 44% ($n = 8/18$) were tree-based; 39% ($n = 7/18$), SVM; 6% ($n = 1/18$), neural networks; and 11% ($n = 2/18$), other classifiers (Table 3 and Fig 4B).

A neural network with a mean AUC and accuracy of 0.88 and 85% achieved the best overall prediction results. No statistically significant difference among all ML classifiers and neural networks was found with the Mann-Whitney *U* test. Tree-based algorithms and SVM performed slightly worse than neural networks with an AUC and accuracy of 0.82 and 82%, and 0.83 and 83%, respectively (Table 3).

In internal validation models, the overall mean AUC and accuracy for deep learning (DL) algorithms (0.82 and 82%) were higher than ML algorithms (0.88 and 86%) and were statistically significant by the Mann-Whitney *U* test ($P = .02$) (Fig 4B).

These results were not observed in external validation pipelines (Table 3).

Performance of Studies on the Prediction of *IDH* Mutation and *MGMT* Methylation Status

Mean AUC and accuracy for the prediction of *IDH* mutation status were 0.88 and 85% in internal validation studies and 0.85 and 86% in external validation studies. For the prediction of *MGMT* promoter methylation, AUC and accuracy in internal validation studies were 0.79 and 80% and 0.89 and 83% in external validation. While a *P* value of .02 indicated a statistically significant difference between AUC values for *IDH* and *MGMT* for internal validation cohorts, Mann-Whitney *U* tests did not show such a difference between accuracies, with a *P* value = .16 (Fig 5).

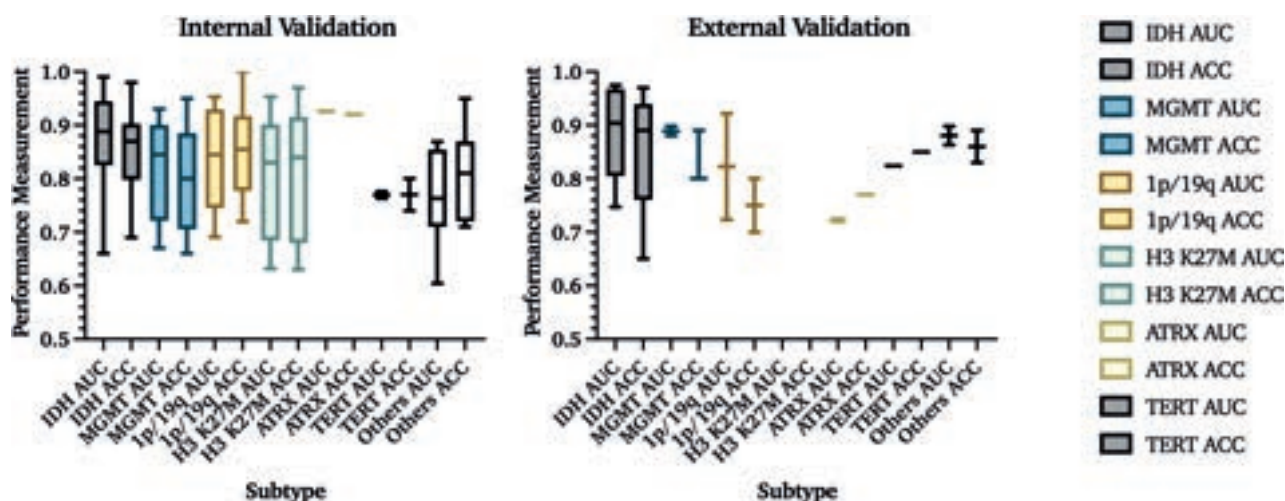


FIG 3. AUC and accuracy (ACC) results from internal and external validation studies. Results from 76 internal and 18 external validation studies are demonstrated on the basis of the molecular subtype that is being predicted. The *central line* in each result indicates the median value of the labeled subtype. Percentages are reported as fractions to provide visualization.

Table 2: Types of algorithms used to predict the molecular subtypes of gliomas and the number of studies that used them

	IDH	MGMT	1p/19q	H3 K27M	ATRX	TERT	Others
Internal validation							
Tree-based	11	4	4	4	0	0	4
SVM	9	1	2	2	1	2	4
Neural networks	14	6	3	0	0	0	2
Others	3	1	0	0	0	0	0
External validation							
Tree-based	4	1	0	0	1	1	1
SVM	2	1	2	0	1	0	1
Neural networks	1	0	0	0	0	0	0
Others	0	1	0	0	0	1	0

Table 3: Performance of algorithms

	Accuracy	AUC
Internal validation		
Tree-based	82% ($n = 23/27$)	0.82 ($n = 23/27$)
SVM	83% ($n = 12/21$)	0.83 ($n = 17/21$)
Neural network	85% ($n = 22/25$)	0.88 ($n = 18/25$)
Others	84% ($n = 3/4$)	0.85 ($n = 2/4$)
External validation		
Tree-based	85% ($n = 6/8$)	0.84 ($n = 7/8$)
SVM	82% ($n = 7/7$)	0.84 ($n = 4/7$)
Neural network	86% ($n = 1/1$)	0.86 ($n = 1/1$)
Others	89% ($n = 1/2$)	0.88 ($n = 2/2$)

Risk of Bias Assessment

A risk of bias assessment was performed per PROBAST guidelines (Fig 6 and Online Supplemental Data).¹⁷

The risk of bias was low in the participant, predictor, and outcome sections of development and validation studies. The risk of bias was high in all of the analysis sections of the studies, regardless of whether these were development or validation studies. The main reasons for high bias in the analysis sections were due to items 4.1 and item 4.4.

Item 4.1 assesses the number of patients relative to imaging features extracted, suggesting an overfitting issue. It scored as “no” in 77.64% ($n = 66/85$) of the studies. Additionally, item 4.4

suggesting a lack of reporting or incorrect handling of missing data, such as simple exclusion, was found in 98% ($n = 83/85$) of studies.

DISCUSSION

Diagnosis and treatment of gliomas are based on pathologic and molecular classifications outlined in the 2021 WHO Classification of CNS tumors. Noninvasive methods for image-based prediction of glioma molecular subtype on preoperative images are the next

frontier in neuro-oncology because they will provide information before surgical intervention and have the potential to change treatment options for patients with brain tumors. Since 2017, published literature has significantly increased, showing high prediction results for this classification task. However, a thorough assessment of this literature to identify algorithms that can be used for clinical translation have yet to be performed.

Our systematic review shows that literature in this field has several limitations, which include low patient numbers (mean = 165.31), limited use of geographically distinct validation data sets (18.95%), and limited use of multicenter hospital data (14.71%). The most predicted molecular biomarkers were *IDH* and *MGMT*, which are critical for classifying glioblastoma from lower-grade gliomas and for predicting response to temozolomide therapy. Testing for these is standard of care in clinical practice. Therefore, these results are more available than other molecular biomarkers.

We show that DL algorithms result in significantly higher AUC values in internal validation studies than ML (P value = .02), leading us to recommend developing DL algorithms for future applications. Our review highlights the feasibility of accurately predicting glioma molecular subtypes; however, only some studies addressed the need for transparency and interpretability of these prediction models and demonstrated a high risk of bias.

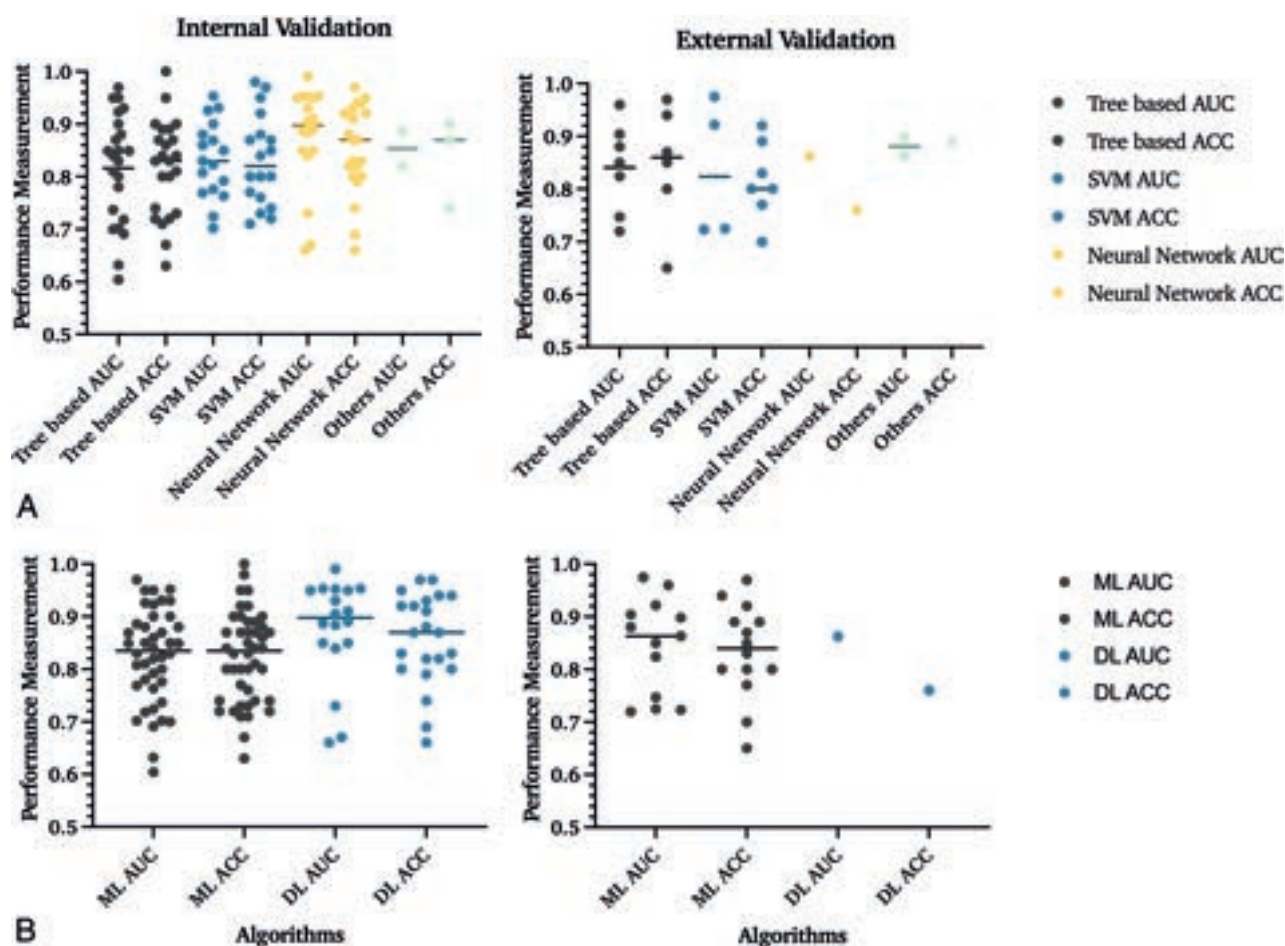


FIG 4. Comparison of performance of different ML algorithms in internal and external validation data sets. **A**, In internal validation studies, 35% ($n = 27/77$) used tree-based; 27% ($n = 21/77$), SVM; 32% ($n = 25/77$), neural network; and 5% ($n = 4/77$), other classifiers. The section named “Others” includes machine and deep learning algorithms, which cannot be classified into these 3 groups, and mixed classifiers with characteristics of multiple techniques. Lines indicate the mean value. **B**, Comparison of ML and DL algorithms. This figure refers to SVM and tree-based algorithms as overall ML algorithms. At the same time, all neural network-based classifiers are DL classifiers. In internal validation studies, 68% ($n = 52/77$) used ML algorithms, and 32% ($n = 25/77$) used DL classifiers. The DL algorithms demonstrated higher AUCs and statistically significant internal validation data sets. In the 95 patient cohorts analyzed, 68 studies used classic ML classifiers for their predictive models, while 26 used DL networks. The comparison of algorithms in external validation data sets was limited due to the small number of studies that validated DL algorithms. ACC indicates accuracy.

A unique aspect of our study is the broad inclusion criteria, which required screening of >12,000 abstracts before full-text review. Our study also contains the largest number of evaluated articles ($n = 85$) with the most characteristic features extracted ($n = 18$) (Online Supplementary Data).³⁴ In our analysis, we extracted not only feature characteristics of patient data sets but also imaging sequences, including advanced imaging modalities, algorithms, and types of ML algorithms with outcome assessment on internal and external validation data sets.

In addition to a thorough feature-extraction process from individual articles, we performed bias analysis with the most relevant assessment tool currently available, PROBAST.¹⁷ Prior studies demonstrated that ML approaches for the evaluation of gliomas have significant deficiencies in reporting quality as assessed by TRIPOD,^{9–11} aligning with our results of high bias. Unlike the previous systematic review, which had more restrictive inclusion criteria, our systematic review included 35 studies that did not report AUC ($n = 21/95$) or accuracy ($n = 14/95$) in evaluating their

respective models. Our less restrictive inclusion criteria, inclusion of studies that did not report both AUC and accuracy allowed us to better evaluate the distribution of different ML approaches in this field. Furthermore, our review did not exclude studies with non-MR imaging modalities or those with missing data, resulting in a total of 44 studies being included in our analysis.³⁴ While our findings generally agree with the results of prior work, we provide additional information on the differences in internal and external validation studies (conclusions), prediction performances of analyzed glioma molecular subtypes (section 4.4), and details of different image modalities with their corresponding performance metrics (Online Supplemental Data).

We recommend future studies evaluating ML algorithms in the imaging of gliomas to develop prediction algorithms based on larger data sets with geographically distinct data for model validation to provide generalizable results.³⁵ Until now, most studies relied on single-center hospital data and publicly available data sets. We recommend building databases of annotated images in

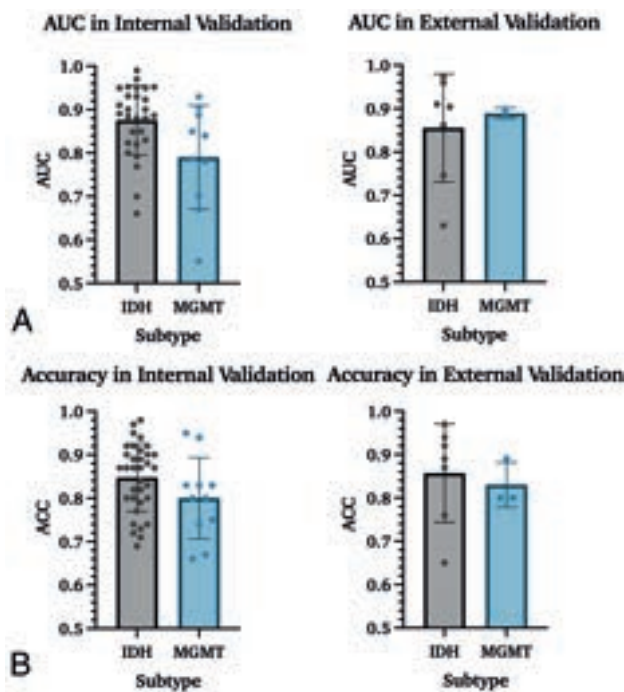


FIG 5. Performance measurements for the prediction of *IDH* and *MGMT* in internal and external data sets. **A.** AUC results for *IDH* mutation status and *MGMT* promoter methylation prediction. These data include all studies with *IDH* and/or *MGMT* prediction results, increasing the number of studies from 59 to 63. **B.** Accuracy results for *IDH* mutation and *MGMT* promoter methylation status.

individual hospitals and considering federated learning.^{36,37} These changes will result in less reliance on publicly available data sets that are highly curated and will provide a diversity of hospital training data that will overcome the low data set patient number and allow translation of algorithms between different hospitals.³⁸ Additionally, we recommend that journal editors require validation of data set results, despite the expected drop in AUC and accuracy. This is well-represented in the recent RSNA-MICCAIS AI challenge, which demonstrated that the highest prediction for *MGMT* methylation based on an unseen external validation set was an AUC of 0.62.³⁹⁻⁴² Our literature review demonstrates 17 articles predicting *MGMT* methylation with AUC results ranging from 0.55 to 0.93 (mean, 0.81) in internal and external validation sets.

Some of the reasons for the inability to replicate these results could be low patient numbers in the reported studies, lack of description of model validation, high bias within the articles with the potential for overtraining, differences in segmentations among data sets, and differences in methods of *MGMT* methylation assessment between different hospitals. Future studies should consider that other pathologic testing protocols for molecular characterization at different hospitals and tumor heterogeneity with sampling bias can be a source of error because it is currently used as the criterion standard in ML algorithm development and, therefore, introduces reference bias in the respective findings.

Limitations

A limitation of this review includes analysis of studies with incomplete data, which resulted in unequal data analysis for

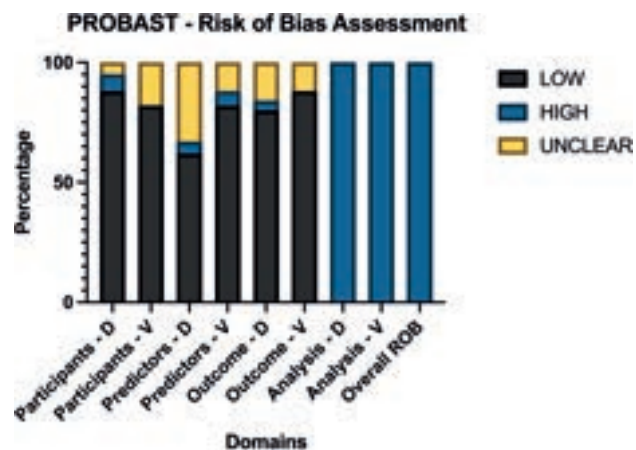


FIG 6. Average PROBAST scores of all 85 included articles. Domains were scored as high ROB if ≥ 1 item for each domain were scored as “No” or “Probably No.” PROBAST questions are assessed so that answering with no indicates ROB for this respective item. If 1 domain was considered a high ROB, the overall ROB of the study was considered high. As a result, each study was overall rated as having a high ROB. ROB indicates risk of bias; D, development studies; V, validation studies; D1, domain 1 participants; D2, domain 2 predictors; D3, domain 3 outcome; D4, domain 4 analysis.

criteria such as the accuracy of results. Additionally, the number of studies with external validation was low, limiting the generalizability of the findings and raising an important point for editors and authors to consider the need to include external validation data sets in their publications. Furthermore, publication bias was not assessed because it was considered beyond the scope of this review. Finally, we compared the best prediction pipeline each study had to offer, which might introduce bias in our findings despite our efforts to minimize bias in our analyses.

CONCLUSIONS

The results of prediction algorithms for molecular subtypes of gliomas in published studies demonstrate high AUC and accuracy. Still, there is an increased risk of bias based on the PROBAST assessment, which can result in poor data reproducibility. Improvement in reporting the quality of articles, development of large hospital-annotated data sets, and performance of external validation studies in future literature are critical for identifying algorithms that can be translated into clinical practice. This issue raises the need to develop novel tools for efficient data curation and annotation within the clinical workflow.

ACKNOWLEDGMENTS

The American Society of Neuroradiology Fellow Award 2018 was presented to M.A. This publication was made possible by KL2 TR001862 (M.A.) from the National Center for Advancing Translational Science, components of the National Institutes of Health, and the National Institutes of Health Roadmap for Medical Research. A.A. is a PhD student in the Investigative Medicine Program at Yale, which is supported by Clinical and Translational Science Awards grant No. UL1 TR001863 from the National Center for Advancing Translational Science, a component of the National Institutes of Health. A.A. holds securities in

Hyperfine. M.L. is an employee and stockholder of Visage Imaging Inc and, unrelated to this work, receives funding from National Institutes of Health/National Cancer Institute R01 CA206180 and is a board member of Tau Beta Pi engineering honor society. Michael Sabel is a consultant for Novocure and Codman.

The content of this article is solely the authors' responsibility and does not necessarily represent the official view of National Institutes of Health.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

- Norden AD, Wen PY. Glioma therapy in adults. *Neurologist* 2006;12:279–92 CrossRef Medline
- Johns Hopkins Medicine. Gliomas. June 27, 2022. <https://www.hopkinsmedicine.org/health/conditions-and-diseases/gliomas>. Accessed June 27, 2022
- 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
- Parker NR, Khong P, Parkinson JF, et al. Molecular heterogeneity in glioblastoma: potential clinical implications. *Front Oncol* 2015;5:55 CrossRef Medline
- Friedmann-Morvinski D. Glioblastoma heterogeneity and cancer cell plasticity. *Crit Rev Oncog* 2014;19:327–36 CrossRef Medline
- Olar A, Aldape KD. Using the molecular classification of glioblastoma to inform personalized treatment. *J Pathol* 2014;232:165–77 CrossRef Medline
- Tillmanns N, Lum AE, Cassinelli S, et al. Identifying clinically applicable machine learning algorithms for glioma segmentation: recent advances and discoveries. *Neurooncol Adv* 2022;4:vdac093 CrossRef Medline
- Bakas S, Reyes M, Jakab A, et al. Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation: Progression Assessment, and Overall Survival Prediction in the BRATS Challenge. 2019. arxiv.org/abs/1811.02629. Accessed June 27, 2022
- Jekel L, Brim WR, von Reppert M, et al. Machine learning applications for differentiation of glioma from brain metastasis: a systematic review. *Cancers(Basel)* 2022;14:1369 CrossRef Medline
- Cassinelli Petersen GI, Shatalov J, Verma T, et al. Machine learning in differentiating gliomas from primary CNS lymphomas: a systematic review, reporting quality, and risk of bias assessment. *AJNR Am J Neuroradiol* 2022;43:526–33 CrossRef Medline
- Bahar RC, Merkaj S, Cassinelli Petersen GI, et al. Machine learning models for classifying high- and low-grade gliomas: a systematic review and quality of reporting analysis. *Front Oncol* 2022;12:856231 CrossRef Medline
- Sanghani P, Ang BT, King NK, et al. Overall survival prediction in glioblastoma multiforme patients from volumetric, shape and texture features using machine learning. *Surg Oncol* 2018;27:709–14 CrossRef Medline
- Bakas S, Shukla G, Akbari H, et al. Overall survival prediction in glioblastoma patients using structural magnetic resonance imaging (MRI): advanced radiomic features may compensate for lack of advanced MRI modalities. *J Med Imaging (Bellingham)* 2020;7:031505 CrossRef Medline
- Clark K, Vendt B, Smith K, et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging* 2013;26:1045–57 CrossRef Medline
- Heus P, Damen J, Pajouheshnia R, et al. Uniformity in measuring adherence to reporting guidelines: the example of TRIPOD for assessing completeness of reporting of prediction model studies. *BMJ Open* 2019;9:e025611 CrossRef Medline
- Mongan J, Moy L, Kahn CE Jr. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): a guide for authors and reviewers. *Radiol Artif Intell* 2020;2:e200029 CrossRef Medline
- Moons KG, Wolff RF, Riley RD, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med* 2019;170:W1–33 CrossRef Medline
- Li YM, Suki D, Hess K, et al. The influence of maximum safe resection of glioblastoma on survival in 1229 patients: can we do better than gross-total resection? *J Neurosurg* 2016;124:977–88 CrossRef Medline
- Hegi ME, Diserens AC, Gorlia T, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med* 2005;352:997–1003 CrossRef Medline
- Fukuma R, Yanagisawa T, Kinoshita M, et al. Prediction of IDH and TERT promoter mutations in low-grade glioma from magnetic resonance images using a convolutional neural network. *Sci Rep* 2019;9:20311 CrossRef Medline
- Ren Y, Zhang X, Rui W, et al. Noninvasive prediction of IDH1 mutation and ATRX expression loss in low-grade gliomas using multiparametric MR radiomic features. *J Magn Reson Imaging* 2019;49:808–17 CrossRef Medline
- Zhang X, Tian Q, Wang L, et al. Radiomics strategy for molecular subtype stratification of lower-grade glioma: detecting IDH and TP53 mutations based on multimodal MRI. *J Magn Reson Imaging* 2018;48:916–26 CrossRef Medline
- Gao M, Huang S, Pan X, et al. Machine learning-based radiomics predicting tumor grades and expression of multiple pathologic biomarkers in gliomas. *Front Oncol* 2020;10:1676 CrossRef Medline
- Lu CF, Hsu FT, Hsieh KL, et al. Machine learning-based radiomics for molecular subtyping of gliomas. *Clin Cancer Res* 2018;24:4429–36 CrossRef Medline
- Chen X, Zeng M, Tong Y, et al. Automatic prediction of MGMT status in glioblastoma via deep learning-based MR image analysis. *Biomed Res Int* 2020;2020:9258649 CrossRef Medline
- Li Y, Wei D, Liu X, et al. Molecular subtyping of diffuse gliomas using magnetic resonance imaging: comparison and correlation between radiomics and deep learning. *Eur Radiol* 2022;32:747–58 CrossRef Medline
- Hedehzadeh M, Maghooli K, MomenGharibvand M, et al. A comparison of the efficiency of using a deep CNN approach with other common regression methods for the prediction of EGFR expression in glioblastoma patients. *J Digit Imaging* 2020;33:391–98 CrossRef Medline
- Akbari H, Bakas S, Pisapia JM, et al. In vivo evaluation of EGFRvIII mutation in primary glioblastoma patients via complex multiparametric MRI signature. *Neuro Oncol* 2018;20:1068–79 CrossRef Medline
- Park JE, Kim HS, Park SY, et al. Prediction of core signaling pathway by using diffusion- and perfusion-based MRI radiomics and next-generation sequencing in isocitrate dehydrogenase wild-type glioblastoma. *Radiology* 2020;294:388–97 CrossRef Medline
- Sun Z, Li Y, Wang Y, et al. Radiogenomic analysis of vascular endothelial growth factor in patients with diffuse gliomas. *Cancer Imaging* 2019;19:68 CrossRef Medline
- Buda M, AlBadawy EA, Saha A, et al. Deep radiogenomics of lower-grade gliomas: convolutional neural networks predict tumor genomic subtypes using MR images. *Radiol Artif Intell* 2020;2:e180050 CrossRef Medline
- Li Y, Liang Y, Sun Z, et al. Radiogenomic analysis of PTEN mutation in glioblastoma using preoperative multi-parametric magnetic resonance imaging. *Neuroradiology* 2019;61:1229–37 CrossRef Medline
- Wagner MW, Hainc N, Khalvati F, et al. Radiomics of pediatric low-grade gliomas: toward a pretherapeutic differentiation of BRAF-mutated and BRAF-fused tumors. *AJNR Am J Neuroradiol* 2021;42:759–65 CrossRef Medline
- Jian A, Jang K, Manuguerra M, et al. Machine learning for the prediction of molecular markers in glioma on magnetic resonance

- imaging: a systematic review and meta-analysis. *Neurosurgery* 2021;89:31–44 CrossRef Medline
35. Cabitza F, Campagner A, Soares F, et al. **The importance of being external. Methodological insights for the external validation of machine learning models in medicine.** *Comput Methods Programs Biomed* 2021;208:106288 CrossRef Medline
 36. Rieke N, Hancox J, Li W, et al. **The future of digital health with federated learning.** *NPJ Digit Med* 2020;3:119 CrossRef Medline
 37. Sheller MJ, Edwards B, Reina GA, et al. **Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data.** *Sci Rep* 2020;10:12598 CrossRef Medline
 38. Pfitzner B, Steckhan N, Arnrich B. **Federated learning in a medical context: a systematic literature review.** *ACM Trans Internet Technol* 2021;21:1–31 CrossRef
 39. Baid U, Ghodasara S, Mohan S, et al. **The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification.** *Computer Vision and Pattern Recognition*. 2021. <https://arxiv.org/abs/2107.02314>. Accessed June 27, 2022
 40. Menze BH, Jakab A, Bauer S, et al. **The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS).** *IEEE Trans Med Imaging* 2015;34:1993–2024 CrossRef Medline
 41. Bakas S, Akbari H, Sotiras A, et al. **Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features.** *Sci Data* 2017;4:170117 CrossRef Medline
 42. **RSNA-MICCAI Brain Tumor Radiogenomic Classification.** July 13, 2021 to October 15, 2021. <https://www.kaggle.com/c/rsna-miccai-brain-tumor-radiogenomic-classification>. Accessed June 27, 2022

Segmentation of Brain Metastases Using Background Layer Statistics (BLAST)

Chris Heyn, Alan R. Moody, Chia-Lin Tseng, Erin Wong, Tony Kang, Anish Kapadia, Peter Howard, Pejman Maralani, Sean Symons, Maged Goubran, Anne Martel, Hanbo Chen, Sten Myrehaug, Jay Detsky, Arjun Sahgal, and Hany Soliman



ABSTRACT

BACKGROUND AND PURPOSE: Accurate segmentation of brain metastases is important for treatment planning and evaluating response. The aim of this study was to assess the performance of a semiautomated algorithm for brain metastases segmentation using Background Layer Statistics (BLAST).

MATERIALS AND METHODS: Nineteen patients with 48 parenchymal and dural brain metastases were included. Segmentation was performed by 4 neuroradiologists and 1 radiation oncologist. K-means clustering was used to identify normal gray and white matter (background layer) in a 2D parameter space of signal intensities from postcontrast T2 FLAIR and T1 MPRAGE sequences. The background layer was subtracted and operator-defined thresholds were applied in parameter space to segment brain metastases. The remaining voxels were back-projected to visualize segmentations in image space and evaluated by the operators. Segmentation performance was measured by calculating the Dice-Sørensen coefficient and Hausdorff distance using ground truth segmentations made by the investigators. Contours derived from the segmentations were evaluated for clinical acceptance using a 5-point Likert scale.

RESULTS: The median Dice-Sørensen coefficient was 0.82 for all brain metastases and 0.9 for brain metastases of ≥ 10 mm. The median Hausdorff distance was 1.4 mm. Excellent interreader agreement for brain metastases volumes was found with an intraclass correlation coefficient = 0.9978. The median segmentation time was 2.8 minutes/metastasis. Forty-five contours (94%) had a Likert score of 4 or 5, indicating that the contours were acceptable for treatment, requiring no changes or minor edits.

CONCLUSIONS: We show accurate and reproducible segmentation of brain metastases using BLAST and demonstrate its potential as a tool for radiation planning and evaluating treatment response.

ABBREVIATIONS: BL = background layer; BLAST = Background Layer Statistics; BM = brain metastases; DL = deep learning; DSC = Dice-Sørensen coefficient; HD = Hausdorff distance; ICC = intraclass correlation coefficient; IQR = interquartile range; SRS = stereotactic radiosurgery; TH = threshold

Brain metastases (BM) are diagnosed in up to 40% of patients with metastatic cancer and usually imply a short survival.¹ However, recent advances in the treatment of BM, with, for

example, stereotactic radiosurgery (SRS) have led to improvements in patient outcomes with less impact on neurocognition and quality of life.² A requirement for SRS is accurate detection and contouring of BM. Additionally, treatment of BM requires measurements of tumor burden at baseline and follow-up to assess treatment response. To aid in this requirement, accurate segmentation of the tumor is needed to provide precise lesion targeting and to monitor changes in the size of the metastases between baseline and follow-up scans.

During the past few years, advances in machine learning methods have led to improvements in automated and semiautomated brain tumor segmentation. Machine learning methods can be grouped into supervised and unsupervised algorithms. Supervised methods, such as those based on deep convolutional neural networks, have recently garnered attention, showing excellent performance in brain tumor segmentation tasks,³⁻⁵ with 1 algorithm now FDA-cleared.⁶ These supervised methods, however, require large numbers of (manual) labels for training,

Received May 24, 2023; accepted after revision August 16.

From the Department of Medical Imaging (C.H., A.R.M., E.W., T.K., A.K., P.H., P.M., S.S.), Sunnybrook Research Institute (C.H., A.R.M., M.G., A.M.), and Department of Radiation Oncology (C.-L.T., H.C., S.M., J.D., A.S., H.S.), Sunnybrook Health Sciences Center, Toronto, Ontario, Canada; and Department of Medical Biophysics (M.G., A.M.), University of Toronto, Toronto, Ontario, Canada.

C. Heyn is supported by a New Investigator Research Grant from the SickKids Foundation and the Canadian Institutes of Health Research.

C. Heyn, H. Soliman, and A.R. Moody are coinventors on a provisional patent assigned to Sunnybrook Research Institute (US Provisional Patent No. 63/537,092) for the methodology described in this article.

Please address correspondence to Chris Heyn, MD, PhD, Department of Medical Imaging, Sunnybrook Health Sciences Center, 2075 Bayview Ave, Toronto, ON, Canada M4N 3M5; e-mail: chris.heyn@utoronto.ca

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

Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A7998>

which is a time-consuming and costly process and can be prone to bias introduced by the training set. Additionally, optimal performance of deep learning (DL) algorithms across multiple institutions commonly requires retraining with additional site-specific data (distributions).⁷

Unsupervised techniques do not require a priori training and can be used to facilitate the creation of ground truth data, which can be used to train DL models. Unsupervised techniques using clustering methods (eg, K-means, fuzzy c-means, and the expectation-maximization method) are iterative algorithms that segment by grouping voxels with similar signal properties (intensities), then estimating and optimizing cluster properties.⁸ K-means clustering groups signal intensity data into k classes by iteratively computing a mean intensity for each class and clustering voxels into the closest class centroid. Brain tumor segmentation with these techniques can be challenging because segmentation performance is highly dependent on initial conditions and the algorithm used.⁹ Furthermore, the signal heterogeneity and the small size of BM relative to background brain are additional challenges for clustering algorithms, to accurately identify and classify these tumors.

In this article, we describe an alternative semiautomated method for segmentation of BM using multiparametric MR imaging. The methodology first establishes a parameter space with origin and axes defined by the signal intensity statistics of background brain (Background Layer Statistics [BLAST]). In the present implementation of the methodology, K-means clustering is used to define the statistics of the background layer on a section of normal brain. Within the parameter space, voxels related to the background layer are then removed from the entire volume and additional operator-defined thresholds are applied to preferentially detect and segment BM. In this study, we evaluate BLAST methodology for segmentation of BM and hypothesize that it is accurate and reproducible.

MATERIALS AND METHODS

Subjects

Consecutive patients with newly diagnosed BM undergoing pretreatment MR imaging from July 1, 2022, to September 30, 2022, were included in this retrospective study. Research ethics board approval was obtained at our institution (Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada). Inclusion criteria were the following: 1) known biopsy-proved primary malignancy at the time of brain MR imaging, 2) 10 or fewer BM, and 3) no relevant treatment history at the time of brain MR imaging, including chemotherapy, radiation therapy, or prior brain surgery. Exclusion criteria were the following: 1) the presence of leptomeningeal disease or hemorrhagic metastasis, 2) the presence of another coexisting acute process such as acute stroke, or 3) severe corruption of MR imaging by motion artifacts.

MR Imaging Acquisition

All imaging was performed on 1.5T (Magnetom Aera or Sola; Siemens) or 3T (Magnetom Vida; Siemens) MR imaging systems using body-transmit and 20-channel head and neck receiver coils. Patients were scanned with the institutional brain tumor imaging protocol including axial RESOLVE DWI (Siemens) (b -values = 0

and 1000 s/mm², TR = 3650–8010 ms, TE = 67.2–72.2 ms, in-plane resolution = 0.54×0.54 mm² to 1.25×1.25 mm², section thickness = 5 mm), axial T2 FLAIR postgadolinium (TR = 9000 ms, TE = 80–108 ms, TI = 2500 ms, in-plane resolution = 0.75×0.75 mm² to 0.83×0.83 mm², section thickness = 3 or 5 mm), and 3D T1 MPRAGE postgadolinium (TR = 1800 or 2240 ms, TE = 2.4 or 3 ms, flip angle = 8°, resolution = $1 \times 1 \times 1$ mm³).

Image Processing and Analysis

T2 FLAIR and trace DWI (b =1000 s/mm²) images were registered to 3D T1 MPRAGE using BRAINSFit (3D Slicer; <https://www.slicer.org/>)¹⁰ with a 6- df rigid registration and linear interpolation to 1-mm isotropic resolution. Brain extraction was performed using a custom-written script in Matlab and the Image Processing Toolbox, Release 2022a and 2023a (MathWorks) which uses a brain mask derived from a binary threshold of the coregistered and interpolated trace DWI (b =1000 s/mm²) data set. Brain-extracted T2 FLAIR and 3D T1 MPRAGE data were bias-field corrected using the N4ITK algorithm (3D Slicer)¹⁰ and saved in a NIfTI format.

Ground truth segmentations of BM were manually acquired on the 3D T1 MPRAGE images (ITK-SNAP; www.itksnap.org)¹¹ by C.H., a neuroradiologist with 8 years of postfellowship experience. All ground truth segmentations were reviewed and edited by H.S., a neuroradiation oncologist with 13 years of postfellowship experience. The ROI was drawn around each metastasis encompassing the entire enhancing portion of the brain metastasis as well as central areas of necrosis and excluding edema, generating a binary mask (for a 1-class segmentation task). For this study, enhancing was defined as having a qualitative signal intensity above the surrounding brain parenchyma on the postgadolinium T1 MPRAGE sequence.

BLAST

Figure 1 illustrates the key concepts in defining the BLAST parameter space and performing segmentations of enhancing tumor. An example of segmentation of enhancing tumor with vasogenic edema is shown in the Online Supplemental Data. The algorithm was implemented using a custom-written script in Matlab. Signal intensities of the skull-stripped, coregistered, and bias-corrected T1 MPRAGE and T2 FLAIR images were first rescaled from 0 to 1. On a section of normal brain, K-means clustering (2 clusters) of T1 MPRAGE and T2 FLAIR signal intensities was used to define the cluster corresponding to normal gray and white matter (background layer). For this study, the normal section (without any evidence of enhancing tumor or edema) was selected beforehand by the investigators. The centroid of this background layer cluster was then used to define the origin of the parameter space. The rescaled signal intensities in T1 MPRAGE and T2 FLAIR were divided by the SD of the background layer to redefine the position of data points in this parameter space as a z score.

Once signal intensities were put into the BLAST parameter space, segmentations were performed by subtracting the background layer and applying additional thresholds in T1 MPRAGE and T2 FLAIR to exclude other nontumor tissues such as blood vessels and dura. Background layer subtraction

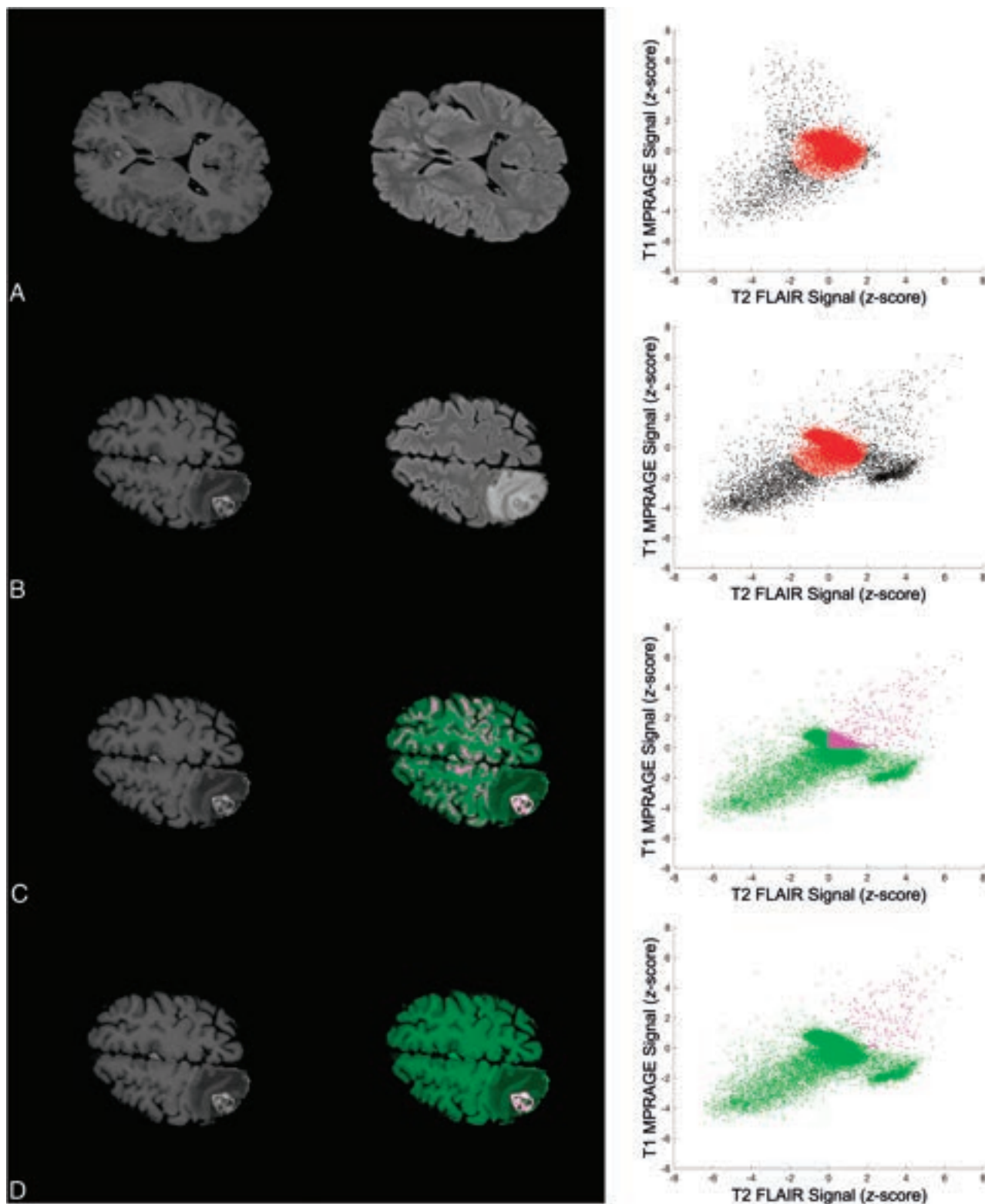


FIG 1. Key steps in defining the BLAST parameter space and performing segmentations with BLAST. Images and segmentations are shown in the left and middle columns, and the corresponding parameter space is shown in the right column. Each point in the scatterplot represents a voxel in the image. A, The statistics of the background layer (gray and white matter) are first defined using K-means clustering on a normal section of the brain. The origin of the parameter space is defined by the centroid of the background layer cluster and the axes expressed as z score relative to the background layer cluster SD. B, An ellipsoid approximating the background layer is applied to all slices (*red ellipse*). Voxels corresponding to enhancing brain tumor, edema, and CSF are found in the upper right, lower right, and lower left corner of parameter space, respectively. C, Thresholds in both T1 MPAGE and T2 FLAIR are applied to exclude nontumoral voxels to perform segmentations. In the sample shown, thresholds in both parameters are set to the mean of the background layer (z score = 0). D, Voxels falling within the background layer ellipsoid are subtracted to further exclude the background layer. The segmentation is selected to save a mask of the 3D-connected enhancing BM voxels. An algorithm to fill in necrotic areas is applied for the final segmentation (not shown).

was achieved by approximating the background layer voxels using an ellipsoid with the center set to the origin of the BLAST parameter space. The remaining voxels in parameter space were projected back into the image space to visualize the resulting segmentation of enhancing tumor. The 3D object consisting of connected enhancing voxels was selected, and an algorithm to “fill in” nonenhancing or necrotic regions was applied to finalize the segmentation. This filling was completed using morphologic dilations and erosion with structuring elements. To produce an outline of the brain metastasis, we applied a line contour algorithm to encompass the segmentation.

Ablation Study

An ablation study was performed to evaluate the contribution of different steps to segmentation performance. The complete BLAST method consisted of application of thresholds in 2D (T1 MPRAGE and T2 FLAIR) and subtraction of the background layer approximated by an ellipsoid with the semiaxis length set to 1.5 (threshold $[TH] \times 2 \text{ BL} = 1.5$). Algorithms consisting of application of thresholds in 2D without background layer subtraction ($TH \times 2 \text{ background layer [BL]} = 0$) and an algorithm consisting of a threshold in 1D (T1 MPRAGE) without background layer subtraction ($TH \times 1 \text{ BL} = 0$) were also evaluated. For $TH \times 2 \text{ BL} = 1.5$ and $TH \times 2 \text{ BL} = 0$, the thresholds in T1 MPRAGE were incremented by 0.5 from the z score = -1 up to z score = 3, and those in T2 FLAIR were incremented by 0.5 from the z score = -2 up to z score = 3. For $TH \times 1 \text{ BL} = 0$, the threshold in T1 MPRAGE was iterated from the z score = -1 up to z score = 5 by increments of 0.5. For each iteration, the segmentation result was saved and the Dice-Sørensen coefficient (DSC) for the tumor volumes was calculated in Matlab using the manual segmentations as the ground truth.

Evaluation of Background Layer Subtraction on Segmentation Performance

To understand the effect of background layer subtraction on segmentation performance, we ran the algorithm with different amounts of background subtraction by varying the ellipsoid semiaxis length from 0 to 3 by increments of 0.5. This step was performed for thresholds in T1 MPRAGE and T2 FLAIR of 0, 0.5, and 1. For each combination, the segmentation result was evaluated by calculating the DSC.

Operator-Generated BLAST Segmentations and Contours

BM segmentation was performed by 5 operators (4 board-certified neuroradiologists with BLAST using background layer subtraction and thresholding in 2D [$TH \times 2 \text{ BL} = 1.5$]: E.W., T.K., A.K., and, P.H., with 1–15 years of postfellowship experience and 1 board-certified neuroradiation oncologist: C.-L.T with 8 years of postfellowship experience treating and contouring BM). All definitive BM were identified by the principal investigators, and the location was provided to the operators. Operators segmented all BM within a patient including dural-based metastases. Bone metastases were excluded because these were removed from the volume by the brain-extraction algorithm.

The size of the metastases was estimated as the length of the major axis of an ellipsoid with the same normalized second

central moment as the segmented 3D object. Volume was calculated from the number of voxels comprising the segmented 3D object. In addition to the DSC, the Hausdorff distance (HD) for the center section of a metastasis was calculated using Matlab.

To evaluate the clinical acceptance of the BLAST-generated contours, we used a 5-point Likert scale adapted from the MD Anderson Cancer Center.¹² Scoring was performed by an experienced neuroradiation oncologist (H.S.) using the contour from the operator with the highest DSC for the case. A score of 4 or 5 on the Likert scale, for example, indicates that the contour is acceptable for clinical practice, while a score of 2 indicates that the contour is completely unusable.

Statistical Analysis

Statistical analysis was performed using GraphPad Prism software, Version 9.5.1 for Mac OS (GraphPad Software). The Shapiro-Wilk test was used to test for normality. For nonparametric group comparison, the Mann-Whitney U test or Friedman test was performed, and the results were considered significant if the P value was less than .05. A linear regression of tumor volume measured by BLAST segmentation versus ground truth segmentation was performed. The difference in tumor volume between BLAST and ground truth segmentation was also measured with a Bland-Altman plot.

Interreader agreement of tumor volume measurements was evaluated by a 2-way random effects model intraclass correlation coefficients (ICC) in Matlab.¹³ ICC < 0.5 (poor), 0.5–0.75 (fair), 0.75–0.9 (moderate), and ≥ 0.9 (excellent) agreement. ICCs were reported with their 95% CIs.

RESULTS

A total of 19 patients with 48 BM met the inclusion and exclusion criteria (Table). The median volume of BM was 0.70 cm³ (interquartile range [IQR], 0.1–2.1 cm³) and the median diameter was 12.8 mm (IQR, 6.7–20.7) mm. Twenty-two BM measured <10 mm, and 6 BM measured <5 mm.

Results from the ablation experiment are included in the Online Supplemental Data. The median DSC scores for $TH \times 2 \text{ BL} = 1.5$, $TH \times 2 \text{ BL} = 0$, and $TH \times 1 \text{ BL} = 0$ were 0.9 (IQR, 0.87–0.92), 0.90 (IQR, 0.88–0.92), and 0.86 (IQR, 0.78–0.9), respectively. DSC scores for $TH \times 2 \text{ BL} = 0$ and $TH \times 2 \text{ BL} = 1.5$ were significantly higher than those for $TH \times 1 \text{ BL} = 0$ ($P < .001$). DSC was not significantly different between $TH \times 2 \text{ BL} = 0$ and $TH \times 2 \text{ BL} = 1.5$ ($P = .94$). An illustrative case in which $TH \times 1 \text{ BL} = 0$ failed to properly segment a brain metastasis abutting the adjacent tentorium is also shown.

Results of varying the levels of background layer subtraction on segmentation performance are included in the Online Supplemental Data. The experiment was run on 46 of the 48 BM. Two BM demonstrated low T2 FLAIR signal, and the signal intensities for the tumor fell below the minimum T2 FLAIR (z score = 0) used in the experiment. For a given threshold in T1 MPRAGE and T2 FLAIR, increasing background layer subtraction by varying the semiaxis length from 0 to 3 resulted in improving segmentation performance. The benefit of background layer subtraction is more pronounced for lower thresholds and reduces at higher thresholds in T1 MPRAGE and T2 FLAIR. Visually, background

Patient demographics and tumor characteristics

Parameter	
Demographics	
No. of patients	19
Average age (yr)	65.7 (SD, 14)
No. women	9 (47.3%)
Primary cancer type	
Lung	
NSCLC	9
SCLC	1
Breast	2
Melanoma	2
Esophagus	1
Gastric	1
Pancreas	1
Vagina	1
Nasopharynx	1
Metastasis information	
Total No.	
Parenchymal	38
Dural	10
Median No./patient	2 (IQR, 1–3.5)
Median size (mm)	21.8 (IQR, 6.7–20.7)
Median volume (cm ³)	0.70 (IQR, 0.10–2.1)

Note:—NSCLC indicates non-small cell lung cancer; SCLC, small cell lung cancer.

layer subtraction reduces the amount of background voxels included in the segmentation mask, allowing better visualization of metastases across the thresholds that were evaluated. While there is greater background layer removal at higher thresholds or with greater background layer subtraction, there is also reduction in the tumor segmentation.

Sample contours for a cross-section of metastases contoured by human operators are shown in Fig 2. Figure 3 shows the relationship between DSC versus metastasis size for the operators and compares this with the results from the ablation analysis (TH × 2 BL = 1.5). The median DSC for all BM and all operators was 0.82 (IQR, 0.73–0.9). For operators, there was a significant difference in DSC ($P < .001$) between metastases of <10 mm (0.70; IQR, 0.65–0.8) and those of ≥10 mm (0.90; IQR, 0.86–0.92). A statistically significant difference in the median DSC for metastases of <10 mm (0.90; IQR, 0.84–0.91) and those of ≥10 mm (0.91; IQR, 0.89–0.94) was not found for TH × 2 BL = 1.5 from the ablation experiment ($P = .05$). The median HD for all metastases and operators was 1.4 mm (IQR, 1–2 mm). The median HD was significantly different ($P < .001$), for metastases of <10 mm (1.4 mm; IQR, 1–1.4 mm) and those of ≥10 mm (1.8 mm; IQR, 1.1–2.2 mm).

Linear regression of tumor volume measured by human operators versus ground truth volume showed excellent fit with $R^2 = 0.9951$ (Fig 4). Bland-Altman analysis showed a bias of 0.14 cm³ toward larger tumor volumes with BLAST compared with the ground truth (95% limits of agreement, −0.38–0.66). There was excellent interreader agreement for tumor volumes with ICC = 0.9978 (95% CI, 0.9967–0.9987). The median segmentation time using BLAST for all operators was 2.8 (IQR, 1.6–4.3) minutes/metastasis and was not influenced by tumor size.

Forty-five contours (94%) were scored a 4 or 5, indicating that the contours were acceptable clinically for treatment, requiring no changes at all or minor edits that are not thought to be

clinically relevant. Only 3 contours (6%) scored a 3, indicating that minor edits were needed.

DISCUSSION

The treatment of brain metastases with SRS requires accurate detection and segmentation, which can be time-consuming and challenging. Furthermore, the follow-up of brain metastases necessitates reliable measurements of tumor burden to evaluate treatment response, which is important in routine clinical practice and clinical trials. To this end, methodologies for accurate and rapid segmentation of brain tumors have the potential to greatly impact the treatment of BM by improving the accuracy of treatments and measuring the response.

In the present work, we show the results of a methodology that provides highly accurate and reproducible segmentation of BM using multiparametric MR images. The methodology is based on using the statistics of normal background brain to identify abnormal tissue. As opposed to using K-means clustering to detect and segment BM directly, K-means clustering is used to identify normal background brain voxels from a section of normal brain, allowing them to be excluded and resulting in segmentation of enhancing tumor. Conveniently, the thresholds used to exclude nontumoral voxels are set relative to the statistics of the centroid of the background brain cluster.

Brain lesion segmentation based on the detection of outlier voxels has been previously described by Seghier et al.¹⁴ In their methodology, a fuzzy clustering procedure was used to identify outlier voxels corresponding to brain pathology from gray and white matter using T1-weighted images alone. The methodology was subsequently adapted to detect and segment BM on postgadolinium T1-weighted sequences, but the performance of the detection and segmentation task was limited by false-positives, mainly from vascular structures (arteries and veins), the dura, and the choroid plexus.¹⁵ These methods require training on a set of normal brains to model the intensity distribution of tissue types. One of the advantages of BLAST is that no training is required to perform segmentations because the normal brain cluster is defined on the basis of the statistics from the subject. Additionally, another main difference between these methods and BLAST is the use of multiparametric data in BLAST to better separate the normal brain cluster corresponding to gray and white matter from outlier voxels corresponding to BM and other normal structures such as blood vessels or the dura. In particular, the use of a black-blood sequence such as T2 FLAIR in combination with postcontrast T1 MPRAGE provides excellent separation of contrast-enhancing blood vessels from tumors in parameter space, resulting in fewer false-positives. This finding was highlighted in the results of the ablation study, which showed the superior performance of thresholding in both T1 MPRAGE and T2 FLAIR compared with thresholding in T1 MPRAGE alone.

The ablation study found no significant effect of background layer subtraction on the overall segmentation performance using BLAST. This finding is probably the result of adequate separation of enhancing tumor from the background layer in parameter space for most of the BM in this study, allowing the use of higher thresholds in T1 MPRAGE and T2 FLAIR. An analysis of the effect of background layer subtraction, however, showed that

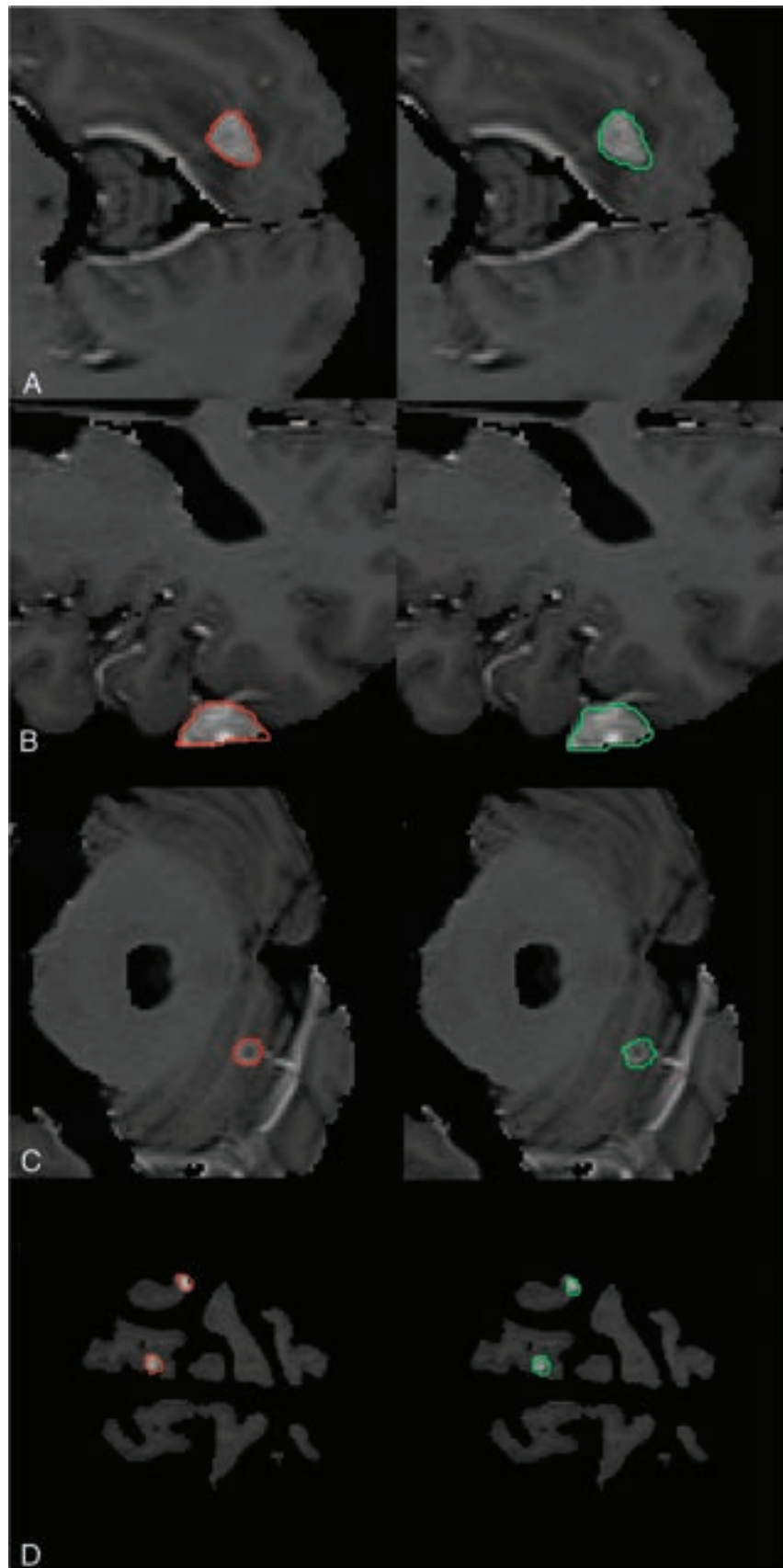


FIG 2. Sample contours derived from BLAST segmentations for parenchymal and dural metastases. The first column shows ground truth segmentations for 5 metastases (red). Corresponding contours created with BLAST are shown in the second column (green) for the operator with the best DSC. The mean DSCs for A, B, and C are 0.90, 0.83, 0.80, respectively. Two small metastases measuring 7 and 8 mm are shown in D with DSCs of 0.67 and 0.82, respectively.

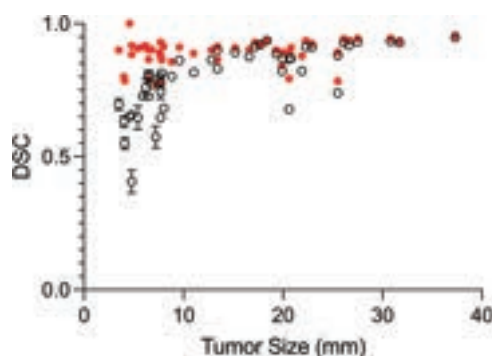


FIG 3. BLAST segmentation performance measured by DSC (mean) as a function of metastasis size for all operators (*black open circles*). Segmentation performance for BLAST generated by automatically iterating through combinations of thresholds with fixed background layer subtraction (*red closed circles*) outperforms operator-generated segmentations for small (<10 mm) metastases ($P < .001$).

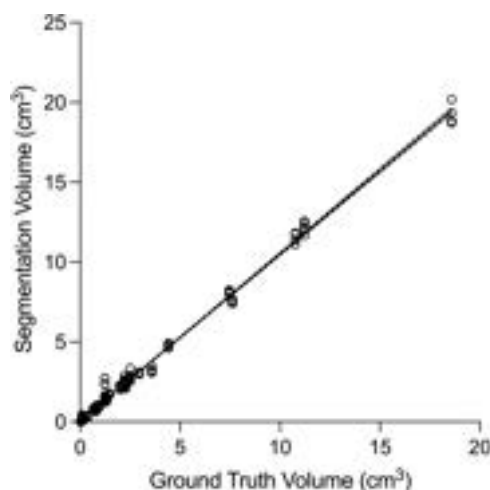


FIG 4. Comparison of tumor volumes measured by operators with BLAST versus ground truth. Linear regression of the data (*dashed lines* indicate 95% CI) shows an excellent fit ($R^2 = 0.9951$). The measured volumes closely approximate the ground truth volumes, with a slope close to unity (1.05; 95% CI, 1.04–1.06).

increasing levels of background layer subtraction improve segmentation performance at lower thresholds, which may be important for segmenting faintly enhancing metastases. Background layer subtraction also reduced the amount of background voxels included in the segmentation mask, possibly aiding users in the detection of metastases. Further work will be required to understand whether background layer subtraction provides any added benefit for segmenting or detecting faintly enhancing tumors or other pathologic processes whose signal characteristics overlap with normal brain (such as edema or nonenhancing tumor). Additionally, the optimal level of background subtraction for these applications will require further experimentation.

To evaluate the performance of segmentations using BLAST, we recruited human operators to use the methodology to generate contours of BM. Volumetric measurement of BM with the BLAST algorithm showed excellent interobserver agreement ($ICC = 0.9978$). Furthermore, tumor volumes measured with

BLAST showed excellent correlation with ground truth manual segmentation. The excellent interobserver agreement and accuracy are features of BLAST that could make this algorithm well-suited for assessing response to treatment.

The results of the human operator study demonstrated a poorer segmentation performance for smaller metastases compared with larger metastases. Most interesting, this difference did not exist when evaluating BLAST performed by iterating a combination of thresholds in T1 MPRAGE and T2 FLAIR with background layer subtraction. The poorer performance of BLAST by human operators for smaller metastases was, therefore, seen to be a limitation of the users and the interface of the current Matlab implementation rather than a limitation of the BLAST algorithm. We expect that this situation will be improved in future versions of the software, which will allow operators to better visualize smaller metastases and allow real-time evaluation of the segmentation as thresholds are manipulated by the operator.

While a head-to-head comparison of BLAST and DL was not performed, compared with recently published state-of-the-art DL methods, BLAST fared well. Using a 2.5D fully convolutional neural network based on the GoogLeNet architecture (<https://www.mathworks.com/help/deeplearning/ref/googlenet.html>) trained on multiparametric MR imaging including 3D T1 BRAVO (GE Healthcare) postgadolinium, 3D T1 CUBE (GE Healthcare) pre- and postgadolinium, and T2 FLAIR, brain metastasis segmentation with a mean DSC = 0.79 (SD, 0.12) has been reported.³ With a 3D U-Net trained on 3D T1 echo-spoiled gradient echo postgadolinium or subtraction images, segmentation of BM with a median DSC = 0.75 and a median HD = 1.5 mm has been achieved.⁴ By means of a self-adaptive nnU-Net model (<https://www.nature.com/articles/s41592-020-01008-z>) trained on 3D T1 postgadolinium images, excellent segmentation performance with an overall mean DSC = 0.822 (SD, 0.095) for all metastases in the test set (mean size = 12.3 [SD, 9.2] mm) and DSC = 0.868 (SD, 0.075) for metastases of ≥ 6 mm has been recently shown.⁵ Further work will be required to evaluate the performance of BLAST compared with DL algorithms for brain tumor segmentation.

The median segmentation time for BLAST (2.8 minutes/metastasis) was longer compared with most DL algorithms, which report whole-brain inference times ranging from 20 seconds to 5 minutes.^{3,4,16,17} Inference times, however, may not be a fair comparison because they do not take into consideration the quality or clinical usability of the segmentations. A recently FDA-approved algorithm based on a 3D U-net and DeepMedic (<https://github.com/deepmedic/deepmedic>) of volumetric MR imaging and CT data showed whole-brain inference times of 90 seconds but an average of 6.1 minutes/case for users to finalize segmentations of BM for treatment planning.¹⁸

The median segmentation time presented herein is primarily a function of operator manipulation of thresholds and re-evaluation of the resulting segmentation, which varied between operators (IQR, 1.6–4.3) minutes/metastasis. With the current Matlab implementation, the process of adjusting the threshold requires a re-computation of the brain metastasis mask after each change in threshold, which can be a lengthy process. In the future, this time can be shortened with improvements in the user interface.

The clinical acceptance of the BLAST-generated contours measured using a 5-point Likert score was excellent. Of the 3 contours that required editing, 2 were small metastases that were located superficially in the brain, and partial volume averaging effects were the cause of the poor segmentations. For the other contour, an adjacent blood vessel altered the shape and size of the contour, requiring minimal editing.

There are several limitations to the presented technique:

- 1) Hemorrhagic metastases were excluded from the study. These hemorrhagic BM have a propensity for lower T2 FLAIR signal within areas of hemorrhage, which overlap in signal intensity with adjacent dura or blood vessels. Segmentation of these metastases could be aided in the future by additional parameters such as pregadolinium T1 MPRAGE, which can better distinguish between the brain metastasis and T2 hypointense normal brain structures such as the dura or blood vessels.
- 2) Another limitation of the algorithm is how the necrotic core is segmented. In the present study, the algorithm fills in the necrotic core on the basis of the boundaries of surrounding enhancing tumor. In some cases, areas of necrosis could be excluded if the rim of enhancing tumor is very poorly enhancing or thin. In the future, these areas of necrosis could be segmented separately with the addition of other parameters, including pregadolinium T1 MPRAGE.
- 3) The present study uses postgadolinium T2 FLAIR, which is not a widespread practice. At our center, T2 FLAIR is routinely acquired postgadolinium administration to save table time and enable a longer delay between gadolinium administration and the acquisition of T1 MPRAGE. The acquisition of T2 FLAIR after gadolinium administration also enables the detection of subtle leptomeningeal enhancement, which can be missed on T1 sequences alone.¹⁹ It is possible that the use of postgadolinium T2 FLAIR could provide an advantage for segmenting enhancing BM because the signal enhancement on T2 FLAIR resulting from intratumoral gadolinium leakage could result in better separation of background brain and the metastasis in T2 FLAIR parameter space. While this article has not addressed this possibility, our experience with BLAST has shown the feasibility of BLAST with pregadolinium T2 FLAIR data as well.
- 4) The current study is a single-center validation study with a small number of patients scanned using MR imaging systems from a single vendor. To address the possibility that the sample size was too small to detect a difference between arms, we calculated that a total sample size of 21 brain metastases would have been required to have an 80% power to assess the equivalence of the contoured volumes between the experimental method and the criterion standard, assuming an equivalence limit of $\pm 20\%$ of the criterion standard volume and an α threshold of .05, assuming that the dispersion of the results in the study reflected the true population dispersion. A larger sample size will be required in future studies to reduce the equivalence limit and increase overall power. The general utility of BLAST in a larger multicenter study of BM will also be needed to evaluate the generalizability of our results.

Despite these limitations, the segmentations produced by BLAST were easy to generate and could be used for BM treatment-planning or response-evaluation. The algorithm could also be adapted for diagnostic metastasis detection. The BLAST algorithm was able to detect and distinguish BM adjacent to arteries and other enhancing structures like venous sinuses and the dura, which can be challenging to detect. While it is possible that the BLAST algorithm alone may be sufficient for segmenting most BM, it can likely have the greatest impact through using the generated segmentations to train new DL models on improved and larger ground truth data sets for metastasis detection and segmentation.

Finally, the general framework of BLAST is not limited to metastases. The algorithm could be used to segment primary brain tumors (including enhancing and nonenhancing tumor subregions) or other brain pathology (such as white matter disease for instance). It could also be used with other MR imaging pulse sequences and in combination with other modalities, including CT and PET.

CONCLUSIONS

We present here an alternative methodology for brain metastasis segmentation, which provides accurate and reproducible segmentations of both parenchymal and dural metastases without extensive a priori training. Combined with the relative simplicity of the algorithm, the methodology could be widely implemented for brain metastasis treatment and response assessment.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

1. Tsao MN, Lloyd N, Wong RK, et al. **Whole brain radiotherapy for the treatment of newly diagnosed multiple brain metastases.** *Cochrane Database Syst Rev* 2012;2012:CD003869 CrossRef Medline
2. Brown PD, Jaeckle K, Ballman KV, et al. **Effect of radiosurgery alone vs radiosurgery with whole brain radiation therapy on cognitive function in patients with 1 to 3 brain metastases: a randomized clinical trial.** *JAMA* 2016;316:401–09 CrossRef Medline
3. Grovik E, Yi D, Iv M, et al. **Deep learning enables automatic detection and segmentation of brain metastases on multisequence MRI.** *J Magn Reson Imaging* 2020;51:175–82 CrossRef Medline
4. Rudie JD, Weiss DA, Colby JB, et al. **Three-dimensional U-Net convolutional neural network for detection and segmentation of intracranial metastases.** *Radiol Artif Intell* 2021;3:e200204 CrossRef Medline
5. Ziyadeh H, Cardenas CE, Yeboa DN, et al. **Automated brain metastases segmentation with a deep dive into false-positive detection.** *Adv Radiat Oncol* 2023;8:101085 CrossRef Medline
6. Wang JY, Qu V, Hui C, et al. **Stratified assessment of an FDA-cleared deep learning algorithm for automated detection and contouring of metastatic brain tumors in stereotactic radiosurgery.** *Radiat Oncol* 2023;18:61 CrossRef Medline
7. Rauschecker AM, Gleason TJ, Nedelec P, et al. **Interinstitutional portability of a deep learning brain MRI lesion segmentation algorithm.** *Radiol Artif Intell* 2022;4:e200152 CrossRef Medline
8. Despotovic I, Goossens B, Philips W. **MRI segmentation of the human brain: challenges, methods, and applications.** *Comput Math Methods Med* 2015;2015:450341 CrossRef Medline
9. Velthuisen RP, Clarke LP, Phuphanich S, et al. **Unsupervised measurement of brain tumor volume on MR images.** *J Magn Reson Imaging* 1995;5:594–605 CrossRef Medline

10. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. **3D Slicer as an image computing platform for the Quantitative Imaging Network.** *Magn Reson Imaging* 2012;30:1323–41 CrossRef Medline
11. Yushkevich PA, Piven J, Hazlett HC, et al. **User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability.** *Neuroimage* 2006;31:1116–28 CrossRef Medline
12. Baroudi H, Brock KK, Cao W, et al. **Automated contouring and planning in radiation therapy: what is 'clinically acceptable'?** *Diagnostics (Basel)* 2023;13:667 CrossRef Medline
13. Salarian A. **Intraclass Correlation Coefficient (ICC).** 2023. (<https://www.mathworks.com/matlabcentral/fileexchange/22099-intraclass-correlation-coefficient-icc>). Accessed April 10, 2023
14. Seghier ML, Ramackhansingh A, Crinion J, et al. **Lesion identification using unified segmentation-normalisation models and fuzzy clustering.** *Neuroimage* 2008;41:1253–66 CrossRef Medline
15. Shearkhani O, Khademi A, Eilaghi A, et al. **Detection of volume-changing metastatic brain tumors on longitudinal MRI using a semiautomated algorithm based on the Jacobian operator field.** *AJNR Am J Neuroradiol* 2017;38:2059–66 CrossRef Medline
16. Bousabarah K, Ruge M, Brand JS, et al. **Deep convolutional neural networks for automated segmentation of brain metastases trained on clinical data.** *Radiat Oncol* 2020;15:87 CrossRef Medline
17. Hsu DG, Ballangrud A, Shamseddine A, et al. **Automatic segmentation of brain metastases using T1 magnetic resonance and computed tomography images.** *Phys Med Biol* 2021;66:175014 CrossRef Medline
18. Lu SL, Xiao FR, Cheng JC, et al. **Randomized multi-reader evaluation of automated detection and segmentation of brain tumors in stereotactic radiosurgery with deep neural networks.** *Neuro Oncol* 2021;23:1560–68 CrossRef Medline
19. Lee EK, Lee EJ, Kim S, et al. **Importance of contrast-enhanced fluid-attenuated inversion recovery magnetic resonance imaging in various intracranial pathologic conditions.** *Korean J Radiol* 2016;17:127–41 CrossRef Medline

Cortical and Subcortical Brain Atrophy Assessment Using Simple Measures on NCCT Compared with MRI in Acute Stroke

Tanaporn Jaroengarmsamer, Faysal Benali, Joachim Fladt, Nishita Singh, Fouzi Bala, Michael Tymianski, Michael D. Hill, Mayank Goyal, Aravind Ganesh,
On behalf of the ESCAPE-NA1 Investigators



ABSTRACT

BACKGROUND AND PURPOSE: Brain atrophy is an important surrogate for brain reserve, the capacity of the brain to cope with acquired injuries such as acute stroke. It is unclear how well atrophy measurements on MR imaging can be reproduced using NCCT imaging. We aimed to compare pragmatic atrophy measures on NCCT with MR imaging in patients with acute ischemic stroke.

MATERIALS AND METHODS: This is a post hoc analysis, including baseline NCCT and 24-hour follow-up MR imaging data from the Safety and Efficacy of Nerinetide (NA-1) in Subjects Undergoing Endovascular Thrombectomy for Stroke (ESCAPE-NA1) trial. Cortical atrophy was measured using the global cortical atrophy scale, and subcortical atrophy was measured using the intercaudate distance-to-inner-table width (CC/IT) ratio. Agreement and correlation between these measures on NCCT and MR imaging were calculated using the Gwet agreement coefficient 1 and Pearson correlation coefficients, respectively.

RESULTS: Among 1105 participants in the ESCAPE-NA1 trial, interpretable NCCT and 24-hour MR imaging were available in 558 (50.5%) patients (mean age, 67.2 [SD, 13.7] years; 282 women). Cortical atrophy assessments performed on NCCT underestimated atrophy severity compared with MR imaging (eg, patients with global cortical atrophy of ≥ 1 assessed on NCCT = 133/558 [23.8%] and on MR imaging = 247/558 [44.3%]; a 20.5% difference). Overall, cortical (ie, global cortical atrophy) atrophy assessments on NCCT had substantial or better agreement with MR imaging (Gwet agreement coefficient 1 of > 0.784 ; $P < .001$). Subcortical atrophy measures (CC/IT ratio) showed strong correlations between NCCT and MR imaging (Pearson correlation = 0.746, $P < .001$).

CONCLUSIONS: Brain atrophy can be evaluated using simple measures in emergently acquired NCCT. Subcortical atrophy assessments on NCCT show strong correlations with MR imaging. Although cortical atrophy assessments on NCCT are strongly correlated with MR imaging ratings, there is a general underestimation of atrophy severity on NCCT.

ABBREVIATIONS: AC1 = agreement coefficient 1; CC/IT = intercaudate distance-to-inner-table width; ESCAPE-NA1 = Safety and Efficacy of Nerinetide (NA-1) in Subjects Undergoing Endovascular Thrombectomy for Stroke; GCA = global cortical atrophy; h-ICD = hemi-intercaudate distance; MTA = medial temporal atrophy

Brain atrophy is considered an important imaging surrogate of brain reserve, the ability of the brain to cope with acquired tissue injuries, such as stroke, demyelination, or trauma.¹⁻⁷ There is compelling evidence on the role that atrophy and other

measures of brain reserve play in moderating functional recovery and neurocognitive sequelae after such brain injuries, leading to burgeoning interest in assessing these markers in research and practice.⁸⁻¹⁰ Several methods to assess brain atrophy on NCCT

Received October 18, 2022; accepted after revision August 3, 2023.

From the Department of Clinical Neurosciences (T.J., F. Benali, J.F., N.S., F. Bala, M.D.H., M.G., A.G.), University of Calgary, Calgary, Alberta, Canada; Faculty of Medicine Ramathibodi Hospital (T.J.), Mahidol University, Bangkok, Thailand; Department of Radiology and Nuclear Medicine (F. Benali), Maastricht University Medical Center+ (MUMC+), Maastricht, the Netherlands; Department of Neurology and Stroke Center (J.F.), University Hospital Basel, Basel, Switzerland; NoNO Inc (M.T.), Toronto, Ontario, Canada; Department of Radiology (M.D.H., M.G.), Department of Community Health Sciences (M.D.H.), and Hotchkiss Brain Institute and the Mathison Centre for Mental Health Research and Education (M.D.H., M.G., A.G.), University of Calgary, Calgary, Alberta, Canada; and Department of Medicine (M.D.H.), University of Calgary Cumming School of Medicine, Calgary, Alberta, Canada.

The first authorship is shared between Drs Tanaporn Jaroengarmsamer and Faysal Benali for equal contributions to conceptualization, data collection and analysis, and manuscript writing of this study.

The senior authorship is shared between Drs Mayank Goyal and Aravind Ganesh for equal contributions to the study design, supervision, analysis, and revision of the study manuscript.

Dr Jaroengarmsamer reported receiving a scholarship from the Prince Mahidol Award Youth Program. Dr Tymianski reported being chief executive officer of NoNO Inc, a company developing neuroprotectants for the treatment of acute ischemic stroke. Dr Hill reported receiving personal fees from Sun Pharma and Merck and receiving nonfinancial support from Hoffmann-La Roche Canada; holding a patent for US 10,916,346 licensed to Circle NVI and a patent for US 62,086,077 licensed to Circle NVI; owning stock in PureWeb Inc; serving as a director of the Canadian Federation of Neurological Sciences, the Canadian Stroke Consortium, and Circle NVI; and receiving grants from Alberta Innovates Health Solutions, the Canadian Institutes of Health Research, the Heart & Stroke Foundation of Canada, Covidien, Boehringer-Ingelheim, Biogen, Stryker, Medtronic, and the National Institutes of Neurological Disorders and Stroke. Dr Goyal reported receiving personal fees from Medtronic, Stryker, MicroVent, and Mentice during the management of the study; receiving unrestricted research grants to the University of Calgary from NoNO, Stryker, and Medtronic; holding a patent for systems of acute

have been developed, including sophisticated automated volumetrics.^{11,12} However, pragmatic visual ratings and measurements are currently the mainstay in clinical practice.^{13–15}

Such pragmatic scales include cortical atrophy assessments, which are based on the width of the sulci and volume of the gyri (global cortical atrophy [GCA]);¹³ hippocampal atrophy assessments, based on the height/volume of the hippocampus (medial temporal atrophy [MTA] scale)¹⁵ and parietal lobe atrophy assessments based on the width of the posterior cingulate and parieto-occipital sulci (Koedam scale).¹⁴ Subcortical atrophy is assessed less commonly but can be quantified using simple measurements and calculations like the intercaudate distance-to-inner-table width (CC/IT) ratio.¹⁶

MR imaging is the preferred imaging tool for the aforementioned scales and measures because they were originally developed as part of the work-up for neurodegenerative disorders.¹⁷ However, NCCT is more widely available and more often used in emergency settings like acute stroke. In addition, some patients cannot undergo MR imaging due to claustrophobia, excessive agitation, metal implants, or other contraindications.^{18,19} Therefore, should atrophy assessments on NCCT be comparable with those on MR imaging, this similarity would help facilitate the routine consideration of brain atrophy in the evaluation and prognostication of patients with acute neurologic injuries like stroke.

Prior studies comparing the use of NCCT and MR imaging for atrophy assessment included patients with neurodegenerative disorders rather than acute injuries like stroke^{12,20} and used less established scales or measurements.²¹ The few studies that included patients with acute stroke share the important limitation of not considering confounding of measurements by mass effect from infarct-related edema, which can complicate around 5% of all ischemic strokes.^{22,23}

In this study, we aimed to compare the pragmatic assessment of cortical and subcortical atrophy using well-known and standardized rating scales applied on NCCT versus MR imaging in a large randomized controlled trial–derived population of patients with acute stroke.

MATERIALS AND METHODS

Patient Sample

This study is a post hoc analysis of the Safety and Efficacy of Nerinetide (NA-1) in Subjects Undergoing Endovascular Thrombectomy for Stroke (ESCAPE-NA1) trial (clinicaltrials.gov: NCT02930018), assessing the efficacy of IV nerinetide in participants with acute ischemic stroke who underwent endovascular thrombectomy within 12 hours from onset.²⁴ Trial participants of ESCAPE-NA1 were enrolled between March 1, 2017, and August 12, 2019. Detailed information about the trial is provided in the Online Supplemental Data. We performed the

stroke diagnosis, licensed to GE Healthcare; and having an ownership interest in Circle NVI. Dr Ganesh reported receiving grants from the Wellcome Trust, the Canadian Institutes of Health Research, Canadian Cardiovascular Society, Alberta Innovates, Campus Alberta Neuroscience, and Sunnybrook Research Institute (INOVAIT network); receiving personal fees from MD Analytics, My Medical Panel, Creative Research Designs, Atheneum, DeepBench, Research on Mind, Figure 1, Alexion; receiving travel awards from American Academy of Neurology, American Heart Association, and University of Calgary; receiving cash awards from the Association of Indian Neurologists in America; holding stock options from SnapDx, Advanced Health Analytics, and The Rounds.com outside the submitted work; holding a patent for US 63/024,239 pending for a system for patient monitoring and

atrophy measurement between February 1, 2022, and July 31, 2022.

Image Acquisition

The imaging included in this study consisted of baseline NCCT and MR imaging at 24 hours after randomization. The 5-mm-thick, axial NCCT was acquired using a minimum power of 120–140 kV and 170–200 mA and a 2-second scanning time. The images were processed using appropriate algorithms to reduce bone artifacts and a high SNR for gray-white differentiation. Contiguous axial sections were obtained from the skull base to the vertex, parallel to the inferior orbitomeatal line. A good-quality NCCT scan was defined as having a well-discriminated lateral margin of the lentiform nucleus on the unaffected side and a well-defined insular ribbon in the absence of previous infarction. Reformats included 5-mm-thick sections of axial, sagittal, and coronal NCCT images and 25-mm axial and sagittal MIP CTA images, all generated from the source. CTA required at least 75 mL of contrast media. The scan was used only to confirm affected side for measuring hemi-intercaudate distance (h-ICD) on the contralateral side. A follow-up brain MR imaging was performed at a mean of 24 (SD, 12) hours after randomization. MR imaging sequences included a minimum of axial DWI, gradient-echo, and FLAIR. Reformation was used to generate axial, coronal, and sagittal images from the 3D FLAIR source. If only axial 2D FLAIR images were acquired, as was the case for some sites in this international trial, then available axial views were reviewed.

Image Analysis

Available baseline NCCT and 24-hour MR imaging data (flow chart, Fig 1) were reviewed by independent core laboratory readers (F. Benali, F. Bala, J.F., N.S., with 4–7 years of experience; Online Supplemental Data) who were blinded to clinical and outcome data. Disagreements were resolved by a senior reader (A.G., with 10 years of experience). The readers were trained to measure atrophy and subcortical atrophy using 20 cases from the data set. However, these cases were re-read for this study after a 30-day washout period, mixed in with all the other data set cases with no indicators of being from the training set. The readers then randomly assessed either NCCT or MR imaging, with no reader assessing both scans from the same patient in the first pass. An additional intrarater intermodality analysis was conducted to assess whether the same rater read NCCT and MR imaging differently.²⁵ Two readers (F. Benali and J.F.) each assessed global cortical atrophy (applying the GCA scale) in a random set of 100 patients using NCCT and MR imaging, with a washout period of 30 days between the NCCT and MR imaging reading sessions to mitigate the likelihood of being biased by prior measurements. We selected a sample size of 100 patients because this would give

delivery of remote ischemic conditioning or other cuff-based therapies; and serving on the editorial boards of Neurology; Neurology: Clinical Practice; and Stroke.

Please address correspondence to Tanaporn Jaroengarmsamer, MD, 270 Rama VI Rd, Thung Phaya Thai, Ratchathewi, Bangkok 10400, Thailand; e-mail: tanaporn.jn@gmail.com

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

Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A7981>

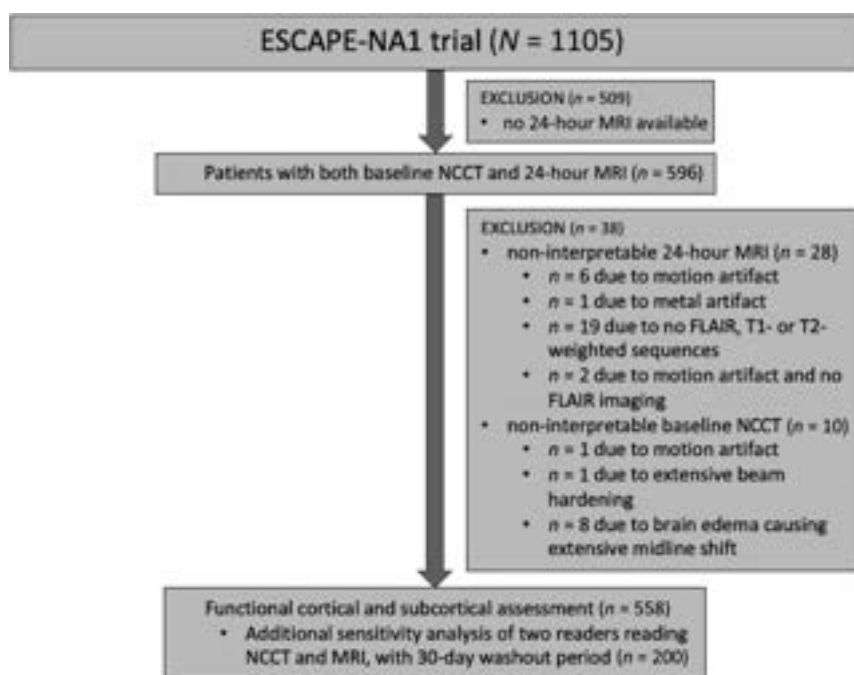


FIG 1. Flow diagram showing inclusion and exclusion of cases.

us 80% power for identifying a disagreement rate of 10% at an α of .001.²⁶

Cortical Atrophy

Cortical atrophy was measured using visual rating scales for the GCA (Online Supplemental Data). Additional sensitivity analyses included regional atrophy assessments, assessed in both hemispheres by applying the aforementioned scale to the frontal, temporal, and occipital lobes. The parietal lobe was assessed using the Koedam scale (Online Supplemental Data),¹⁴ and the hippocampus was assessed using the MTA scale (Online Supplemental Data).¹⁵ If the atrophy was asymmetric in any of the mentioned scales, the most severe side was included.

Subcortical Atrophy

Subcortical atrophy was assessed by calculating the ratio of the intercaudate distance-to-the-inner-table width, the CC/IT ratio.¹⁶ To avoid measurement error due to potential infarct-associated edema, the intercaudate distance was measured on the contralateral hemisphere (ie, h-ICD) and then multiplied by 2 (Online Supplemental Data). The h-ICD was defined as the minimum distance between the caudate head and the septum pellucidum at the level of the foramen of Monro and has been used previously^{22,23} to assess subcortical atrophy in stroke. In case of very severe edema and midline shift, participants were excluded from all analyses.

Statistical Analyses

Baseline variables of the included participants were summarized using mean (SD) or median and interquartile range for numeric data; and numbers (No.) and percentages (%) for categoric and dichotomous variables.

For correlation analyses regarding cortical atrophy assessment in the overall data set, we analyzed the intermodality agreement

between NCCT and MR imaging ratings. When we examined the distribution of our data, most cortical atrophy data fell into the “no atrophy” and “mild atrophy” categories. With such a data distribution, use of conventional methods for interrater reliability like the Cohen κ can result in misleadingly low κ values, even in the setting of good agreement, also known as the κ paradox.²⁷ The Gwet agreement coefficient 1 (AC1)²⁸ is particularly useful in this setting because it provides a more stable coefficient with varying prevalence or marginal probability.²⁹ We used a linear-weighted Gwet AC1 for our agreement calculations. The Gwet coefficient was also used for assessment of intermodality agreement for the random subset of 200 cases for which both NCCT and MR imaging were read by the same readers (100 cases each).

Although Gwet did not describe the level of agreement with categorized AC1

results, Landis and Koch’s level of agreement was used to interpret AC1.^{28,30} Such categorizations have been used in discussing the results of prior studies using the Gwet AC1.³⁰ An AC1 of 0.21–0.40 was considered fair agreement, 0.41–0.60 was considered moderate, 0.61–0.80 was considered substantial, and 0.81–1.00 was considered almost perfect.³¹ We performed agreement analyses for binary assessment of atrophy: no atrophy (atrophy grade 0) versus any atrophy (atrophy grade 1 or more) (Online Supplemental Data).

For correlation analysis in the measurement of subcortical atrophy between NCCT and MR imaging, we used Pearson correlation analyses. A Pearson correlation of 0.1–0.39 was considered as weak, a coefficient of 0.4–0.69 was considered moderate, and a coefficient of 0.7–0.89 was considered strong.³² Scatterplots were first drawn to visualize any outliers or linear correlations. In addition, we used Bland-Altman plots to visualize the agreements between NCCT and MR imaging. The patterns of the plots help interpret the correlation between NCCT and MR imaging beyond numeric data.^{33,34} For example, a Bland-Altman plot with concentrated data reveals that the both sets of data are highly correlated, whereas the plot with parallel data shows that the difference among data are constant. A narrow difference in the former plot means that the concentration of data is high and the bias between measurements is low. One of the radiologic indicators of early surgical decompression in patients with an ischemic brain is the bicaudate ratio of <0.16 .³⁵ We performed Gwet AC1 analysis for binary comparison of the CC/IT ratio (<0.16 and ≥ 0.16). All P values $\leq .05$ were considered significant, and analyses were performed using STATA/MP 16.1 (StataCorp).

RESULTS

Among the 1105 patients enrolled in the ESCAPE-NA1 trial, 1102 (99.7%) had interpretable NCCT and 568 (51.4%) had interpretable 24-hour MR imaging. Interpretable baseline NCCT and 24-

hour MR imaging were available in 566 (51.2%) patients. We excluded 8 cases for which subcortical atrophy was not assessable due to severe brain edema (Fig 1). Of the remaining 558 cases, 282 were women, with a mean age of 67.2 (SD, 13.7) years (Table 1). Multiplanar reformation of FLAIR images was available in 399 (71.5%) patients.

Cortical Atrophy

When using MR imaging, the rate of participants with any atrophy (GCA ≥ 1) was higher compared with when GCA assessments were performed using NCCT (ie, 245/558 [43.9%] and 133/558 [23.8%] for MR imaging and NCCT, respectively). This finding was also true for other cortical atrophy scales, such as the Koedam scale (ie, Koedam ≥ 1 with NCCT = 172/558 [30.8%] and with MR imaging = 261/558 [46.8%]) and the MTA scale (ie, MTA of ≥ 1 with NCCT = 123/558 [22.0%] and with MR imaging = 213/558 [38.2%]). The agreement values of binary (no versus any atrophy) assessments (Gwet AC1 = 0.500–0.554) were lower than the values of standard assessments (Gwet AC1 = 0.784–0.833). GCA assessment showed almost perfect agreement between NCCT and MR imaging (Gwet AC1 = 0.815; $P < .001$) (Table 2). Regional cortical atrophy assessments showed substantial-to-almost-perfect agreement (Gwet AC1 > 0.784 , $P < .001$) (Online Supplemental Data). Additional intrarater intermodality analyses of the GCA (for cases read by the same person) showed moderate-to-substantial agreement (Gwet AC1 = 0.775 and 0.547; $P < .001$).

Subcortical Atrophy

Subcortical atrophy assessments (CC/IT ratios) were similar when using NCCT compared with MR imaging (CC/IT ratio of ≥ 0.16

with NCCT = 87 [15.6%] and with MR imaging = 93 [16.7%]) with a Pearson correlation coefficient of 0.746 ($P < .001$). Bland-Altman plots depicting NCCT assessments compared with MR imaging showed a low difference in highly concentrated areas (mean difference of the ratio was 0.0002; 95% limits of agreement, 0.005–0.005), suggesting highly correlated measurements (Fig 2).^{33,34}

DISCUSSION

In this post hoc analysis of 558 participants in the ESCAPE-NA1 trial who underwent baseline NCCT and 24-hour follow-up MR imaging, we show that pragmatic visual ratings of cortical atrophy on NCCT strongly correlate with atrophy assessments on MR imaging, though atrophy grades tended to be lower when assessed on NCCT compared with 24-hour MR imaging across different rating scales including GCA, Koedam, and MTA. Moreover, we found a strong correlation between NCCT and MR imaging for subcortical atrophy measurements using the intercaudate distance on the nonaffected hemisphere.

Compared with previous literature mostly involving considerably smaller patient samples, our study shows similar or higher agreement between atrophy ratings performed on NCCT and MR imaging. One observational study including 70 patients with stroke that assessed cortical and subcortical atrophy on NCCT and MR imaging using a more complex 5-point visual rating scale²¹ found lower agreement than in our study (weighted κ = 0.43–0.61 for cortical atrophy and 0.53–0.70 for subcortical atrophy). In a prospective cross-sectional study of 30 patients suspected of having dementia, the GCA and MTA scales on NCCT and MR imaging²⁰ showed agreement similar to our results (weighted κ = 0.83 for the GCA scale, 0.86 and 0.88 for the MTA scale on the right and left side, respectively). Notably, in this study, imaging acquisitions were from patients recruited in a memory clinic rather than the acute emergency setting. In addition, a recently conducted retrospective study including 214 patients with neurodegenerative disorders used computational atrophy analyses including automated volumetrics of different brain regions (ie, frontal, temporal, parietal, and occipital) and showed strong correlations of the obtained volumes in all regions (Pearson correlation > 0.82) between NCCT and MR imaging.¹²

Our study adds to previous literature providing empiric evidence that a simple visual assessment of brain atrophy on routine NCCT is feasible in patients with acute neurologic conditions such as stroke or traumatic brain injury and represents a valuable

diagnostic tool in the emergency setting when the numbers of acutely unwell patients are relatively high and performance of MR imaging or sophisticated volumetric analyses is not practical.

Pragmatic visual atrophy assessment may have important prognostic value complementing routine neuroimaging beyond the confirmation of acute stroke diagnosis and depiction of ischemic penumbra to inform treatment planning as well as recovery expectations for patients and treating physicians.

Table 1: Baseline characteristics (n = 558)

Clinical Characteristics	
Age (mean) (yr)	67.2 (SD, 13.7)
Sex (female) (No.) (%)	282 (50.5)
Hypertension (No.) (%)	386 (69.2)
Current smoker (No.) (%)	130 (23.3), n = 557
Dyslipidemia (No.) (%)	245 (43.9)
Atrial fibrillation (No.) (%)	172 (30.8)
Coronary artery disease (No.) (%)	121 (21.7)
Diabetes mellitus (No.) (%)	92 (16.5)
Any prior stroke (No.) (%)	76 (13.6)
Peripheral vascular disease (No.) (%)	19 (3.4)
Systolic blood pressure (mean) (mm Hg)	147.4 (SD, 25.1) n = 557
ASPECTS (median) (IQR)	8 (7–9) n = 552

Note:—IQR indicates interquartile range.

Table 2: Agreement and correlation between NCCT and MR imaging for cortical and subcortical atrophy assessments, respectively^a

	NCCT	MRI	Gwet AC1 ^a
GCA (median) (min-max) ^b	0 (0–2)	0 (0–3)	0.815
GCA ≥ 1 (No.) (%)	133 (23.8)	245 (43.9)	0.500
Koedam (median) (min-max) ^b	0 (0–2)	0 (0–3)	0.784
Koedam ≥ 1 (No.) (%)	172 (30.8)	261 (46.8)	0.512
MTA (median) (min-max) ^b	0 (0–4)	0 (0–4)	0.833
MTA ≥ 1 (No.) (%)	123 (22.0)	213 (38.2)	0.554
CC/IT ratio (median) (IQR) ^b	0.12 (0.09–0.14)	0.12 (0.09–0.15)	0.746 (Pearson correlation)
CC/IT ratio ≥ 0.16 (No.) (%) ^c	87 (15.6)	93 (16.7)	0.523

Note:—min indicates minimum; max, maximum.

^a All P values are $< .001$.

^b Full categoric scale is provided. Linear-weighted Gwet AC1.

^c Cutoff is chosen on the basis of prior literature.³⁵

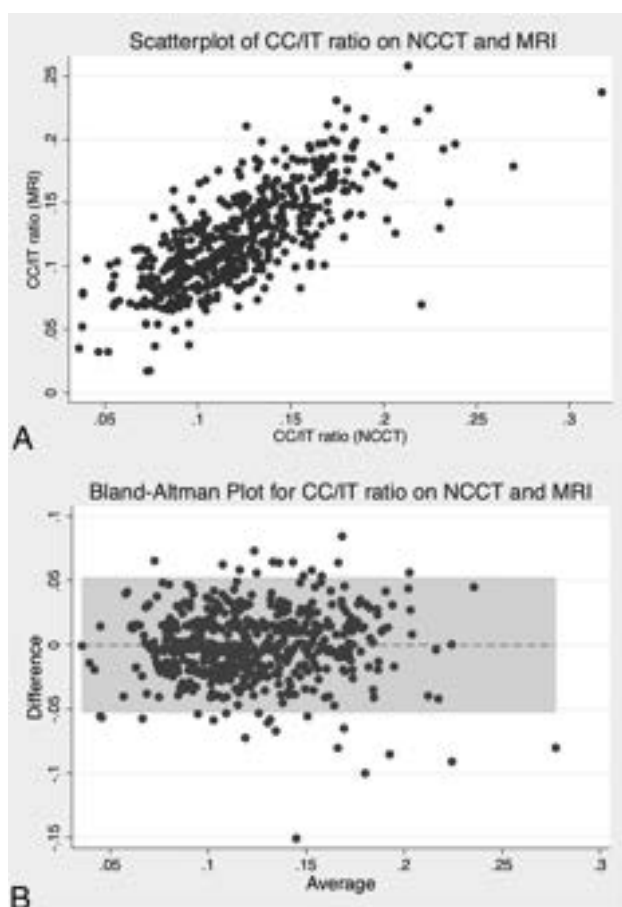


FIG 2. Scatterplot and Bland-Altman plot for the CC/IT ratio ($n = 558$). Scatterplot (A) and Bland-Altman plot (B) for the CC/IT ratio measured on NCCT and MR imaging. The Bland-Altman plot had a mean difference of 0.0002 (95% limits of agreement, -0.005 – 0.005). Highly concentrated areas could be observed, which can be seen in highly correlated measurements.

While an important value of our article is in demonstrating how simple measurements of brain atrophy on NCCT compare with MR imaging measurements, the specific setting of acute stroke used for our study is also of interest, given the emerging recognition of atrophy and other such measures of brain frailty as important mediators of treatment outcome in acute stroke.^{36–38} Because there is now considerable impetus for future acute stroke studies to factor such variables into the analyses, our finding that cortical and subcortical atrophy can be feasibly measured on NCCT in the acute setting is reassuring. In addition, future studies could examine how the underestimation of atrophy burden on NCCT affects the prediction of poststroke outcomes.

Limitations

First, because all raters in our study had several years of experience with neuroimaging readings and received similar training using the scales at the start of the study, we cannot comment on to what extent interrater differences in experience or training could contribute to discrepancies between NCCT and MR imaging assessments. Second, even though we used the h-ICD to account for any stroke-related edema, possible evolution of edema during the 24-hour interval between baseline NCCT and follow-up MR

imaging could theoretically confound comparisons. However, cases with extensive midline shift and/or sulcal or ventricular effacement that were noninterpretable on the basis of the judgment of the reader were excluded and accounted for $<1\%$ of the cases in the data set. Third, because this was a multicenter international trial, scanners differed from one site to another, and we did not impose specific MR imaging parameters beyond the expectations for a minimum set of sequences. However, this choice allowed us to perform a more pragmatic comparison of ratings performed on CT and MR imaging that is more generalizable to the heterogeneities of real-world practice.

CONCLUSIONS

Brain atrophy can be feasibly evaluated using simple measures in emergently acquired NCCT imaging. Subcortical atrophy measurement using CC/IT ratios are highly reproducible on both NCCT and MR imaging axial planes. Cortical atrophy assessments on NCCT have substantial or better agreement with MR imaging ratings but generally underestimate the severity of atrophy.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

1. Rabinstein AA, Albers GW, Brinjikji W, et al. **Factors that may contribute to poor outcome despite good reperfusion after acute endovascular stroke therapy.** *Int J Stroke* 2019;14:23–31 CrossRef Medline
2. Pedraza MI, de Lera M, Bos D, et al. **Brain atrophy and the risk of futile endovascular reperfusion in acute ischemic stroke.** *Stroke* 2020;51:1514–21 CrossRef Medline
3. Luijten SP, Compagne KC, van Es AC, et al. **Brain atrophy and endovascular treatment effect in acute ischemic stroke: a secondary analysis of the MR CLEAN trial.** *Int J Stroke* 2021 Oct 28. [Epub ahead of print] CrossRef Medline
4. Lauksio I, Lindström I, Khan N, et al. **Brain atrophy predicts mortality after mechanical thrombectomy of proximal anterior circulation occlusion.** *J Neurointerv Surg* 2021;13:415–20 CrossRef Medline
5. Lamballais S, Zijlmans JL, Vernooij MW, et al. **The risk of dementia in relation to cognitive and brain reserve.** *J Alzheimers Dis* 2020;77:607–18 CrossRef Medline
6. Fuchs TA, Benedict RH, Bartnik A, et al. **Preserved network functional connectivity underlies cognitive reserve in multiple sclerosis.** *Hum Brain Mapp* 2019;40:5231–41 CrossRef Medline
7. Yoon HJ, Kim SG, Kim SH, et al; For the Alzheimer's Disease Neuroimaging Initiative. **Associations between brain reserve proxies and clinical progression in Alzheimer's disease dementia.** *Int J Environ Res Public Health* 2021;18:12159 CrossRef
8. Benali F, Singh N, Fladt J, et al. **Brain frailty mediates the relationship between age and 90-day functional outcome after endovascular therapy: analysis of the Escape-NA1 trial.** *Stroke* 2023;54 (Suppl1):AWMP57 CrossRef
9. Hussein NE, Katzan IL, Rost NS, et al; American Heart Association Stroke Council; Council on Cardiovascular and Stroke Nursing; Council on Cardiovascular Radiology and Intervention; Council on Hypertension; and Council on Lifestyle and Cardiometabolic Health. **Cognitive impairment after ischemic and hemorrhagic stroke: a scientific statement from the American Heart Association/American Stroke Association.** *Stroke* 2023;54:e272–91 CrossRef Medline
10. Bigler ED. **The clinical significance of cerebral atrophy in traumatic brain injury.** *Arch Clin Neuropsychol* 1987;2:293–304 Medline
11. Caspers J, Heeger A, Turowski B, et al. **Automated age- and sex-specific volumetric estimation of regional brain atrophy: workflow and feasibility.** *Eur Radiol* 2021;31:1043–48 CrossRef Medline

12. Kaipainen AL, Pitkänen J, Haapalinna F, et al. **A novel CT-based automated analysis method provides comparable results with MRI in measuring brain atrophy and white matter lesions.** *Neuroradiology* 2021;63:2035–46 CrossRef Medline
13. Pasquier F, Leys D, Weerts JG, et al. **Inter- and intraobserver reproducibility of cerebral atrophy assessment on MRI scans with hemispheric infarcts.** *Eur Neurol* 1996;36:268–72 CrossRef Medline
14. Koedam EL, Lehmann M, van der Flier WM, et al. **Visual assessment of posterior atrophy development of an MRI rating scale.** *Eur Radiol* 2011;21:2618–25 CrossRef Medline
15. Torisson G, van Westen D, Stavenow L, et al. **Medial temporal lobe atrophy is underreported and may have important clinical correlates in medical inpatients.** *BMC Geriatr* 2015;15:65 CrossRef Medline
16. van Zagten M, Kessels F, Boiten J, et al. **Interobserver agreement in the assessment of cerebral atrophy on CT using bicaudate and Sylvian-fissure ratios.** *Neuroradiology* 1999;41:261–64 CrossRef Medline
17. Kipps CM, Davies RR, Mitchell J, et al. **Clinical significance of lobar atrophy in frontotemporal dementia: application of an MRI visual rating scale.** *Dement Geriatr Cogn Disord* 2007;23:334–42 CrossRef Medline
18. Boldor N, Vaknin S, Myers V, et al. **Reforming the MRI system: the Israeli National Program to shorten waiting times and increase efficiency.** *Isr J Health Policy Res* 2021;10:57 CrossRef Medline
19. Frija G, Blazović I, Frush DP, et al. **How to improve access to medical imaging in low- and middle-income countries?** *EclinicalMedicine* 2021;38:101034 CrossRef Medline
20. Wattjes MP, Henneman WJ, van der Flier WM, et al. **Diagnostic imaging of patients in a memory clinic: comparison of MR imaging and 64-detector row CT.** *Radiology* 2009;253:174–83 CrossRef Medline
21. Ferguson KJ, Cvorovic V, MacLulich AM, et al. **Visual rating scales of white matter hyperintensities and atrophy: comparison of computed tomography and magnetic resonance imaging.** *J Stroke Cerebrovasc Dis* 2018;27:1815–21 CrossRef Medline
22. Lee SH, Oh CW, Han JH, et al. **The effect of brain atrophy on outcome after a large cerebral infarction.** *J Neurol Neurosurg Psychiatry* 2010;81:1316–21 CrossRef Medline
23. Butzkueven H, Kolbe SC, Jolley DJ, et al. **Validation of linear cerebral atrophy markers in multiple sclerosis.** *J Clin Neurosci* 2008;15:130–37 CrossRef Medline
24. Hill MD, Goyal M, Menon BK, et al; ESCAPE-NA1 Investigators. **Efficacy and Safety of Nerinetide for the Treatment of Acute Ischaemic Stroke (ESCAPE-NA1): a multicentre, double-blind, randomised controlled trial.** *Lancet* 2020;395:878–87 CrossRef Medline
25. Schmaranzer F, Kallini JR, Miller PE, et al. **The effect of modality and landmark selection on MRI and CT femoral torsion angles.** *Radiology* 2020;296:381–90 CrossRef Medline
26. Liao JJ. **Sample size calculation for an agreement study.** *Pharm Stat* 2010;9:125–32 CrossRef Medline
27. Zec S, Soriani N, Comoretto R, et al. **High agreement and high prevalence: the paradox of Cohen's Kappa.** *Open Nurs J* 2017;11:211–18 CrossRef Medline
28. 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
29. Gwet KL. **Testing the difference of correlated agreement coefficients for statistical significance.** *Educ Psychol Meas* 2016;76:609–37 CrossRef Medline
30. Wongpakaran N, Wongpakaran T, Wedding D, et al. **A comparison of Cohen's kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples.** *BMC Med Res Methodol* 2013;13:61 CrossRef Medline
31. Landis JR, Koch GG. **The measurement of observer agreement for categorical data.** *Biometrics* 1977;33:159–74
32. Schober P, Boer C, Schwarte LA. **Correlation coefficients: appropriate use and interpretation.** *Anesth Analg* 2018;126:1763–68 CrossRef Medline
33. Giavarina D. **Understanding Bland Altman analysis.** *Biochem Med (Zagreb)* 2015;25:141–51 CrossRef Medline
34. Bland JM, Altman DG. **Statistical methods for assessing agreement between two methods of clinical measurement.** *Lancet* 1986;1:307–10 Medline
35. Park J, Goh DH, Sung JK, et al. **Timely assessment of infarct volume and brain atrophy in acute hemispheric infarction for early surgical decompression: strict cutoff criteria with high specificity.** *Acta Neurochir (Wien)* 2012;154:79–85 CrossRef Medline
36. Benali F, Fladt J, Jaroengarmsamer T, et al. **Association of white matter disease with functional recovery and 90-day outcome after EVT: beyond chronological age.** *Stroke Vasc Interv Neurol* 2023;3:e000734 CrossRef
37. Bretzner M, Bonkhoff AK, Schirmer MD, et al; MRI-GENIE and GISCOME Investigators and the International Stroke Genetics Consortium. **Radiomics-derived brain age predicts functional outcome after acute ischemic stroke.** *Neurology* 2023;100:e822–33 CrossRef Medline
38. Benali F, Fladt J, Jaroengarmsamer T, et al. **Association of brain atrophy with functional outcome and recovery trajectories after thrombectomy: Post-hoc analysis of the ESCAPE-NA1 Trial.** *Neurology* 2023;17 CrossRef Medline

Prospective, Longitudinal Study of Clinical Outcome and Morphometric Posterior Fossa Changes after Craniocervical Decompression for Symptomatic Chiari I Malformation

Alaaddin Ibrahimy, Tianxia Wu, Jessica Mack, Gretchen C. Scott, Michaela X. Cortes, Fredric K. Cantor, Francis Loth, and John D. Heiss



ABSTRACT

BACKGROUND AND PURPOSE: The time course of changes in posterior fossa morphology, quality of life, and neurologic function of patients with Chiari I malformation after craniocervical decompression requires further elaboration. To better understand the pace of these changes, we longitudinally studied patients with Chiari I malformation, with or without syringomyelia, before and after the operation for up to 5 years.

MATERIALS AND METHODS: Thirty-eight symptomatic adult patients (35 women, 3 men) diagnosed with Chiari I malformation only ($n=15$) or Chiari I malformation and syringomyelia ($n=23$) and without previous Chiari I malformation surgery were enrolled in a clinical study. Patients underwent outpatient study visits and MR imaging at 7 time points (ie, initial [before the operation], 3 months, 1 year, 2 years, 3 years, 4 years, and 5 years) during 5 years. The surgical procedure for all patients was suboccipital craniectomy, C1 laminectomy, and autologous duraplasty.

RESULTS: Morphometric measurements demonstrated an enlargement of the CSF areas posterior to the cerebellar tonsils after the operation, which remained largely stable through the following years. There was a decrease in pain and improved quality of life after the operation, which remained steady during the following years. Reduction in pain and improved quality of life correlated with CSF area morphometrics.

CONCLUSIONS: Most changes in MR imaging morphometrics and quality of life measures occurred within the first year after the operation. A 1-year follow-up period after Chiari I malformation surgery is usually sufficient for evaluating surgical efficacy and postoperative MR imaging changes.

ABBREVIATIONS: CMI = Chiari I malformation; KPS = Karnofsky Performance Scale; LS = least square; PFD = posterior fossa and foramen magnum decompression surgery; PCF = posterior cranial fossa; PFDD = PFD with dural opening and duraplasty; RM-ANCOVA = repeated measures ANCOVA; RM-ANOVA = repeated-measures ANOVA; TP = time point

In Chiari I malformation (CMI), tonsillar herniation and posterior fossa underdevelopment disrupt normal CSF dynamics.¹ Typical CMI symptoms are posterior headache and suboccipital neck pain, aggravated by straining, coughing, and the Valsalva

maneuver. Symptoms of visual disturbances, dizziness, imbalance, fatigue, and cognitive impairment also occur. Patients with associated syringomyelia often have spinal cord symptoms and signs.² Surgical indications in CMI include syringomyelia-related neurologic deficits and cough-related headache.²⁻⁴ Posterior fossa and foramen magnum decompression surgery (PFD) reduces CMI symptoms, restores CSF flow, and treats syringomyelia.^{5,6}

PFD for CMI varies from bony decompression to operations opening the dura, arachnoid, and fourth ventricle, reducing tonsillar size, and including duraplasty.⁶ Intraoperative MR imaging or sonography can assess whether posterior bone removal decompressed the cerebellar tonsils and opened the CSF pathways or if additional decompression measures are necessary.⁷⁻⁹ Most symptoms improve after PFD surgery.¹⁰⁻¹³ Children improve more than adults.¹⁴ Complications like CSF leakage, swallowing dysfunction, aseptic meningitis, and infection worsen surgical outcomes.¹⁵

Received June 13, 2023; accepted after revision August 16.

From the Surgical Neurology Branch (A.I., J.M., G.C.S., M.X.C., F.K.C., J.D.H.) and Clinical Trials Unit (T.W.), National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland; Department of Biomedical Engineering (A.I.), Yale University, New Haven, Connecticut; and Departments of Mechanical and Industrial Engineering, and Bioengineering (F.L.), Northeastern University College of Engineering, Boston, Massachusetts.

This work was supported by the National Institutes of Health, National Institute of Neurological Disorders and Stroke, grant No. NIH-ZIANS003052-15.

Please address correspondence to Alaaddin Ibrahimy, M.S.E., Yale University, 40 Temple St, 7G, New Haven, CT 06510; e-mail: alaaddin.ibrahimy@yale.edu

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

Indicates article with online supplemental data.

Indicates article with supplemental online video.

<http://dx.doi.org/10.3174/ajnr.A7993>

Table 1: Number of subjects at each TP category for clinical outcomes and morphometric measurements

TP categories	−0.5	0.5	1	2	3	4	5
TP (yr)	−0.5	0.5	1	2	3	4	5
Relative to surgery (mo)	−9 to −1	0–9	9–18	18–30	30–42	42–54	>54
Number of subjects for each TP							
Morphometric measurements	38	35	31	23	19	17	14
KPS score	36	34	32	24	22	16	14
Average pain	34	33	32	23	21	14	14
Mean of affective pain	34	33	32	23	21	14	14
Mean of continuous pain	33	33	32	23	21	14	14
Mean of intermittent pain	34	33	32	23	21	14	14
Mean of neuropathic pain	34	33	32	23	21	14	14
ASIA total score	35	35	33	23	24	17	14
Ambulatory score	35	35	33	24	24	17	14
Cognitive subtotal	35	35	32	24	22	16	14
McCormick class score	35	35	32	24	24	17	14
Motor subtotal	35	35	32	24	22	16	14
Total FIM	35	35	32	24	22	16	14

Note:—ASIA indicates American Spinal Injury Association; FIM, Functional Independence Measure.

PFD affects anatomic morphologies and CSF flow.^{7,16–19} Favorable treatment outcomes follow PFD, which enlarges the retrotonsillar CSF space and relieves CSF flow obstruction.^{13,18,20} Tonsillar ectopia decreases after PFD, almost 50% in studies by Heiss et al^{17,18} and Nikoobakht,¹³ but 7%, 19%, and 33%, respectively, in studies by Bond et al,⁷ Quon et al,¹⁹ and Eppelheimer et al.¹⁶ The CSF pathway expands in various amounts after PFD surgery. In patients undergoing PFD with dural opening and duraplasty (PFDD), the retrotonsillar CSF space enlarged 11–13 mm and the ventral foramen magnum CSF space enlarged 0.8–2 mm.^{17,18} PFDD enlarged the ventral CSF width more in patients with syringomyelia.^{16,18} In another duraplasty study, the retrotonsillar space widened 9.7 mm but the ventral CSF width did not change significantly.¹⁶ In a study of PFD, most (86%) without duraplasty, the foramen magnum CSF pathways enlarged only 1 mm dorsally and 0.2 mm ventrally.⁷

Khalsa et al,²¹ in 2017, reported that more significant (5.89% versus 1.54%) posterior fossa enlargement after PFD was associated with relief of CMI-related headache and reduced tonsillar ectopia. Noudel et al²⁰ reported that complete recovery was associated with a 15% increase in posterior cranial fossa (PCF) volume, and partial recovery, with a 7% increase in PCF volume after PFD.

This prospective, longitudinal study of patients with CMI with and without syringomyelia treated with PFDD aimed to achieve the following: 1) quantify the change in the posterior fossa and upper cervical spine morphometrics between before PFDD surgery (baseline, denoted “−0.5 time point [TP]”) and TPs up to 5 years after PFDD surgery, and 2) determine the change in symptomatology between before PFDD surgery (baseline) and TPs after surgery. The study also evaluated whether the Karnofsky Performance Scale (KPS) score, the primary performance outcome measure of the study, was related to retrotonsillar CSF space expansion after the operation, the primary morphometric outcome measure of the study, or other morphometric changes (exploratory measures). Establishing the timing of functional improvement and morphologic changes following PFDD

may guide neurosurgeons advising patients with CMI about expected recovery and MR imaging after PFDD.

MATERIALS AND METHODS

Study Population and Ethics

Thirty-eight adult patients (35 women, 3 men) diagnosed with CMI only ($n = 15$) or CMI and syringomyelia ($n = 23$) and without previous CMI surgery were enrolled in a clinical protocol and underwent PFDD surgery. The NIH Combined Neuroscience Institutional Review Board approved the research protocol: 10N0143. Consent was obtained from all patients. All patients were evaluated and treated at a single clinical site. The mean age at surgery was 40.6 (SD, 12.6) years (range, 20.3–64.9 years). The protocol specified outpatient study visits and MR imaging 7 times (ie, initial [before the operation], 3 months, 1 year, 2 years, 3 years, 4 years, and 5 years after the operation). MR images and clinical data were collected between 2010 and 2021. MR images were stored in the PACS server. Clinical data were recorded on case report forms, transcribed, and stored in a secure database. All 38 patients attended at least 1 scheduled visit after the operation, and 14 completed the 5-year protocol (Table 1).

Surgical Procedure

All patients underwent the operation in a prone, horizontal position with the head in a neutral or gently flexed posture. The surgical procedure included a suboccipital craniectomy, C1 laminectomy, and sometimes removal of the superior part of the lamina of C2. The bony opening decompressed the inferior cerebellar hemispheres and cerebellar tonsils. A Y-shaped durotomy was performed. The arachnoid was preserved. Intradural structures were not manipulated. An autologous dural graft of the occipital pericranium measuring 4×4 cm was sewn to the durotomy edges using a 4–0 braided nylon suture.¹⁸

Image Registration and Midsagittal Plane Selection

We used Matlab (MathWorks), FSL (<http://www.fmrib.ox.ac.uk/fsl>), and bash scripts in parallel to register 3D T1 MR images to the Montreal Neurological Institute images (MNI152_T1_brain.nii.gz, from FSL) and identify the midsagittal plane. DICOM (.dcm) images were initially converted to NIfTI (.nii.gz) format. The “dcm2nii” Unix command compressed file sizes. FSL commands extracted the brain from MR images. A linear image registration tool (FLIRT command; FMRIB Linear Image Registration Tool; FLIRT; <http://www.fmrib.ox.ac.uk/fsl/flwiki/FLIRT>) registered the skull-free image to the atlas image (MNI152_T1_brain.nii.gz). The FSL transformation matrix, including the superior-inferior ROI, was applied to the original images, registering them in the atlas image orientation. The final midsagittal MR image in the midsagittal plane of the atlas was saved as a Matlab file for morphometric measurements.

Morphometric Measurements and Software

We made 9 morphometric measurements using 2 custom software programs developed in Matlab (Table 2, Fig 1, and Supplemental Online Video). In previous studies, these measurements significantly changed after PFDD surgery.^{16,22} PFDD removes the opisthion. A described method reconstructed the McRae line on postsurgical images.^{16,18}

Clinical Outcomes Measures

The primary clinical outcome measure was the KPS score change between before and after the operation. The KPS was graded on a 0–100 scale at study visits, representing the patient's ability to perform everyday activities, work, and function without assistance.²³

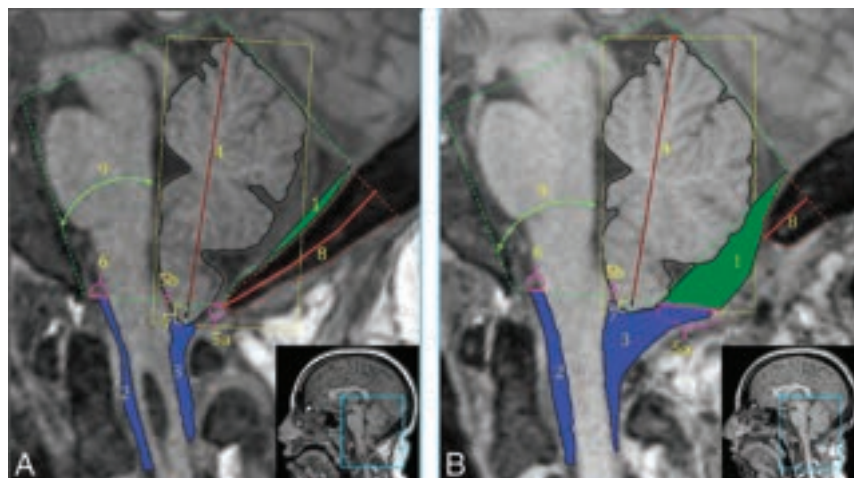


FIG 1. Morphometric measurements on a presurgery MR image (A) and a postsurgery MR image (B) for the same patient 1 year later. 1) CSF area posterior to PCF, 2) anterior CSF area, 3) posterior CSF area, 4) cerebellar height, 5a) posterior dorsal width + (5b) anterior dorsal width = total dorsal CSF width, 6) ventral CSF width, 7) cerebellar tonsillar position, 8) occipital bone length, and 9) cerebellum-clivus bone angle. Note that the syrinx in the presurgical MR image resolved after surgery (Supplemental Online Video).

Exploratory clinical outcomes included the following: 1) the Short-form McGill Pain Questionnaire, a self-reported pain questionnaire used in longitudinal studies;²⁴ 2) the McCormick Clinical/Functional classification, a 4-grade scale ranging from grade I (normal or insignificant functional neurologic deficit) to grade IV (severe deficit, quadriplegia, usually not independent);²⁵ 3) the Functional Ambulatory Score, a 4-grade scale ranging from 1 (ambulant without a walking aid) to 4 (complete motor and sensory paraplegia); 4) the American Spinal Injury Association motor grading scale,²⁶ a 0–100 scale based on a sum of the strengths of 5 upper- and 5 lower-extremity muscles graded from 5 (normal) to 0 (total paralysis); and 5) the Functional Independence Measure, a 13-item questionnaire assessing motor and cognitive functioning on a 1 (total assistance)

to 7 (complete independence) ordinal scale.²⁷ The Chicago Chiari Outcome Scale complication component graded complications as 1 (persistent complication, poorly controlled), 2 (persistent complication, well-controlled), 3 (transient complication), and 4 (uncomplicated course).²⁸

Morphologic Outcomes Measures

The primary morphometric outcome measure was the dorsal CSF space, denoted by 5 in Fig 1. The dorsal CSF space (5) summated the anterior-posterior widths of the foramen magnum CSF space (5a) and the retrotonsillar space (5b). Eight other morphometric measures shown in Table 2 were exploratory.

Data Distribution

TPs represented the order of scheduled study visits, including MR imaging scans and clinical data collection. The

Table 2: Description of morphometric measurements

Morphometric Measurement		Description
Area		
1	CSF area posterior to PCF (mm ²)	CSF area posterior to a line drawn between the IOP and opisthion, and posterior to the cerebellum
2	Anterior CSF area (mm ²)	CSF area in the upper cervical spinal canal anterior to the spinal cord between the FM and inferior limit of the C2 vertebra
3	Posterior CSF area (mm ²)	CSF area in the upper cervical spinal canal posterior to the spinal cord between the FM and inferior limit of the C2 vertebra
Length		
4	Cerebellar height (mm)	Distance between the most superior point of the superior vermis and the most inferior point of the tonsil
5	Total dorsal CSF width (mm)	Width of the CSF space measured at the level of the FM, anterior to the cerebellum, posterior to the brainstem and posterior the cerebellum, anterior to the subarachnoid space
6	Ventral CSF width (mm)	Width of the CSF space measured at the level of the FM, anterior to the brainstem, posterior to the subarachnoid space
7	Cerebellar tonsillar position (mm)	The perpendicular distance between the most inferior tip of the cerebellar tonsils and the McRae line
8	Occipital bone length (mm)	Length between a point at the midpoint of the occipital bone at the level of the tentorium and IOP and the most inferior tip of the occipital bone
Angle		
9	Cerebellum-clivus bone angle (degree)	The angle subtended by the major axis of the cerebellum with the posterior margin of the clivus bone

Note:—IOP indicates internal occipital protuberance; FM, foramen magnum.

TP before surgery was -0.5 , with the negative sign connoting before surgery. The TP values without negative sign prefixes connoted visits after surgery.

Statistical Analysis

The initial statistical plan of analysis was a repeated measures analysis of variance (RM-ANOVA) or analysis of covariance (RM-ANCOVA), assuming normality. The Box-Cox transformation was applied to the 12 clinical variables with skewed distributions. However, the transformation did not work for 6 of them, and the Friedman test was performed. The Shapiro-Wilk method tested the normality assumption.

For each morphometric and transformed clinical variable, either RM-ANOVA or RM-ANCOVA was used to evaluate the change with time. The TP (years from the operation) variable was treated as categorical because the change in all variables across time did not show linear trends. Age and diagnosis (with and without syringomyelia) were considered covariates, and the significance level of .1 was applied to covariate selection (including 2 interactions: TP \times age and TP \times diagnosis). For the compound symmetry, the heterogeneous covariance structure was selected based on the Akaike information criterion to account for the correlation among the repeated measures from the same subject and the heterogeneity of time variance. The Dunnett method was applied to multiple comparisons using TP -0.5 or TP 0.5 as the control. The Bonferroni correction was used for multiple comparisons. For the Box-Cox transformed variables, the back-transformed least square (LS) means and 95% CI were reported.

Spearman correlation coefficients were calculated to assess the association between the clinical and morphometric variables based on the following: 1) the relative change: between pre- (TP = -0.5) and postsurgery (TP = 0.5): $100 \times (\text{post to pre})/\text{pre}$, and 2) the measurements at TP = -0.5 and TP = 0.5 . The P values were not adjusted for the multiple statistical tests of the exploratory variables. SAS Version 9.4 (SAS Institute) was used for statistical analysis, and P values $< .05$ were considered statistically significant.

RESULTS

Morphometric Measurements

Age and diagnosis (CMI only versus CMI and syringomyelia) did not significantly affect any morphometric variables except for the anterior CSF area and cerebellum-clivus bone angle (F -test, P value $> .1$). RM-ANOVA was, therefore, used for all morphometric variables except the anterior CSF area and cerebellum-clivus bone angle (Fig 2). The RM-ANCOVA analyzed the anterior CSF area (diagnosis as a covariate) and the cerebellum-clivus bone angle (age as a covariate). The RM-ANOVA or RM-ANCOVA (Fig 2) showed a significant change across TPs (F -test, $df=6$, ventral CSF width: $P=.023$, others: $P<.0001$) in all variables except the anterior CSF area in the CMI-only group (F -test, $df=6$, $P<.1499$). The difference between presurgery (TP = -0.5 , control) and each postsurgery TP was significant (adjusted $P<.0001$). Furthermore, the change between the first (TP = 0.5) and postsurgical TPs for the CSF area posterior to the PCF was also significant ($P<.05$). For the posterior CSF area, a significant change was observed at TPs 1 and 2 ($P<.05$). For the other 7 variables, the difference between the first

postsurgery TP (TP = 0.5 , control) and each postsurgical TP (TP = $1-5$) was not significant (adjusted $P>.05$).

The total dorsal CSF width increased by 12 mm after PFDD surgery (Fig 2). The posterior CSF area increased its preoperative size >2.5 -fold after the operation. The CSF area posterior to the PCF increased 3-fold by 3–6 months after the operation (TP = 0.5), 4-fold at TP = 1, and remained stable thereafter.

The ventral CSF width for the entire CM cohort did not change significantly following PFDD ($P>.05$). The RM-ANCOVA confirmed significant interaction between diagnosis and TP ($P=.0374$). The anterior CSF area following PFDD surgery increased by $>25\%$ in the CMI and syringomyelia group but $<10\%$ in the CMI-only group (Fig 2).

The change from baseline for the cerebellar tonsillar position inferior to the McRae line, cerebellar height, occipital bone length, and cerebellum-clivus bone angle occurred within the first 3–6 months after surgery (Fig 2). The cerebellar tonsillar position inferior to the McRae line decreased 2 mm 3–6 months after surgery (TP = 0.5) and 0.8 mm more in following years. The cerebellar height decreased by 2 mm following PFDD surgery. The occipital bone length was 23 mm shorter following PFDD surgery. PFDD surgery produced a 3° – 4° reduction in the cerebellum-clivus bone angle 3–6 months after the operation (TP = 0.5).

Clinical Outcomes

Clinical outcomes significantly changed across TPs (F -test, $df=6$, $P<.01$; RM-ANOVA). Age and diagnosis had no significant effect (F -test P value $> .1$). The difference between presurgery (TP = -0.5 as a control) and postsurgery TP = 0.5 was significant (adjusted $P<.05$) for the KPS score and pain variables (Fig 3). After the PFDD surgery, the mean KPS score increased from 81 to 88 (Fig 3). Average pain and other pain measures decreased by almost 50% 3–6 months following PFDD surgery and remained stable thereafter (Fig 3). The other 6 clinical variables (American Spinal Injury Association total score, Functional Ambulatory Score, Total Functional Independence Measure Cognitive Subtotal, Motor Subtotal, McCormick Class Score) did not change significantly following PFDD surgery (Friedman test, $df=6$, $P>.05$).

Five transient treatment complications occurred. There were endotracheal tube–related tongue abrasions in 3 patients, one of whom also had postoperative pneumonia. Another patient developed atrial fibrillation requiring cardioversion 4 days postoperatively. There were no CSF leaks or wound infections.

Correlation between Clinical Outcomes and Morphometric Measurements

There was no significant correlation at TP = -0.5 between any clinical outcomes and morphometric measurements. At TP = 0.5 , the posterior CSF area was weakly correlated with average pain and mean of intermittent pain ($n=33$, $r=0.375$, $P=.0314$; $r=0.431$, $P=.013$, respectively), and the total dorsal CSF width was weakly correlated with average pain, mean of intermittent pain, and mean of neuropathic pain ($n=33$, $r=0.39$, $P=.0248$; $r=0.413$, $P=.0168$; $r=0.366$, $P=.0362$, respectively). The relative change in the KPS score had a weak negative correlation with occipital bone length ($n=28$, $r=-0.483$, $P=.0093$). The relative change of the cerebellum-clivus bone angle was weakly associated

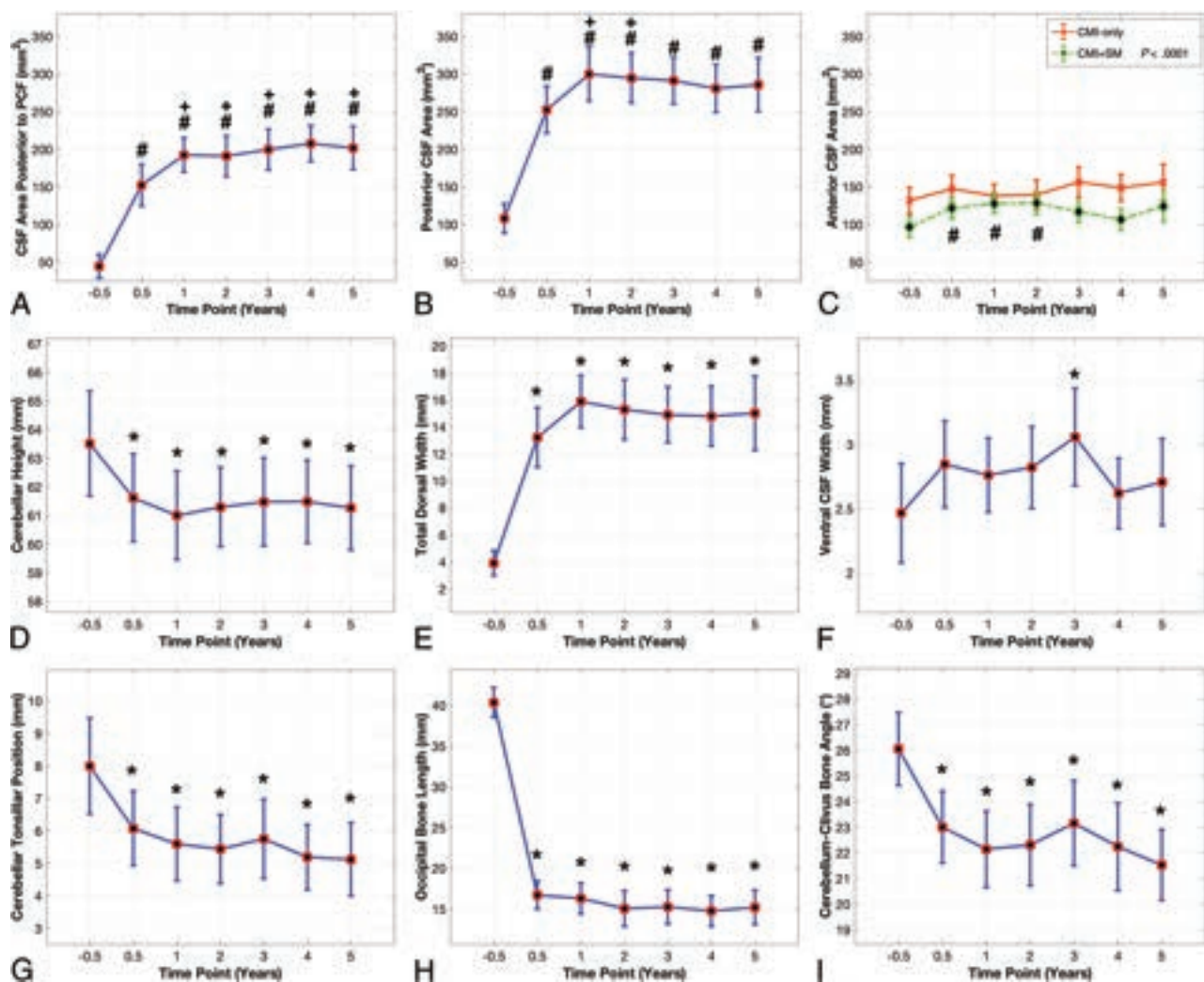


FIG 2. Morphometric length, angle, and CSF area measurements (LS-means and 95% CI). A, CSF area posterior to the PCF. B, Posterior CSF area. A and B, RM-ANOVA *F*-test ($df = 6$) of TP: P value $< .0001$. C, Anterior CSF area for CMI-only, and for patients with CMI plus syringomyelia; C, RM-ANCOVA (diagnosis as a covariate), *F*-test ($df = 6$) for the interaction (diagnosis \times TP): $P = .0374$. RM-ANOVA *F*-test ($df = 6$) for TP: $P > .05$ for CMI-only, no multiple comparisons: $P < .0001$ for CMI and syringomyelia (SM). D, Cerebellar height. E, Total dorsal width. F, Ventral CSF width. G, Cerebellar tonsillar position, H, Occipital bone length. I, Cerebellum-clivus bone angle. RM-ANOVA *F*-test ($df = 6$) of TP: P value $< .0001$ except for ventral CSF width ($P = .023$, F). Hash tag indicates the Dunnett adjusted P value $< .05$ at the TP compared with TP = -0.5; plus sign, the Dunnett adjusted P value $< .05$ at the TP compared with TP = +0.5; asterisk, the Dunnett adjusted P value $< .01$ at the TP compared with TP = -0.5.

with the mean of affective pain (exploratory tests; $n = 28$, $r = 0.407$, $P = .0314$).

The Online Supplemental Data contain detailed statistical results like LS-means, 95% confidence intervals, and P values from the Dunnett multiple comparisons for morphometric measurements and clinical variables.

DISCUSSION

This study evaluated clinical signs and symptoms and posterior fossa morphology in patients with CMI before, at 3–6 months, and then annually for 5 years after craniocervical decompression and duraplasty. Clinical symptoms improved after the operation. We used 9 previously validated morphometric measures.¹⁶ Seven of the 9 measures changed significantly between the pre- and 3- to 6-month post-PFDD visits (Table 1).^{16–18,21} All changes occurred within 3–6 months after the operation, except for the CSF area

posterior to the PCF, which increased 25% between 3 and 6 months and 1 year after the operation (Fig 2).

In the present study, 60% of subjects had syringomyelia. The anterior CSF area significantly increased (27%) following PFDD surgery for the CMI with syringomyelia group, but not in the CMI-only group.^{16,18} Ventral CSF width combining CMI groups increased by 0.3 mm postsurgery, a statistically insignificant change.^{7,16} The semiautomated measurements used were more accurate than manual measurements.¹⁸

Among the 12 clinical symptoms, the KPS score increased significantly following PFDD surgery. Average pain and all 4 pain subcategories (mean of affective pain, mean of continuous pain, mean of intermittent pain, and mean of neuropathic pain) decreased significantly following PFDD surgery. Craniectomy reduced the occipital bone length by a mean of 24 mm, or 60%. More significant length reduction correlated weakly with KPS score improvement.^{16,18}

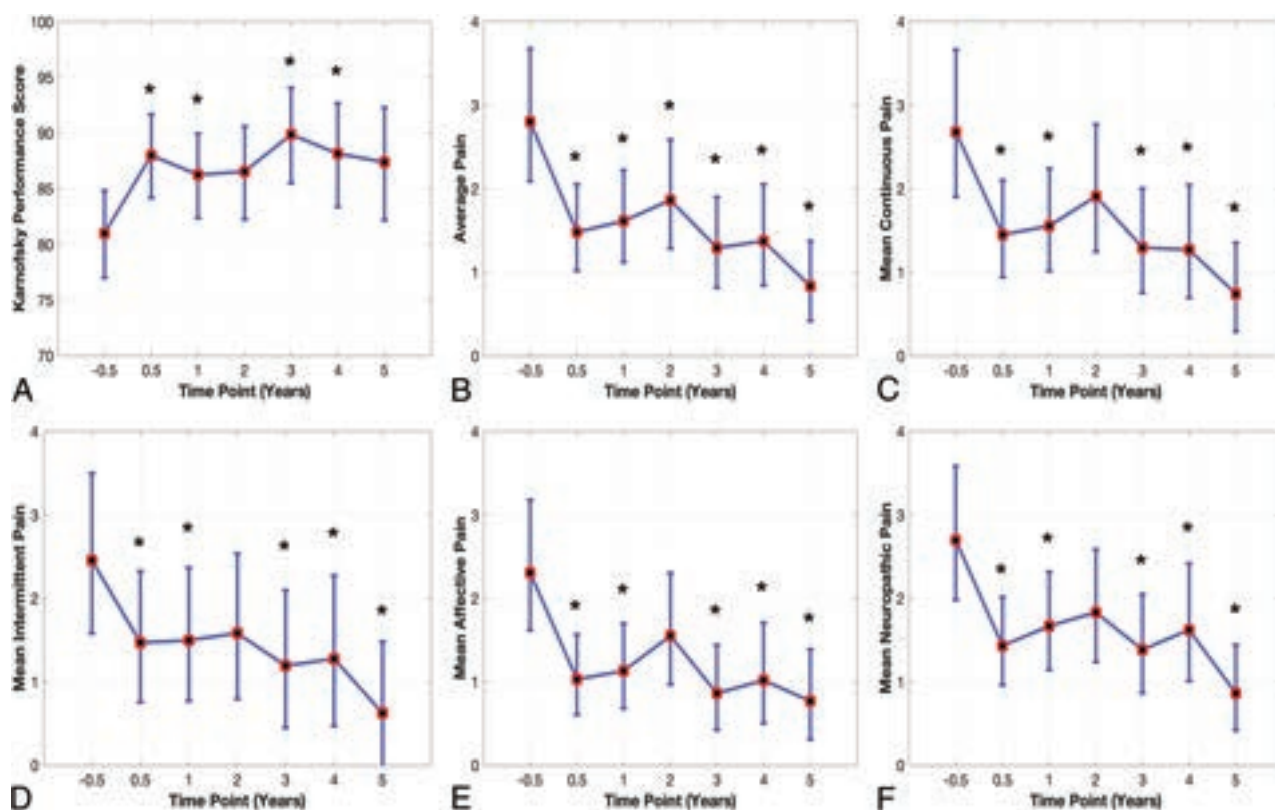


FIG 3. Clinical outcomes (back-transformed LS-means and 95% CI): A, KPS score. B, Average pain. C, Mean continuous pain. D, Mean intermittent pain. E, Mean affective pain. F, Mean neuropathic pain. RM-ANOVA F -test ($df = 6$) of TP: P value $< .005$. The asterisk indicates the Dunnett adjusted P value $< .05$ at the TP compared with TP = -0.5 .

The posterior CSF area doubled by 3–6 months after the operation, reached 2.5 times its preoperative size by 1 year after the operation, and remained stable thereafter.^{16,18} After PFDD surgery, total dorsal CSF width increased by 9–12 mm, >3.5 times its preoperative size, because of increased posterior CSF area, upward tonsillar migration, and reduced occipital bone length. The dura expanded posteriorly into the craniectomy. At TP = 0.5 postsurgery, the posterior CSF area weakly correlated with average pain and mean of intermittent pain. Total dorsal CSF width weakly correlated with average pain, mean of intermittent pain, and mean of neuropathic pain. Patients with a larger posterior CSF area after PFDD had less pain relief.

The cerebellum-clivus bone angle decreased 3° – 5° after PFDD decompression because tonsillar compression and entrapment relief allowed the inferior cerebellum to rotate posteriorly and superiorly.¹⁸ Eppelheimer et al¹⁶ reported less (2.2°) counter-clockwise rotation. The change in cerebellar rotation weakly correlated with the relative change in mean affective pain.

Limitations

Patients infrequently did not complete functional measurement questionnaires. Clinical outcomes did not consider surgically related changes in syrinx size.^{2,3,17}

CONCLUSIONS

After posterior fossa decompression, CMI symptoms and functional performance of patients improved, and their retrotonsillar CSF space enlarged. Symptomatic and MR imaging morphometric

changes occurred within the first postoperative year. One year is sufficient for evaluating surgical efficacy and postoperative MR imaging changes after PFDD surgery.









Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

- Bolognese PA, Brodbelt A, Bloom AB, et al. Professional profiles, technical preferences, surgical opinions, and management of clinical scenarios from a panel of 63 international experts in the field of Chiari I malformation. *World Neurosurg* 2020;140:e14–22 CrossRef Medline
- Milhorat TH, Chou MW, Trinidad EM, et al. Chiari I malformation redefined: clinical and radiographic findings for 364 symptomatic patients. *Neurosurgery* 1999;44:1005–17 CrossRef Medline
- Batzdorf U. Considerations regarding decompressive surgery for Chiari I malformation. *World Neurosurg* 2011;75:222–23 CrossRef Medline
- Chavez A, Roguski M, Killeen A, et al. Comparison of operative and non-operative outcomes based on surgical selection criteria for patients with Chiari I malformations. *J Clin Neurosci* 2014;21:2201–06 CrossRef Medline
- Arnautovic A, Splavski B, Boop FA, et al. Pediatric and adult Chiari malformation type I surgical series 1965–2013: a review of demographics, operative treatment, and outcomes. *J Neurosurg Pediatr* 2015;15:161–77 CrossRef Medline
- Yilmaz A, Kanat A, Musluman AM, et al. When is duraplasty required in the surgical treatment of Chiari malformation type I based on tonsillar descending grading scale? *World Neurosurg* 2011;75:307–13 CrossRef Medline

7. Bond AE, Jane JA Sr, Liu KC, et al. **Changes in cerebrospinal fluid flow assessed using intraoperative MRI during posterior fossa decompression for Chiari malformation.** *J Neurosurg* 2015;122:1068–75 CrossRef Medline
8. Cui LG, Jiang L, Zhang HB, et al. **Monitoring of cerebrospinal fluid flow by intraoperative ultrasound in patients with Chiari I malformation.** *Clin Neurol Neurosurg* 2011;113:173–76 CrossRef Medline
9. Milhorat TH, Bolognese PA. **Tailored operative technique for Chiari type I malformation using intraoperative color Doppler ultrasonography.** *Neurosurgery* 2003;53:899–905; discussion 905–06 CrossRef Medline
10. De Vlieger J, Dejaegher J, Van Calenbergh F. **Posterior fossa decompression for Chiari malformation type I: clinical and radiological presentation, outcome and complications in a retrospective series of 105 procedures.** *Acta Neurol Belg* 2019;119:245–52 CrossRef Medline
11. Giammattei L, Messerer M, Daniel RT, et al. **Long-term outcome of surgical treatment of Chiari malformation without syringomyelia.** *J Neurosurg Sci* 2020;64:364–68 CrossRef Medline
12. Mueller D, Oro JJ. **Prospective analysis of self-perceived quality of life before and after posterior fossa decompression in 112 patients with Chiari malformation with or without syringomyelia.** *Neurosurg Focus* 2005;18:1–6 CrossRef Medline
13. Nikoobakht M, Shojaei H, Gerszten PC, et al. **Craniometrical imaging and clinical findings of adult Chiari malformation type I before and after posterior fossa decompression surgery with duraplasty.** *Br J Neurosurg* 2019;33:481–85 CrossRef Medline
14. Gilmer HS, Xi M, Young SH. **Surgical decompression for Chiari malformation type I: an age-based outcomes study based on the Chicago Chiari Outcome Scale.** *World Neurosurg* 2017;107:285–90 CrossRef Medline
15. Klekamp J, Batzdorf U, Samii M, et al. **The surgical treatment of Chiari I malformation.** *Acta Neurochir (Wien)* 1996;138:788–801 CrossRef Medline
16. Eppelheimer MS, Biswas D, Braun AM, et al. **Quantification of changes in brain morphology following posterior fossa decompression surgery in women treated for Chiari malformation type I.** *Neuroradiology* 2019;61:1011–22 CrossRef Medline
17. Heiss JD, Patronas N, DeVroom HL, et al. **Elucidating the pathophysiology of syringomyelia.** *J Neurosurg* 1999;91:553–62 CrossRef Medline
18. Heiss JD, Suffredini G, Bakhtian KD, et al. **Normalization of hind-brain morphology after decompression of Chiari malformation type I.** *J Neurosurg* 2012;117:942–46 CrossRef Medline
19. Quon JL, Grant RA, DiLuna ML. **Multimodal evaluation of CSF dynamics following extradural decompression for Chiari malformation type I.** *J Neurosurg Spine* 2015;22:622–30 CrossRef Medline
20. Noudel R, Gomis P, Sotoares G, et al. **Posterior fossa volume increase after surgery for Chiari malformation type I: a quantitative assessment using magnetic resonance imaging and correlations with the treatment response.** *J Neurosurg* 2011;115:647–58 CrossRef Medline
21. Khalsa SS, Siu A, DeFreitas TA, et al. **Comparison of posterior fossa volumes and clinical outcomes after decompression of Chiari malformation type I.** *J Neurosurg Pediatr* 2017;19:511–17 CrossRef Medline
22. Biswas D, Eppelheimer MS, Houston JR, et al. **Quantification of cerebellar crowding in type I Chiari malformation.** *Ann Biomed Eng* 2019;47:731–43 CrossRef Medline
23. Gupta SK, Gahlot S, Singh R, et al. **Spinal tumors and tumor-like masses: relevance of initial imaging, Karnofsky performance status, age, location, and cord edema.** *J Clin Imaging Sci* 2019;9:21 CrossRef Medline
24. Melzack R. **The short-form McGill Pain Questionnaire.** *Pain* 1987;30:191–97 nCrossRef Medline
25. McCormick PC, Torres R, Post KD, et al. **Intramedullary ependymoma of the spinal cord.** *J Neurosurg* 1990;72:523–32 CrossRef Medline
26. Ditunno JF Jr, Young W, Donovan WH, et al. **The international standards booklet for neurological and functional classification of spinal cord injury: American Spinal Injury Association.** *Paraplegia* 1994;32:70–80 Medline
27. Caro CC, Mendes PV, Costa JD, et al. **Independence and cognition post-stroke and its relationship to burden and quality of life of family caregivers.** *Top Stroke Rehabil* 2017;24:194–99 CrossRef Medline
28. Aliaga L, Hekman KE, Yassari R, et al. **A novel scoring system for assessing Chiari malformation type I treatment outcomes.** *Neurosurgery* 2012;70:656–64; discussion 664–65 CrossRef Medline

Progressive Changes in Cerebral Apparent Diffusion Values in Fabry Disease: A 5-Year Follow-up MRI Study

 Koen P.A. Baas,  Albert J. Everard,  Simon Körver,  Laura van Dussen,  Bram F. Coolen,  Gustav J. Strijkers,  Carla E.M. Hollak, and  Aart J. Nederveen



ABSTRACT

BACKGROUND AND PURPOSE: White matter lesions are commonly found in patients with Fabry disease. Existing studies have shown elevated diffusivity in healthy-appearing brain regions that are commonly associated with white matter lesions, suggesting that DWI could help detect white matter lesions at an earlier stage. This study explores whether diffusivity changes precede white matter lesion formation in a cohort of patients with Fabry disease undergoing yearly MR imaging examinations during a 5-year period.

MATERIALS AND METHODS: T1-weighted anatomic, FLAIR, and DWI scans of 48 patients with Fabry disease (23 women; median age, 44 years; range, 15–69 years) were retrospectively included. White matter lesions and tissue probability maps were segmented and, together with ADC maps, were transformed into standard space. ADC values were determined within lesions before and after detection on FLAIR images and compared with normal-appearing white matter ADC. By means of linear mixed-effects modeling, changes in ADC and Δ ADC (relative to normal-appearing white matter) across time were investigated.

RESULTS: ADC was significantly higher within white matter lesions compared with normal-appearing white matter ($P < .01$), even before detection on FLAIR images. ADC and Δ ADC were significantly affected by sex, showing higher values in men (60.1 [95% CI, 23.8–96.3] $\times 10^{-6} \text{ mm}^2/\text{s}$ and 35.1 [95% CI, 6.0–64.2] $\times 10^{-6} \text{ mm}^2/\text{s}$), respectively. Δ ADC increased faster in men compared with women (0.99 [95% CI, 0.27–1.71] $\times 10^{-6} \text{ mm}^2/\text{s/month}$). Δ ADC increased with time even when only considering data from before detection (0.57 [95% CI, 0.01–1.14] $\times 10^{-6} \text{ mm}^2/\text{s/month}$).

CONCLUSIONS: Our results indicate that in Fabry disease, changes in diffusion precede the formation of white matter lesions and that microstructural changes progress faster in men compared with women. These findings suggest that DWI may be of predictive value for white matter lesion formation in Fabry disease.

ABBREVIATIONS: ERT = enzyme replacement therapy; IQR = interquartile range; LME = linear mixed effect; MNI = Montreal Neurological Institute; NAWM = normal-appearing white matter; TP = time point; WML = white matter lesion

Fabry disease is a rare, X-linked lysosomal storage disorder caused by a mutation in the galactosidase alpha (*GLA*) gene.¹ This mutation causes a deficiency of α -galactosidase A activity, resulting in the accumulation of glycosphingolipids, mainly globotriaosylceramide, primarily in the cardiovascular system, brain,

and kidneys.² Accumulation in the arterial walls is believed to contribute to cerebrovascular and cardiac diseases, which are among the leading causes of death in patients with Fabry disease. Among other neurologic findings, white matter lesions (WMLs) are found in nearly one-half of all patients.^{3–5} WML burden increases with age and is associated with cognitive decline.^{5,6} Although the prevalence, severity, and progression of WMLs are comparable in men and women, men develop WMLs at a younger age.^{4,7} Moreover, a higher prevalence of WMLs in men with classic compared with nonclassic Fabry disease has been reported.⁸

The pathophysiology of WMLs in Fabry disease is complex and incompletely understood. Cell damage in the cerebrovascular system may impair the autoregulation of cerebral perfusion, resulting in a hyperdynamic circulation and endothelial dysfunction caused by shear stress and incompressible vessel walls.^{4,9,10} Combined with glycosphingolipid storage in endothelial cells, these effects might cause the release of reactive oxygen species, further compromising

Received May 3, 2023; accepted after revision August 16.

From the Departments of Radiology and Nuclear Medicine (K.P.A.B., A.J.N.), Endocrinology and Metabolism (S.K., L.v.D., C.E.M.H.), and Biomedical Engineering and Physics (B.F.C., G.J.S.), Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, the Netherlands; Faculty of Science (A.J.E.), Vrije Universiteit Amsterdam, Amsterdam, the Netherlands; and Amsterdam Cardiovascular Sciences (B.F.C., G.J.S.), University of Amsterdam, Amsterdam, the Netherlands.

Please address correspondence to Aart J. Nederveen, PhD, Department of Radiology and Nuclear Medicine, Amsterdam University Medical Centers, University of Amsterdam, Meibergdreef 9, 1105 AZ, Amsterdam, the Netherlands; e-mail: a.j.nederveen@amsterdamumc.nl



Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A8001>

endothelial cells and accelerating vascular dysfunction.^{11,12} Patients with Fabry disease with and without WMLs have shown increased CBF but decreased glucose metabolism in the WM, indicating microangiopathy.¹³ Last, increased CSF protein levels have been reported, indicating blood-brain barrier dysfunction.¹⁴

MR imaging is the criterion standard imaging technique to detect brain alterations in Fabry disease, including WMLs. DWI can accurately quantify microstructural WM changes and has been used in patients with Fabry disease.^{5,15,16} These studies showed elevated mean diffusivity in healthy-appearing brain regions that are commonly associated with WMLs.^{5,15,16} Moreover, WM mean diffusivity was found to be correlated with plasma globotriaosylceramide levels.⁵ These findings suggest that DWI could help to detect WML formation at an earlier stage and potentially provide a tool to evaluate treatment options like enzyme replacement therapy (ERT). However, these studies were of a cross-sectional design, prohibiting the assessment of longitudinal changes required to investigate whether changes in diffusion precede WML formation. Therefore, this study aimed to explore whether DWI can detect microstructural changes before WMLs appear on FLAIR-weighted images in a cohort of patients with Fabry disease undergoing yearly MR imaging examinations during a 5-year period.

MATERIALS AND METHODS

For this retrospective, observational study, we adhered to the guidelines of the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) statement. Data collection has been described before.¹⁷ The Amsterdam University Medical Centers (location in the Academic Medical Center) are the national referral center for patients with Fabry disease in the Netherlands. Follow-up at the outpatient clinic ranged from 6 months to once every 2 years and included brain MR imaging scans. Follow-up data from patients with a confirmed Fabry disease diagnosis were collected in a local database after patients provided written informed consent. From this database, follow-up data collected between 2013 and 2019 were extracted. All patients were classified as having the classic or nonclassic phenotype using predefined criteria.^{8,18,19} Patients were included if good-quality FLAIR-weighted, T1-weighted, and DWI scans from at least 3 visits were available. According to Dutch law, no approval of the study protocol was needed because this was a retrospective study and patients were not subjected to procedures or rules of behavior in addition to regular clinical follow-up.

Image Acquisition

MR imaging scans were acquired on 3T scanners (Ingenia; Philips Healthcare). T1-weighted anatomic scans were acquired with the following parameters: TR/TE: 9/4.1 ms, flip angle: 8°, resolution: $0.5 \times 0.5 \times 0.9 \text{ mm}^3$, FOV: $256 \times 256 \times 170 \text{ mm}^3$, and number of averages: 1. DWI scans were acquired with 2 b-values, 0 and 1000 s/mm^2 . Other acquisition parameters were the following: TR/TE: 6340/83 ms, flip angle: 90°, resolution: $0.9 \times 0.9 \times 3.0 \text{ mm}^3$, FOV: $229 \times 229 \times 147 \text{ mm}^3$, and number of averages: 2. FLAIR acquisition parameters were as follows: TR/TE: 4800/356 ms, flip angle: 90°, resolution: $1.04 \times 1.04 \times 1.12 \text{ mm}^3$, FOV: $250 \times 250 \times 360 \text{ mm}^3$, and number of averages: 2. MR imaging software upgrades introduced some variation in acquisition

parameters of the DWI scan (Online Supplemental Data). Periodic image-quality tests using a vendor-provided phantom were performed during the course of the study, including before and after software upgrades. Any observed abnormalities were addressed by the vendor before continuation. Patient positioning and planning of the scans were performed according to standard clinical procedures. FLAIR scans were assessed by a neuroradiologist for WMLs according to the Fazekas scale. The Fazekas scale rates WMLs in the periventricular and deep brain separately from zero (no WMLs) to 3 (severe confluent WMLs) and adds both scores, resulting in a total score that can range from zero to 6.

Postprocessing

Postprocessing was performed in Matlab (Version 2019b; MathWorks). An overview of the processing steps is schematically shown in Fig 1.

T1-weighted scans were segmented into GM, WM, and CSF probability maps using CAT12 (<https://neuro-jena.github.io/cat12-help/>) and nonlinearly registered to standard Montreal Neurological Institute (MNI) space using SPM12 (<http://www.fil.ion.ucl.ac.uk/spm/software/spm12/>) (Fig 1A). After registration to MNI space, tissue probability maps were thresholded at 0.8.

WMLs were automatically segmented on the FLAIR images using the lesion segmentation toolbox (<https://www.applied-statistics.de/lst.html>) lesion-prediction algorithm²⁰ and manually corrected (Fig 1B). This segmentation was performed only on FLAIR images that were scored >0 on the Fazekas score to avoid false-positive detections. WML probability maps were thresholded at 0.2, which was deemed appropriate on the basis of visual inspection (Online Supplemental Data). Corrected WML maps were registered to MNI space.

ADC maps were fitted from the DWI data using the FMRIB Software Library (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>) (Fig 1C) and registered to MNI space using the previously obtained transformation matrix.

After transformation to MNI space, normal-appearing white matter (NAWM) was defined by subtracting the WML masks of every available time point (TP) from the WM masks (Fig 1D). WML and NAWM masks were eroded by 1 voxel, and 3-voxel dilated GM and CSF masks were subtracted from the resulting masks to prevent partial volume effects. Unconnected lesions of <5 voxels were excluded (not shown). The resulting lesion masks were termed “full lesion.” Additionally, masks were created that included only WMLs that developed between 2 study visits, denoted as “new lesion.” New lesion masks were created by subtracting the single-voxel dilated full lesion masks from all previous TPs from the current full lesion mask. Again, lesions of <5 voxels were excluded.

The median ADC within the created masks was calculated at each TP (before and after the first appearance on FLAIR-weighted images; Online Supplemental Data). Naming conventions as in the Online Supplemental Data will be used. The TP of first detection on a FLAIR-weighted image was defined the reference TP (time = 0) for that ROI, to which all ADC values were temporally shifted. The difference between ADC within WMLs and NAWM (ΔADC) was calculated by subtracting the median ADC within the NAWM from the lesion ADC.

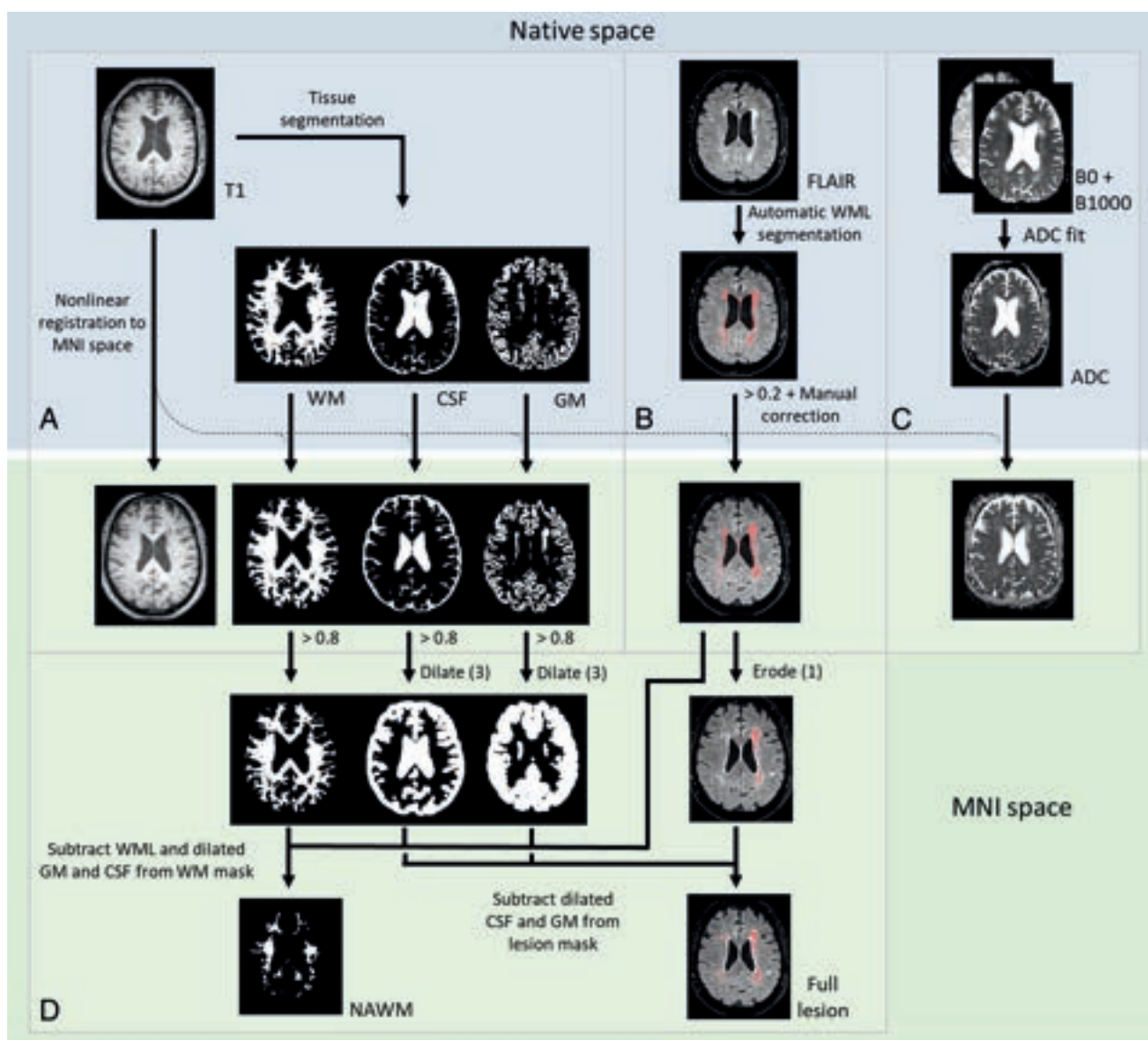


FIG 1. Overview of postprocessing steps. T1-weighted scans were segmented in native space and registered to MNI space (A). WMLs were automatically segmented on FLAIR images, thresholded, and manually corrected (B). ADC maps were fitted from DWIs with b-value = 0 and 1000 s/mm² (C). WML masks and ADC maps were registered to MNI space using the transformation matrix from the T1 registration. NAWM masks were created by subtracting the WML masks from the WM segmentation (D). WML masks were eroded with 2 voxels. Last, GM and CSF masks were dilated with 3 voxels and subtracted from the resulting WML mask and NAWM mask.

Statistical Analysis

Statistical analyses were performed in R Studio (Version 2022.2.3.492; <http://rstudio.org/download/desktop>). Descriptive statistics are reported as mean (SD) or median and interquartile range (IQR) if not normally distributed. ADC values within full and new lesions were compared with corresponding NAWM ADC values using 2-tailed paired *t* tests with a Bonferroni multiple comparison correction.

Linear mixed effect (LME) modeling was used to investigate changes in ADC and Δ ADC with time.²¹ This investigation was only for new lesions because the age of lesions originating from before the first MR imaging scan was unknown. Subjects were assigned as the random effect and TP of lesion segmentation as a crossed random effect to account for multiple TPs per patient

and multiple lesion segmentations per patient, respectively.²² The random effects structure was determined for each model separately as described in the Online Supplemental Data. Next, a priori-determined fixed and interaction effects were added. Fixed effects included time (in months, respectively, to the reference point), sex, and age. Age at baseline was used and centered around the mean. Finally, interaction effects between time and sex and between time and age were included.

$$\begin{aligned} \text{ADC} \sim & \text{Time} + \text{Sex} + \text{Age}_{\text{baseline}} + (\text{Time}|\text{Subject}) \\ & + (1|\text{TP}_{\text{seg}}) + \text{Time} : \text{Sex} + \text{Time} : \text{Age}_{\text{baseline}} \end{aligned}$$

If the model could not converge, the structure was simplified as described in the Online Supplemental Data. Model assumptions

Table 1: Patient characteristics and WML scores

	All	Men	Women
Patient characteristics			
No. of patients (%)	48	25 (52%)	23 (48%)
Classic phenotype (No.) (%)	48 (100%)	25 (100%)	23 (100%)
No. of scan data sets (No.)	229	117	112
Age at first MR imaging (median) (range) (yr)	44 (15–69)	31 (15–55)	46 (22–69)
Patients <18 yr (No.) (%)	2 (4.2%)	2 (8.0%)	0 (0.0%)
Ever ERT (No.) (%)	48 (100%)	25 (100%)	23 (100%)
Months treated (mean)	89 (SD, 34)	92 (SD, 39)	86 (SD, 30)
Events before first MR imaging			
Cerebrovascular event (No.) (%)	2 (4.2%)	1 (4.0%)	1 (4.3%)
TIA (No.) (%)	1 (2.1%)	1 (4.0%)	0 (0.0%)
Kidney function at first MR imaging			
eGFR in mL/min/1.73 m ² (median) (range)	101 (40–154)	105 (49–154)	100 (40–127)
eGFR <60 mL/min/1.73 m ² (No.) (%)	5 (10.4%)	3 (12%)	2 (8.7%)
WML scores			
Fazekas first MR imaging (median) (range)	1 (0–6)	0 (0–6)	1 (0–6)
Fazekas first MR imaging (mean)	1.5 (SD, 1.7)	1.4 (SD, 2.0)	1.5 (SD, 1.3)
Fazekas > 0 first MR imaging (No.) (%)	29 (60%)	11 (44%)	18 (78%)
Fazekas last MR imaging (median) (range) ^a	1 (0–6)	1 (0–6)	1 (0–6)
Fazekas last MR imaging (mean)	1.5 (SD, 1.6)	1.6 (SD, 1.9)	1.4 (SD, 1.3)
Fazekas > 0 last MR imaging (No.) (%)	33 (69%)	15 (60%)	18 (78%)

Note:—eGFR indicates estimated glomerular filtration rate.

^a For 32 patients, additional MR imaging after Fazekas scoring was performed. For these patients, the last known Fazekas score was used.

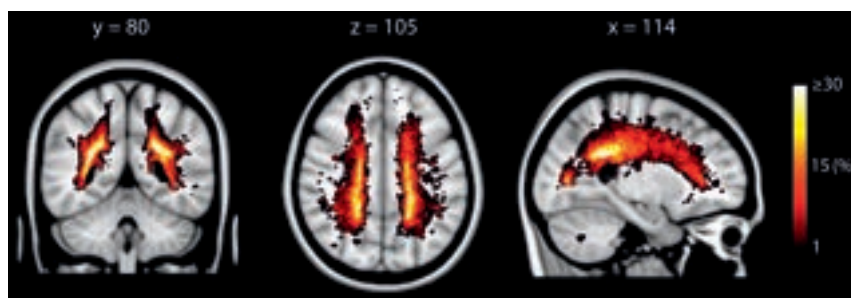


FIG 2. Lesion prevalence in patients with Fabry disease. Y, Z, and X coordinates refer to standard MNI space.

were tested by fitting of the residuals against the observed data and inspecting their normality. Estimated effects and their 95% CIs are reported.

Models were also evaluated using only the ADC values before detection on FLAIR-weighted images (time < 0). To investigate NAWM ADC in patients with and without lesions, we included a grouping variable “presence of lesions” as a fixed effect, indicating whether WMLs, at any TP, had been found in a patient.

Data Availability

Data are available on reasonable request but not publicly available. Because of the rarity of the disease, even anonymized data can be linked to a specific individual.

RESULTS

Data from 49 patients with Fabry disease initially fulfilled the inclusion criteria. Five FLAIR scans and three T1 scans were excluded because of artifacts detected during image analysis, leading to the exclusion of 1 patient because <3 complete TPs were available. After exclusion, 48 patients with Fabry disease (23 women, all classic phenotypes, all on ERT) with a median of 5

TPs (range, 3–6 TPs) were left, totaling 229 included scan data sets (Table 1). The median interval between TPs was 364 days (range, 196–1134 days). At baseline, 29 (60%) patients had WMLs (18 women; median age, 49 years; range, 21–69 years). Four patients without WMLs at baseline developed WMLs during the study, totaling 33 (69%) patients (18 women; median age, 47 years; range, 21–69 years) who had WMLs by the end of the study. In patients who had lesions at baseline, lesion volume increased significantly during the study (median, 0.6 [IQR, 1.6] mm³ and 1.05 [IQR, 3.1] mm³; $Z = 15$, $P < .01$).

WMLs were predominantly found in the periventricular WM (Fig 2). At time point 1 (TP1), the mean ADC within NAWM was 765 (SD, 32) $\times 10^{-6}$ mm²/s compared with 1113 (SD, 138) $\times 10^{-6}$ mm²/s ($t[24] = -14$, $P < .01$) within the full lesion at that TP (full lesion TP1). ADC within full lesion ROIs from all TPs are shown in the Online Supplemental Data.

When we combined new lesion ADC values with the same time respective to their reference point, ADC was significantly higher compared with NAWM up to 5 years before detection on FLAIR-weighted images (Fig 3A; $t[8] = 3.9$, $P = .047$). Similarly, Δ ADC was significantly higher than zero up to 5 years before detection (Fig 3B; $t(8) =$

3.9, $P = .047$). Increased ADC values within WMLs before detection on FLAIR-weighted images also seemed apparent on FLAIR scans and ADC maps of individual patients who developed new WMLs during the study (Fig 4).

LME modeling revealed that sex had a small-but-significant positive fixed effect on NAWM ADC values, showing a 22.1 (95% CI, 3.7–40.4) $\times 10^{-6}$ mm²/s higher ADC intercept for men compared with women ($P = .02$; Online Supplemental Data). Neither time, age, nor the presence of lesions had a significant effect on NAWM ADC.

ADC and Δ ADC within new lesions were significantly affected by sex, with higher intercepts in men compared with women (60.1 [95% CI, 23.8–96.3] $\times 10^{-6}$ mm²/s, $P < .01$ and 35.1 [95% CI, 6.0–64.2] $\times 10^{-6}$ mm²/s, $P = .02$ respectively; Tables 2 and 3). ADC but not Δ ADC was also higher and increased faster in older patients, indicated by the fixed effect of age (2.1 [95% CI, 0.6–3.7] $\times 10^{-6}$ mm²/s/month, $P < .01$) and the interaction effect between age and time (0.04 [95% CI, 0.00–0.07] $\times 10^{-6}$ mm²/s/year/month, $P = .03$). The effects of time, sex, and age on ADC are shown in Fig 5. Δ ADC values increased significantly faster in men compared with women (0.99 [95% CI, 0.27–1.71] $\times 10^{-6}$

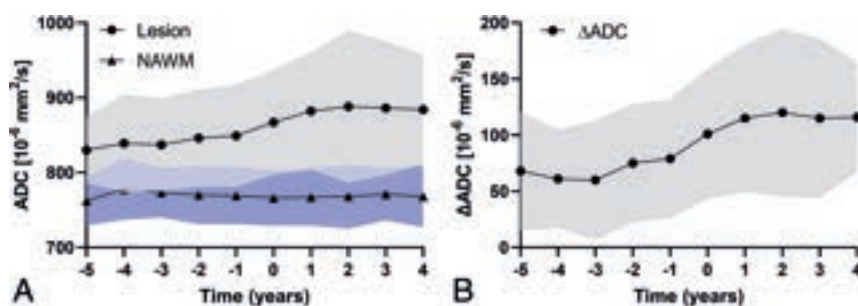


FIG 3. ADC (A) and Δ ADC (B) in new lesion ROIs compared with NAWM. ADC within lesions with the same time (in years) relative to their reference point were combined to calculate mean values (data points) and their SDs (shaded areas). For example, ADC_{TP1} within the new lesion TP2 and ADC_{TP2} within new lesion TP3 both have time = -1 but are compared with the NAWM ADC from TP1 and TP2, respectively. Paired *t* tests showed that the ADC within lesions was significantly higher than the ADC within NAWM and that Δ ADC was significantly higher than zero at every time. NAWM indicates normal appearing white matter.

$mm^2/month$, $P < .01$). Time itself had a significant fixed effect only on Δ ADC (0.49 [95% CI, 0.03 – 0.95] $\times 10^{-6} mm^2/month$, $P = .04$; Online Supplemental Data).

When we evaluated the models using only the TPs before detection, none of the effects describing ADC remained significant, though sex nearly did (35 [95% CI, -3.6 – 74.5] $\times 10^{-6} mm^2$, $P = .075$; Tables 2 and 3). For Δ ADC, only the significant fixed effect of time remained (0.57 [95% CI, 0.01 – 1.14] $\times 10^{-6} mm^2/month$, $P = .047$; Online Supplemental Data).

DISCUSSION

In this 5-year follow-up study, ADC within WMLs of patients with Fabry disease was assessed before and after detection on FLAIR-weighted images in a unique longitudinal data set. ADC values were already increased compared with NAWM in regions that at later TPs were positively identified as WMLs on FLAIR MR imaging. This increase in ADC with time was affected by sex, with men showing a faster increase in ADC than women.

With Fabry disease being an X-linked disease, men are generally more severely affected by it than women. Indeed, this study shows that ADC within WMLs is higher and diverges faster from NAWM values in men compared with women, a finding not observed before.^{7,23,24} However, 1 study reported a comparable WML load between men and women, despite the men being younger, suggesting an earlier onset of WMLs in men.⁷ A different study comparing diffusivity was limited by its small sample size ($n = 27$), cross-sectional setup, and combined inclusion of classic and nonclassic patients.²⁴ Although here only slightly more patients with WMLs were included ($n = 33$), the longitudinal design resulted in 472 observations for the analysis.

Higher ADC values were found in regions confirmed as WML at later TPs, consistent with earlier research showing increased diffusivity in NAWM areas commonly associated with WMLs¹⁶ and indicating that changes in diffusion precede WML detection on FLAIR images. When only the TPs before detection were entered into the model, no effects remained statistically significant for ADC. For Δ ADC, only time remained statistically significant. This finding may be explained by a lack of statistical power to detect changes in ADC while still detecting ADC

increasing faster compared with NAWM. We did not develop a prediction model for WMLs based on ADC but merely showed an association between increased ADC and WML development at later TPs. To determine the predictive value of DWI, a model should be verified using a separate data set. Nevertheless, our findings add to the existing evidence that changes in diffusion precede the formation of WMLs in patients with Fabry disease.

The increasing ADC values with age in WMLs were not observed when considered relative to NAWM (Δ ADC). This finding suggests that the observed effect of age is not WML-specific but rather affects WM globally. However,

time had no significant effect on NAWM ADC. These ambiguous results emphasize the need for caution when interpreting the NAWM ADC results from this patient population. On the other hand, heterogeneous age-related microstructural alterations across the WM²⁵ might prevent the observation of a time effect when considering global WM ADC.

A limitation of this study is the absence of an age- and sex-matched control group. Therefore, ADC within lesions could be compared only with NAWM ADC from other brain regions. NAWM ADC values were similar to literature values from healthy control studies (804 [SD, 110] $\times 10^{-6} mm^2/s$).²⁶ However, it is possible that WM is affected globally in patients with Fabry disease and that areas detected as lesions are just more severely affected than NAWM. This possibility would align with the pathogenesis of patients with MS as observed in studies using magnetization transfer imaging²⁷ and DTI.²⁸ A potential continuation of this study could address this limitation by including a carefully matched healthy control group. Nonetheless, the assessment of diffusion within WMLs both before and after their detection on FLAIR images remained unaffected by the absence of a control group.

All patients included here received ERT, hampering the analysis of the effect of ERT. Treatment duration was explored as a parameter in the LME models but did not yield a statistically significant finding, possibly due to selection bias. Earlier studies reported ambiguous results on the effect of ERT on WMLs. Two studies did not find an effect,^{3,29} while a third study found that WML burden was more likely to remain stable in patients on ERT.³⁰ Although its ability to evaluate ERT is uncertain, this study shows that DWI could still be useful in clinical decision-making. For example, ERT may be delayed until signs of disease progression are observed. In such cases, DWI could serve as a more sensitive marker for detecting early changes. Moreover, if DWI reveals that WMLs continue to progress despite ERT, this finding may be a signal to consider discontinuing the therapy. However, a more thorough validation of the predictive value of DWI must be performed before it can be relied on for such applications.

Some other limitations should be acknowledged. Due to the rarity of the disease, the number of included patients is relatively low, limiting statistical power. However, the longitudinal study

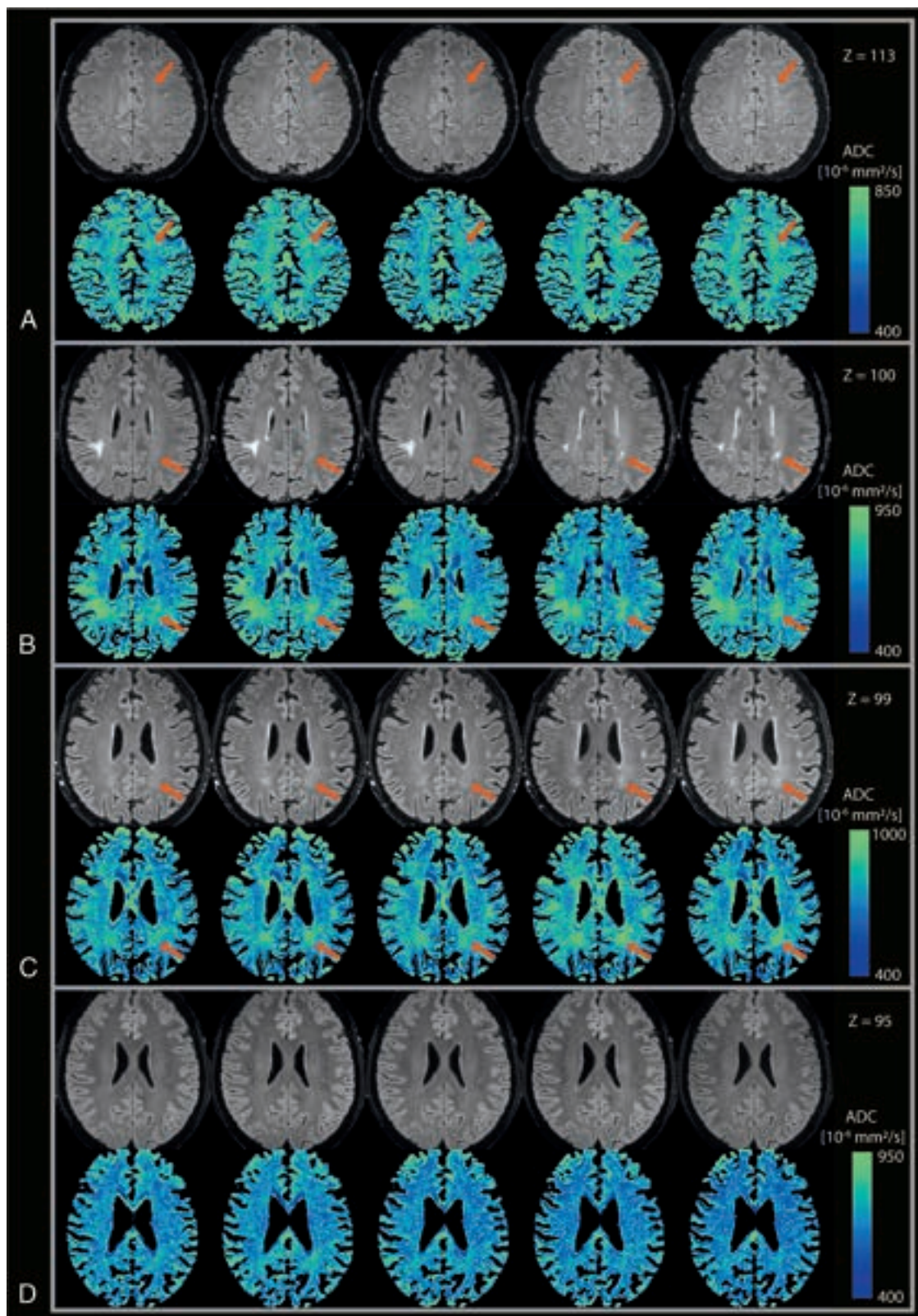


FIG 4. Yearly FLAIR scans and ADC maps. Three of these patients developed WMLs during the study (A–C), while 1 patient did not (D). Scans are shown in chronologic order from left to right. New WMLs are indicated by the *orange arrows*. The patient in B also had a large WM hyperintensity at the right parieto-occipital sulcus, which was diagnosed as a result of infarction before the study. Color scales are optimized for each patient individually. Z values indicate the Z coordinate in standard MNI space.

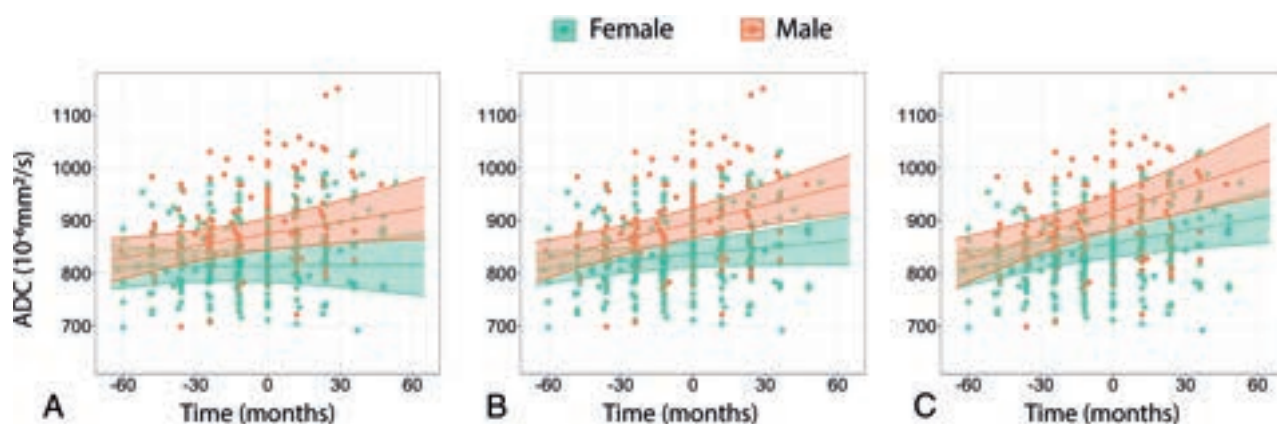


FIG 5. Visualization of LME models describing the ADC within new lesions. The graphs show how time and sex would affect the ADC of 36-, 46-, and 56-year-old patients superimposed on the measured data (A–C respectively). ADC is higher and increases faster in men compared with women and increases faster in older patients.

Table 2: Summary of the obtained LME models describing ADC and Δ ADC

Predictors	ADC			Δ ADC		
	Estimates	CI	P	Estimates	CI	P
(Intercept)	836	810–862	<.01	84.4	63.1–105.7	<.01
Time (mo)	0.41	–0.08–0.91	.10	0.49	0.03–0.95	.04
Male sex	60.1	23.8–96.3	<.01	35.1	6.0–64.2	.02
Age (baseline, mean centered, year)	2.1	0.6–3.7	<.01	1.10	–0.16–2.36	.09
Time:male sex	0.72	–0.05–1.50	.07	0.99	0.27–1.71	<.01
Time:age	0.04	0.00–0.07	.03	0.03	–0.00–0.06	.09
Observations	472			464		
Marginal R ²	0.227			0.199		
Conditional R ²	0.689			0.613		

Table 3: Summary of the obtained LME models describing ADC and Δ ADC before lesion detection on FLAIR-weighted images

Predictors	ADC before Detection			Δ ADC before Detection		
	Estimates	95% CI	P	Estimates	95% CI	P
(Intercept)	834	806–861	<.01	83.2	59.9–106.5	<.01
Time (mo)	0.44	–0.17–1.04	.15	0.57	0.01–1.14	.047
Male sex	35.4	–3.6–74.5	.08	10.7	–21.8–43.1	.52
Age (baseline, mean centered) (yr)	1.0	–0.8–2.8	.25	0.32	–1.22–1.86	.68
Time:male sex	–0.16	–1.10–0.79	.74	0.11	–0.71–0.94	.79
Time:age	–0.00	–0.05–0.05	.85	–0.00	–0.05–0.04	.87
Observations	233			225		
Marginal R ²	0.114			0.035		
Conditional R ²	0.601			0.452		

design partly overcomes this limitation. Given the low sample size, exclusive inclusion of patients with the classic phenotype, and the primary focus on investigating the potential predictive value of DWI, a comparison among patients with different *GLA* gene mutations or phenotypes was not conducted, which remains a topic for future research. Second, because of the large number of MR images, a semiautomated approach was used to segment the WMLs rather than manual segmentation by a trained neuro-radiologist. Last, the acquisition protocol used in this study was not initially designed for a scientific investigation but rather for initial exploration, explaining why a more comprehensive technique like DTI was not used in this patient group.

CONCLUSIONS

This work presents the analysis of a unique, longitudinal data set of DWI and FLAIR MR imaging in patients with Fabry disease. Adding to the existing body of literature, these results indicate that changes in diffusion precede the formation of WMLs, suggesting that DWI may be of predictive value for WML formation in Fabry disease. Moreover, these results show that WML diffusivity progresses faster in men compared with women.

ACKNOWLEDGMENT

ChatGPT (<https://openai.com/chatgpt>) has been used to improve the clarity, grammar, and overall quality of the manuscript.

REFERENCES

1. Fabry H. **Angiokeratoma corporis diffusum: Fabry disease—historical review from the original description to the introduction of enzyme replacement therapy.** *Acta Paediatr* 2002;91:3–5 CrossRef Medline
2. Zarate YA, Hopkin RJ. **Fabry's disease.** *Lancet* 2008;372:1427–35 CrossRef Medline
3. Jardim LB, Aesse F, Vedolin LM, et al. **White matter lesions in Fabry disease before and after enzyme replacement therapy: a 2-year follow-up.** *Arq Neuropsiquiatr* 2006;64:711–17 CrossRef
4. Körver S, Vergouwe M, Hollak CEM, et al. **Development and clinical consequences of white matter lesions in Fabry disease: a systematic review.** *Mol Genet Metab* 2018;125:205–16 CrossRef
5. Ulivi L, Kanber B, Prados F, et al. **White matter integrity correlates with cognition and disease severity in Fabry disease.** *Brain* 2020;143:3331–42 CrossRef Medline
6. Murphy P, Williams F, Davagnanam I, et al. **Cognitive dysfunction and white matter hyperintensities in Fabry disease.** *J Inherit Metab Dis* 2022;45:782–95 CrossRef Medline
7. Rost NS, Cloonan L, Kanakis AS, et al. **Determinants of white matter hyperintensity burden in patients with Fabry disease.** *Neurology* 2016;86:1880–86 CrossRef Medline
8. Arends M, Wanner C, Hughes D, et al. **Characterization of classical and nonclassical Fabry disease: a multicenter study.** *J Am Soc Nephrol* 2017;28:1631–41 CrossRef Medline
9. Fellgiebel A, Keller I, Marin D, et al. **Diagnostic utility of different MRI and MR angiography measures in Fabry disease.** *Neurology* 2009;72:63–68 CrossRef Medline
10. Rombach SM, Twickler TB, Aerts JM, et al. **Vasculopathy in patients with Fabry disease: current controversies and research directions.** *Mol Genet Metab* 2010;99:99–108 CrossRef Medline
11. Wei EP, Kontos HA, Beckman JS. **Mechanisms of cerebral vasodilation by superoxide, hydrogen peroxide, and peroxynitrite.** *Am J Physiol* 1996;271(3 Pt 2):H1262–66 CrossRef Medline
12. Moore DF, Kaneski CR, Askari H, et al. **The cerebral vasculopathy of Fabry disease.** *J Neurol Sci* 2007;257:258–63 CrossRef Medline
13. Moore DF, Altarescu G, Barker WC, et al. **White matter lesions in Fabry disease occur in “prior” selectively hypometabolic and hyperperfused brain regions.** *Brain Res Bull* 2003;62:231–40 CrossRef Medline
14. Böttcher T, Rolfs A, Tanislav C, et al. **Fabry disease: underestimated in the differential diagnosis of multiple sclerosis?** *PLoS One* 2013;8:e71894 CrossRef Medline
15. Albrecht J, Dellani PR, Müller MJ, et al. **Voxel based analyses of diffusion tensor imaging in Fabry disease.** *J Neurol Neurosurg Psychiatry* 2007;78:964–69 CrossRef Medline
16. Paavilainen T, Lepomäki V, Saunavaara J, et al. **Diffusion tensor imaging and brain volumetry in Fabry disease patients.** *Neuroradiology* 2013;55:551–58 CrossRef Medline
17. Körver S, Longo MG, Lima MR, et al. **Determinants of cerebral radiological progression in Fabry disease.** *J Neurol Neurosurg Psychiatry* 2020;91:756–63 CrossRef Medline
18. Tol van der L, Cassiman D, Houge G, et al. **Uncertain diagnosis of Fabry disease in patients with neuropathic pain, angiokeratoma or cornea verticillata: consensus on the approach to diagnosis and follow-up.** *JIMD Rep* 2014;17:83–90 CrossRef Medline
19. Smid BE, Van Der Tol L, Cecchi F, et al. **Uncertain diagnosis of Fabry disease: consensus recommendation on diagnosis in adults with left ventricular hypertrophy and genetic variants of unknown significance.** *Int J Cardiol* 2014;177:400–08 CrossRef Medline
20. Schmidt P. **Bayesian Inference for Structured Additive Regression Models for Large-Scale Problems with Applications to Medical Imaging.** *Mathematics* January 17, 2017. <https://www.semanticscholar.org/paper/Bayesian-inference-for-structured-additive-models-Schmidt/edc315de56ff336b75f36b30bad74fab231d39ca>. Accessed June 20, 2022
21. Bates D, Mächler M, Bolker BM, et al. **Fitting linear mixed-effects models using lme4.** *Journal of Statistical Software* 2015;67 CrossRef
22. Baayen RH, Davidson DJ, Bates DM. **Mixed-effects modeling with crossed random effects for subjects and items.** *Journal of Memory and Language* 2008;59:390–412 CrossRef
23. Fellgiebel A, Müller MJ, Mazanek M, et al. **White matter lesion severity in male and female patients with Fabry disease.** *Neurology* 2005;65:600–02 CrossRef Medline
24. Fellgiebel A, Mazanek M, Whybra C, et al. **Pattern of microstructural brain tissue alterations in Fabry disease: a diffusion-tensor imaging study.** *J Neurol* 2006;253:780–87 CrossRef Medline
25. Burzynska AZ, Preuschhof C, Bäckman L, et al. **Age-related differences in white matter microstructure: region-specific patterns of diffusivity.** *Neuroimage* 2010;49:2104–12 CrossRef Medline
26. Sener RN. **Diffusion MRI: apparent diffusion coefficient (ADC) values in the normal brain and a classification of brain disorders based on ADC values.** *Comput Med Imaging Graph* 2001;25:299–326 CrossRef Medline
27. Filippi M, Campi A, Dousset V, et al. **A magnetization transfer imaging study of normal-appearing white matter in multiple sclerosis.** *Neurology* 1995;45:478–82 CrossRef Medline
28. Ciccarelli O, Werring DJ, Wheeler-Kingshott CA, et al. **Investigation of MS normal-appearing brain using diffusion tensor MRI with clinical correlations.** *Neurology* 2001;56:926–33 CrossRef Medline
29. Jardim L, Vedolin L, Schwartz IV, et al. **CNS involvement in Fabry disease: clinical and imaging studies before and after 12 months of enzyme replacement therapy.** *J Inherit Metab Dis* 2004;27:229–40 CrossRef Medline
30. Fellgiebel A, Gartenschläger M, Wildberger K, et al. **Enzyme replacement therapy stabilized white matter lesion progression in Fabry disease.** *Cerebrovasc Dis* 2014;38:448–56 CrossRef Medline

Partial (SAVE) versus Complete (Solumbra) Stent Retriever Retraction Technique for Mechanical Thrombectomy: A Randomized In Vitro Study

Magda Jablonska, Jiahui Li, Riccardo Tiberi, Pere Canals, Santiago Ortega, Alejandro Tomasello, and Marc Ribo



ABSTRACT

BACKGROUND AND PURPOSE: Mechanical thrombectomy has become a first-line treatment for acute ischemic stroke. Several techniques combining stent retrievers and distal aspiration catheters have been described. We aimed to characterize the efficacy of 2 commonly used techniques according to clot characteristics.

MATERIALS AND METHODS: Soft (mean stiffness = 95.77 [SD, 5.80] kPa) or stiff (mean stiffness = 205.63 [SD, 6.70] kPa) clots (3 × 10 mm and 2 × 10 mm, respectively) were embolized to the distal M1 segment of the MCA in an in vitro model. The technique was randomly allocated (1:1): stent retriever assisted vacuum-locked extraction (SAVE) versus complete retraction (Solumbra). The primary end point was the percentage of first-pass recanalization. Secondary end points were periprocedural distal embolization measures.

RESULTS: A total of 130 mechanical thrombectomies were performed (50 for soft clots and 15 for stiff clots per arm). Overall, the rate of first-pass recanalization was 35% with Solumbra and 15% with SAVE ($P < .01$). For stiff clots, the first-pass recanalization was equal for both methods (27%; $P = 1.00$). With soft clots, the first-pass recanalization was higher with Solumbra (38%) than with SAVE (12%; $P < .01$). When we used soft clots, the maximum embolus size (mean, 1.19 [SD, 0.9] mm versus 2.16 [SD, 1.48] mm; $P < .01$) and total area of emboli (mean, 1.82 [SD, 2.73] versus 3.34 [SD, 3.2]; $P = .01$) were also lower with Solumbra than with SAVE.

CONCLUSIONS: Clot characteristics may influence the efficacy of the thrombectomy technique. In occlusions caused by soft clots, complete retrieval into the distal aspiration catheters achieved higher rates of first-pass recanalization and lower embolization.

ABBREVIATIONS: DAC = distal aspiration catheter; %FPR = percentage of first-pass recanalization; MT = mechanical thrombectomy; SAVE = stent retriever assisted vacuum-locked extraction; Solumbra = total retraction; SR = stent retriever

Stroke is the second leading global cause of mortality, with an annual fatality rate of approximately 3.3 million worldwide.¹ Endovascular treatment has proved its efficacy in patients with acute ischemic stroke due to large-vessel occlusion.^{2–4} First-pass recanalization (FPR), defined as a single pass of the device to achieve complete or successful recanalization without the use of rescue therapy,⁵ has been related to improved functional outcomes and identified as an independent predictor of no disability at 90 days.⁶

A better understanding of the device-clot interactions during all steps of mechanical thrombectomy (MT) is needed to optimize the technique to achieve the highest rates of FPR. In vitro studies have demonstrated that clot composition and biomechanical characteristics may impact device performance and the outcome of a procedure.^{7,8}

Stent retrievers (SRs) have undoubtedly shown their safety and efficacy in the endovascular treatment of stroke.⁹ Moreover, different treatment strategies combining a SR with aspiration through a distal access catheter (DAC) have been proposed to increase recanalization rates. Strategies commonly referred to as stent retriever assisted vacuum-locked extraction (SAVE) and complete retraction (Solumbra) are probably the most frequently used by neurointerventionalists.

The Solumbra technique consists of fully retracting the entire SR inside the DAC under continuous aspiration at the level of the occlusion site before retracting the DAC.¹⁰ Full ingestion of the clot at the occlusion site potentially avoids embolization of the whole clot or fragments to uninvolved arterial territories such as the anterior cerebral artery during retrieval.

Received May 31, 2023; accepted after revision August 17.

From the 2nd Department of Radiology (M.J.), Medical University of Gdansk, Gdansk, Poland; Stroke Unit (M.J., J.L., R.T., P.C., M.R.), Stroke Research (J.L., R.T., P.C., A.T., M.R.), and Neurointerventional Radiology Department (A.T.), Vall d'Hebron University Hospital, Barcelona, Spain; and Department of Neurology (S.O.), The University of Iowa Hospitals and Clinics, Iowa City, Iowa.

Please address correspondence to Marc Ribo, MD, Stroke Unit, Vall d'Hebron University Hospital, Barcelona, Spain, Hospital Vall d'Hebron, Passeig de la Vall d'Hebron, Barcelona 119-129, Spain; e-mail: marc.ribo@hvh.cat



Indicates article with online supplemental data.



Indicates article with supplemental online videos.

<http://dx.doi.org/10.3174/ajnr.A7996>

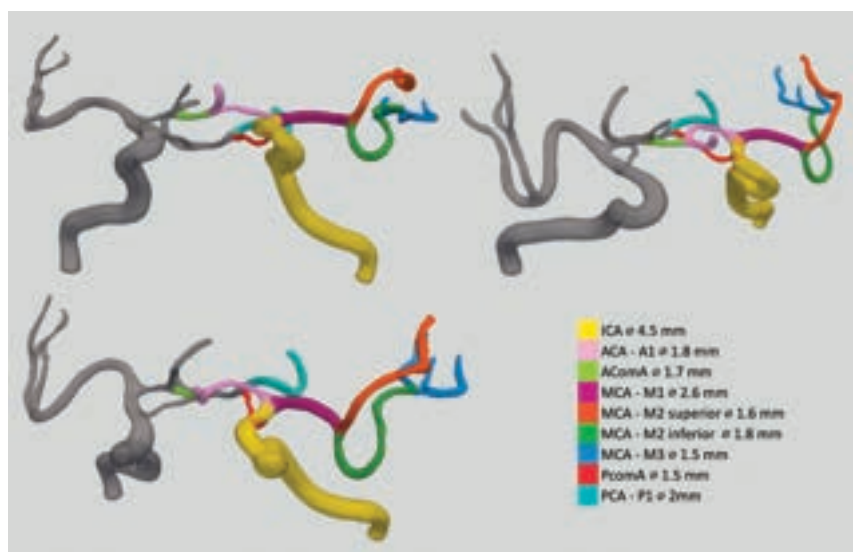


FIG 1. Virtual model of the circle of Willis segmented vessels. ACA indicates anterior cerebral artery; AcomA, anterior communicating artery; PcomA, posterior communicating artery; PCA, posterior cerebral artery.

In contrast, with the SAVE technique, the SR is only partially retracted inside the DAC, with the intention of pinching the clot between the SR and the tip of the DAC; continuous aspiration is also applied to this area to increase clot entrapment.¹¹ Then the whole system is retracted into the guiding catheter under this configuration. The main potential advantage is the reduction of the probability of stripping the clot at the time of the ingestion into the exiting catheter. The clot is never fully retracted into the DAC and is only ingested into the guiding catheter, which has a larger inner diameter at the level of the ICA. Therefore, the probability of clot shaving during this critical step with embolization of the resulting fragments is theoretically reduced.

To date, there is no clear evidence supporting an increased efficacy of one technique over the other, and both are routinely used in clinical practice according to the neurointerventionalist's preferences. We hypothesized that the efficacy of each technique might be related to the mechanical characteristics of the clot.

This randomized, in vitro study aimed to compare 2 MT techniques, SAVE versus Solumbra, in terms of FPR and periprocedural distal embolization.

MATERIALS AND METHODS

Clot Analogs

Two types of clot analogs, namely soft and stiff, were embolized to create an occlusion within the distal M1 segment of the MCA in a 3D-printed neurovascular model. Clot analogs used in the experiments were within the comparable stiffness range of aged thrombus retrieved from patients with stroke (elasticity $[E]_{0\%-45\%} = 170$ [SD, 39] kPa).¹² The Secant modulus of elasticity at 0%–45% strain ($E_{0\%-45\%}$) was a mean of 95.77 [SD, 5.80] kPa and 205.63 [SD, 6.70] kPa for soft and stiff clots, respectively. Three replicates of the Unconfined Compression Test were performed per clot type, as previously described.¹³

The production of the clots comprises the preparation of 2 solutions: for stiff clots, 3% sodium alginate and 40% calcium chloride, and for soft clots, 1% sodium alginate and 2% calcium chloride. The alginate solution was stained using red food coloring to increase the visibility of the clot within the model. Such a prepared alginate solution was subsequently drawn into prepared silicone tubes with a fixed diameter of 3 mm for soft clots and 2 mm for stiff clots. The next step involved casting the alginate solution into the solution with calcium ions and waiting 5 minutes to let the reaction occur. The obtained material was then cut into equal 10-mm-long fragments to create ready-to-use clot analogs.

Neurovascular Model

Experiments were conducted using a 3D-printed neurovascular model (FlowCAT, Barcelona, Spain), realistically reflecting human neurovascular anatomy obtained from anonymized CT scans (Fig 1). The model included a fragment of the descending aorta, aortic arch, bilateral carotid arteries, and a complete circle of Willis. The manufacturing procedure has been previously described.¹⁴ The model functions in a closed circulatory system with the mean water temperature maintained at 36.6 [SD, 1]°C and the continuous flow rate fixed at 800 mL/min (Fig 2).

In Vitro MT Technique

Before each experiment, preprocedural data were recorded, consisting of clot dimensions before embolization, clot length, and location after the initial embolization inside the model.

The clot was inserted into the model through an access port at the proximal portion of the ICA. Flow was subsequently activated in the model, resulting in an embolization of the clot to the distal M1 segment of the MCA. After primary embolization, the experiments were randomized to 1 of the 2 study arms: complete (Solumbra) versus partial (SAVE) stent retrieval into the DAC (1:1). The type of clot was known to the operator.

Initially, 30 in vitro MTs were performed using stiff clots, followed by 30 MTs with soft clots. After a preplanned interim analysis, further experiments with stiff clots were abandoned for futility reasons, and 60 additional experiments with soft clots were performed.

The same MT setup was used for conducting every experiment.

The 8F sheath was plugged into a silicone tube that acts as an extension of the descending aorta to mimic femoral access. A long sheath (Cerebase DA 8F; Cerenovus) was placed in the distal C3 ICA segment. A triaxial system composed of a DAC (React 71; Medtronic), a microcatheter (Phenom 21; Medtronic), and a microguidewire (Synchro 14; Stryker) was advanced. Permanent saline flush lines were connected to the long sheath, DAC, and microcatheter, simulating the clinical setting. After we crossed the clot with the microguidewire and microcatheter into 1 of the

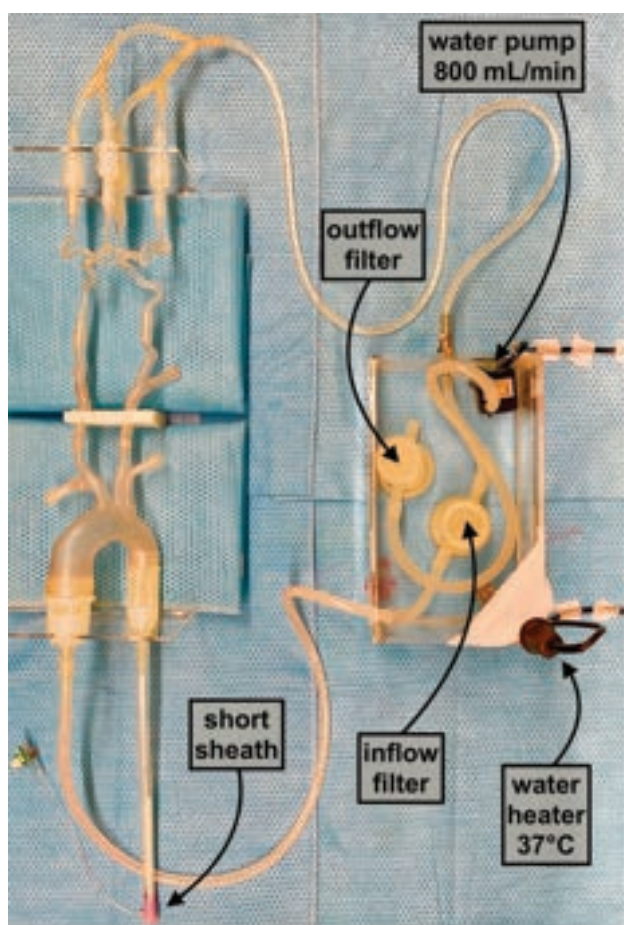


FIG 2. Flow loop setup including the neurovascular model, continuous pump, temperature regulator, and inflow and outflow filters.

M2 MCA branches, the microguidewire was exchanged for an SR (Solitaire 4 × 40; Medtronic). SR deployment was performed using the push-and-fluff technique¹⁵ from the M2 to the M1 MCA segment, aligning the proximal ends of the clot and the stent. The SR (Medtronic, MN) was replaced after every 10 experiments.

After SR deployment, the DAC was advanced up to the proximity of the clot. Manual aspiration with a 60-mL syringe (VacLok; Merit Medical) was initiated through the DAC, and the MT technique was performed according to the study arm allocation. Each experiment was video recorded. Only 1 MT attempt was performed per clot. FPR was defined as no remnant clot fragments of >1 mm either in the model (partial recanalization) or in the filter (distal embolization). Partial recanalization (TICI 2b50) was defined as the presence of particles of >1 mm in the filter but no remnant in the model. After the experiment, the number, size, and location of distal emboli were registered. The primary end point was the percentage of first-pass recanalization (%FPR). Secondary end points were periprocedural distal embolization measures. After each experiment, the entire model and equipment were flushed to ensure no residual contaminants were left.

Filter Analysis

The model included an outflow filter designed to trap particles of >100 μm to precisely evaluate the periprocedural distal

embolization of small particles. Before each experiment, a new filter was installed in the model. After each experiment, the filter was removed, and the contents were stained with Congo red dye to increase the visibility and accurate assessment of the smallest particles. A photograph of each filter was obtained using a high-resolution camera (VZ-R; IPEVO). An independent researcher, who was blinded to all experimental data, analyzed the images by processing an algorithm developed in Matlab R2020a (MathWorks) as previously described.¹⁴ The algorithm automatically provided the following data about emboli collected in the filter: Feret diameter of the largest embolus (maximum [max]-E), total count of emboli (count-E), total count of emboli larger than 1 mm (count-E > 1mm), and total area of emboli (area-E), thus providing a detailed characterization and quantification of distal emboli generated at each pass.

Statistical Analysis

Statistical analyses were performed using SPSS, Version 23.0 software (IBM).

Emboli variables were expressed as mean (SD). The χ^2 test was used to compare both techniques regarding the %FPR. A *t* test was used to analyze the distal emboli parameters. The statistical significance was defined as $P < .05$.

RESULTS

We performed a total of 130 MTs at the level of the left distal M1 MCA segment: 50 cases per arm for soft clots and 15 cases per arm for stiff clots. Overall, FPR was achieved in 25% (33/130) of the cases: 35% (23/65) for Solumbra and 15% (10/65) for SAVE ($P < .01$). Partial recanalization (TICI 2b50) was achieved in 73% (48/65) for Solumbra and 63% (41/65) for SAVE ($P = .186$). Analyses of emboli showed better results with Solumbra technique: Lower maximal embolus diameter (mean, 1.24 [SD, 1] mm versus 2.03 [SD, 1.43] mm; $P < .01$), lower total area of emboli (mean, 1.98 [SD, 3] versus 3.36 [SD, 3.38]; $P = .01$), and lower count E > 1 mm (mean, 1.12 [SD, 1.71] mm versus 1.71 [SD, 1.85] mm; $P = .06$) compared with SAVE (Online Supplemental Data).

Stiff Clots

There was no statistical difference between the 2 techniques in FPR: Solumbra 27% (4/15) versus SAVE 27% (4/15) and partial recanalization (TICI 2b50) (47% [7/15] versus 47% [7/15]; $P = 1$). No significant differences were identified in distal emboli analyses (Online Supplemental Data and Fig 3).

Soft Clots

With soft clots, the Solumbra technique achieved higher rates of FPR (38% [19/50] versus 12% [6/50]; $P < .01$) and partial recanalization (TICI 2b50) (82% [41/50] versus 68% [34/50]; $P = .106$).

The Solumbra technique also showed a lower maximal embolus size (mean, 1.19 [SD, 0.9] mm versus 2.16 [SD, 1.48] mm; $P < .01$) and a lower total area of emboli (mean, 1.82 [SD, 2.73] versus 3.34 [SD, 3.18]; $P = .01$). Additional emboli measures are shown in Fig 3.

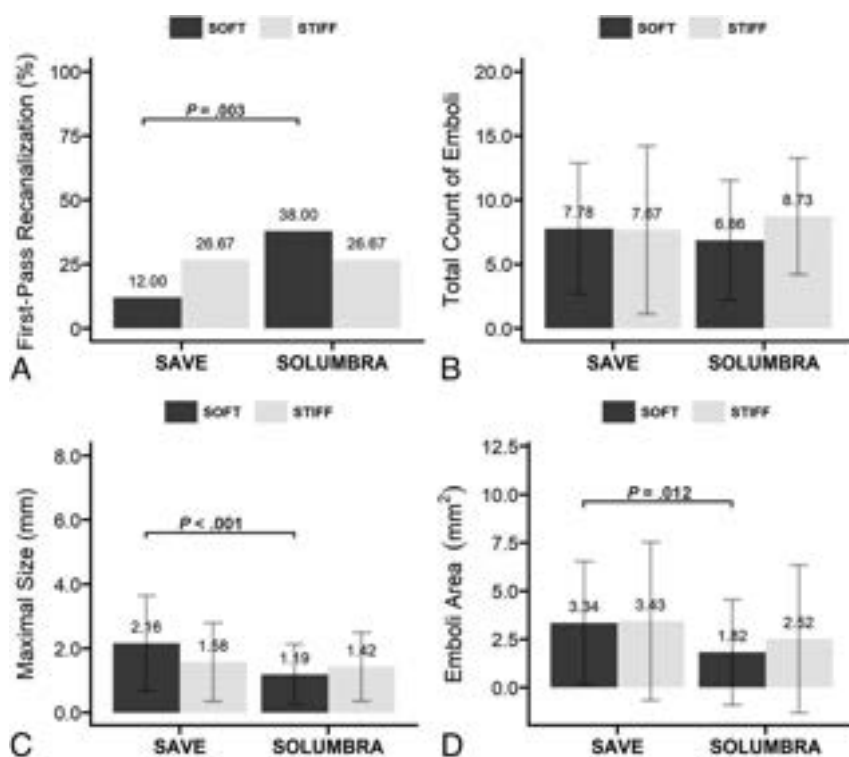


FIG 3. Recanalization rate and distal emboli parameters according to the treatment arm and the type of clot. A, %FPR. B, Total count of emboli (count-E). C, Feret diameter of the largest embolus (max-E). D, Total area of emboli (area-E).

DISCUSSION

This randomized in vitro thrombectomy study assessing 2 different DAC + SR combination techniques demonstrates that clot characteristics influence angiographic outcomes according to the technique used. While complete and partial retrieval of the SR into the DAC have similar efficacy with stiff clots, complete retrieval showed superior outcomes with soft clots.

Several studies indicated that MT outcome and device performance may be influenced by clot composition and its biomechanical characteristics.¹⁶ While soft clots are prone to fragmentation¹⁷ but are more easily ingested into smaller-diameter catheters, stiff clots are characterized by high rigidity, probably less friability, increased resistance to be ingested into the catheters, and higher friction against the vessel wall.¹³ The physical characteristics of the synthetic clot analogs used in the present experiments mimic these features and remain within the reported biomechanical ranges of human clots obtained from thrombectomies.¹² The adjunctive effect of aspiration, generating a protective flow reversal, might also be modified by clot characteristics: Soft clots, easily ingested into the DAC, allow continuous transmission of the negative pressure to the tip of the catheter, generating a sustained local flow reversal. On the other hand, stiff clots that remain stuck or pinched at the tip of the catheter interfere with the transmission of the negative pressure and interrupt the local flow reversal. Thus, small clot fragments generated by the tip of the catheter during ingestion (clot-shaving effect) or while dragging the clot with the SR through the vessel tortuosities (roll-out effect) have higher probabilities of embolizing to distal or new vascular territories. Our hypotheses are confirmed by the frequencies of occurrence of the aforementioned

phenomena for individual MT techniques. Clot shaving was observed more frequently while performing the Solumbra technique than with SAVE (17% versus 13%) (Online Supplemental Data), and the roll-out effect was more common in the SAVE group than in the Solumbra group (20% versus 2%) (Online Supplemental Data).

Given these characteristics, total retrieval of the SR and the clot (Solumbra) at the initial clot location (ie, M1 MCA segment) should theoretically be favored in occlusions caused by soft clots for different reasons. First, a softer clot will still be ingested into the DAC (even with the DAC being smaller than the base catheter), and second, eventual shaved clot fragments will be locally aspirated by the uninterrupted flow reversal (Online Supplemental Data). On the other hand, a partial retrieval of the SR into the DAC (SAVE) seems apparently more adequate for stiff clots because the ingestion of a less deformable clot will be finally made into a larger catheter (the base catheter positioned in the ICA) and the local flow reversal being interrupted

by the clot itself seems less relevant in clots that, by nature, are predominantly nonfriable (Online Supplemental Data).

This hypothesis is mainly supported by the present findings, in which the total retrieval of the SR at the level of the occlusion led to improved FPR rates and a lower distal emboli count, mainly with soft clots. In contrast, when retrieving stiff clots, the efficacy was overall significantly lower than with soft clots, but the 2 studied methods had very similar results. It is possible that if even stiffer clots were to be used, the results might have shown a complete shift, indicating a higher efficacy of the partial SR retrieval technique.

The results of our study may not have a direct impact on clinical practice until the nature and characteristics of the clot cannot be definitively predicted before retrieval. However, indirect recommendations can still be made. Our preferred protocol would be to perform a total retrieval of the SR in the first attempt, considering that softer clots are more prevalent and that in the event that the clot is of stiff, it will remain unfragmented in the original location. If the initial technique fails, the neurointerventionalist could assume that the clot is of a stiffer nature, and a second attempt, partially retrieving the SR, could increase the probability of success. Proceeding inversely with frontline partial retrieval of the SR technique may result in clot fragmentation and embolization of softer clots outside of the DAC during retrieval and before final ingestion into the sheath (Online Supplemental Data).

Additionally, the skills of the interventionalist could be another factor that influences the efficacy, especially when adopting the SAVE technique. The SAVE technique requires that the SR remain partially retracted into the DAC during the whole

maneuver until the system is totally retracted into the base catheter. However, it is frequently observed that due to the elongation of the DAC, which is not paralleled by the SR wire, the SR is progressively (on occasions totally) retracted into the DAC during the operation, increasing the chances of clot shaving and rolling before ingestion into the base catheter (Online Supplemental Data). The frequency of elongation of the DAC during the SAVE technique leading to premature clot ingestion was found to be 22%, being more common for stiff clots compared with soft clots (53% versus 12%). To minimize this phenomenon, the interventionist should release the tension of the DAC as much as possible before initiating the SAVE technique and control the relative position of the SR and the DAC during the whole procedure, steps that require substantial skills and expertise of the interventionist. In our study, there were also unsuccessful attempts for each technique, in which the clot remained in its original location (Solumbra 9%, SAVE 5%; $P = .289$). We provide supplemental videos that demonstrate the mechanisms of MT failure for both techniques (Online Supplemental Data).

Our results are in contrast to those of a recently published in vitro study comparing multiple MT methods in terms of FPR, which demonstrated the superiority of the SAVE technique regarding the %FPR.¹⁸ The differences may be due to the lower number of experiments performed using the present techniques with the same thrombectomy setup (SAVE $n = 6$; Solumbra $n = 6$), a different method of assessing distal emboli, and, especially, the use of a smaller DAC size (0.061 versus 0.071 in the present study), which directly affects aspiration efficacy and clot ingestion.¹⁹ Using a larger-bore DAC predominantly favors the Solumbra technique because the clot is ingested by the DAC only with this technique. Under the SAVE technique, the diameter of the DAC may have a limited impact because clot ingestion occurs at the base catheter. In any case, the contradictory results suggest that further research to define the optimal device combination for each clot type is still required.

Limitations

As in every in vitro study, our experiments have some limitations. Although the anatomy in the neuroanatomic model is reproduced very realistically, the material from which the model is made cannot fully reflect the plasticity of the human vasculature and mimic the interactions among the vessel wall, devices, and the clot. In addition, clot analogs were used in the study, which do not fully reflect the heterogeneous structure and wide range of characteristics of human thrombus.

In our study, we used only 1 type of SR and DAC; thus, the results should not be applied universally and may be different with other device combinations.

Another limitation of our study arises from the use of a single clot size. This choice was motivated by the desire to minimize the accumulation of confounding variables and maintain consistency in our experimental conditions. Nevertheless, we are well-aware that the diverse sizes of clots could substantially impact the effectiveness of the different MT techniques. Thus, we advocate further research in this sense to enrich our understanding and optimize treatment outcomes.

Despite these limitations, given the randomized design, the rates of FPR comparable with the rates achieved in clinical practice, and the large quantity of experiments, it is reasonable to assume that the conditions of our experiment replicated real clinical settings to a high degree.

CONCLUSIONS

Clot characteristics may influence the efficacy of the thrombectomy technique. In occlusions caused by soft clots, complete retrieval of the SR into the DAC achieved higher rates of FPR and lower distal embolization.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

1. Feigin VL, Brainin M, Norrving B, et al. **World Stroke Organization (WSO): Global Stroke Fact Sheet 2022.** *Int J Stroke* 2022;17:18–29 CrossRef Medline
2. Goyal M, Demchuk AM, Menon BK, et al; ESCAPE Trial Investigators. **Randomized assessment of rapid endovascular treatment of ischemic stroke.** *N Engl J Med* 2015;372:1019–30 CrossRef Medline
3. Balamir JS, Sutherland BA, Edmunds LD, et al. **A systematic review and meta-analysis of randomized controlled trials of endovascular thrombectomy compared with best medical treatment for acute ischemic stroke.** *Int J Stroke* 2015;10:1168–78 CrossRef Medline
4. Goyal M, Menon BK, van Zwam WH, et al; HERMES collaborators. **Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials.** *Lancet* 2016;387:1723–31 CrossRef Medline
5. Bai X, Zhang X, Yang W, et al. **Influence of first-pass effect on recanalization outcomes in the era of mechanical thrombectomy: a systematic review and meta-analysis.** *Neuroradiology* 2021;63:795–807 CrossRef Medline
6. Zaidat OO, Castonguay AC, Linfante I, et al. **First-pass effect: a new measure for stroke thrombectomy devices.** *Stroke* 2018;49:660–66 CrossRef Medline
7. Johnson S, Chueh J, Gounis MJ, et al. **Mechanical behavior of in vitro blood clots and the implications for acute ischemic stroke treatment.** *J Neurointerv Surg* 2020;12:853–57 CrossRef Medline
8. Gunning GM, McArdle K, Mirza M, et al. **Clot friction variation with fibrin content; implications for resistance to thrombectomy.** *J Neurointerv Surg* 2018;10:34–38 CrossRef Medline
9. Puñal-Ribóo J, Atienza G, Blanco M. **Safety and efficacy of mechanical thrombectomy using stent retrievers in the endovascular treatment of acute ischaemic stroke: a systematic review.** *Interv Neurol* 2015;3:149–64 CrossRef Medline
10. Munich SA, Vakharia K, Levy EI. **Overview of mechanical thrombectomy techniques.** *Clin Neurosurg* 2019;85:S60–67 CrossRef Medline
11. Maus V, Behme D, Kabbasch C, et al. **Maximizing first-pass complete reperfusion with SAVE.** *Clin Neuroradiol* 2018;28:327–38 CrossRef Medline
12. Chueh JY, Wakhloo AK, Hendricks GH, et al. **Mechanical characterization of thromboemboli in acute ischemic stroke and laboratory embolus analogs.** *AJNR Am J Neuroradiol* 2011;32:1237–44 CrossRef Medline
13. Boodt N, Snouckaert van Schauburg PR, Hund HM, et al. **Mechanical characterization of thrombi retrieved with endovascular thrombectomy in patients with acute ischemic stroke.** *Stroke* 2021;52:2510–17 CrossRef Medline
14. Li J, Tiberi R, Canals P, et al. **Double stent-retriever as the first-line approach in mechanical thrombectomy: a randomized in vitro evaluation.** *J Neurointerv Surg* 2023 Jan 10. [Epub ahead of print] CrossRef Medline
15. Wiesmann M, Brockmann MA, Heringer S, et al. **Active push deployment technique improves stent/vessel-wall interaction in**

- endovascular treatment of acute stroke with stent retrievers. *J Neurointerv Surg* 2017;9:253–56 CrossRef Medline**
16. Jolugbo P, Ariëns RA. **Thrombus composition and efficacy of thrombolysis and thrombectomy in acute ischemic stroke.** *Stroke* 2021;52:1131–42 CrossRef Medline
17. Fereidoonhezahad B, Dwivedi A, Johnson S, et al. **Blood clot fracture properties are dependent on red blood cell and fibrin content.** *Acta Biomater* 2021;127:213–28 CrossRef Medline
18. Gregory G, Federico C, Lefevre PH, et al. **Comparison of mechanical thrombectomy techniques in an in vitro stroke model: how to obtain a first pass recanalization?** *J Neuroradiol* 2022;50:438–43 CrossRef Medline
19. Pérez-García C, et al. **Impact of aspiration catheter size on first-pass effect in the combined use of contact aspiration and stent retriever technique.** *Stroke Vasc Neurol* 2021;6:553–60 CrossRef Medline

Peritumoral Signal on Postcontrast FLAIR Images: Description and Proposed Biomechanism in Vestibular Schwannomas

John C. Benson, Matt L. Carlson, and John I. Lane

ABSTRACT

BACKGROUND AND PURPOSE: Anecdotally, postcontrast FLAIR images of vestibular schwannomas can show peritumoral hyperintense signal, hypothesized to represent gadolinium extravasation. This study assessed the incidence of this phenomenon in a cohort of patients with treatment-naïve sporadic vestibular schwannomas.

MATERIALS AND METHODS: A retrospective review of 20 patients with presumed vestibular schwannoma based on characteristic imaging findings and with dedicated internal auditory canal imaging, including postcontrast T1WI and postcontrast FLAIR, was performed. Tumor size and location were recorded, as was the presence or absence of a fundal fluid cleft. Images were reviewed for the presence of peritumoral hyperintense signal on FLAIR images (a “halo”) and for both subjectively and objectively elevated signal in the ipsilateral cochlea and fundus.

RESULTS: Patients were randomly selected from an institutional vestibular schwannoma registry. Eleven (55.0%) were women. A peritumoral halo was present in 90% of patients, averaging 1.0 (SD, 0.2) mm in thickness. The maximum mean FLAIR signal in the ipsilateral fundus (205.9 [SD, 110.2]) was significantly greater than on the contralateral side (121.6 [SD, 27.8]) ($P = .02$). Maximum mean ipsilateral intracochlear signal (167.8 [SD, 104.5]) was also significantly greater than on the contralateral side (113.4 [SD, 40.1]) ($P = .04$).

CONCLUSIONS: A peritumoral halo on postcontrast FLAIR images was present in 90% of our cohort with randomly selected, treatment-naïve sporadic vestibular schwannomas. Although its mechanism is unknown, this signal is hypothesized to represent gadolinium extravasation, given an ipsilateral increased signal in the adjacent internal auditory canal fundus and cochlea.

ABBREVIATIONS: CPA = cerebellopontine angle; IAC = internal auditory canal; VS = vestibular schwannoma

Vestibular schwannomas (VSs) are benign nerve sheath neoplasms that account for approximately 80% of cerebellopontine angle (CPA) tumors in adults.¹ The lifetime prevalence of developing a VS is 1 in 500.² Even small tumors may lead to ipsilateral nonfunctional hearing and tinnitus, while medium and large-sized vestibular schwannomas may result in additional symptoms such as trigeminal neuropathy or neuralgia or brainstem compression with attendant hydrocephalus.³ VSs that grow with time have been shown to exhibit higher levels of intratumoral inflammation.⁴ Although physical compression from a growing tumor adequately explains the mechanism of trigeminal symptoms and noncommunicating hydrocephalus, even very small tumors that do not fill the cross-diameter of the internal auditory canal (IAC) (ie, do not compress the cochlear

nerve) are often associated with progressive sensorineural hearing loss. Moreover, many studies indicate that the degree of hearing loss is poorly associated with tumor size at diagnosis,⁵ and progressive hearing loss is expected to occur even in tumors that do not exhibit growth.⁶

The imaging appearance of VSs has been extensively reported. In general, the tumors tend to extend from the IAC into the CPA, often with an “ice cream cone” configuration.⁷ VSs typically demonstrate homogeneous tumoral enhancement, though 5%–15% may have intratumoral cystic changes.⁸ Intratumoral hemorrhage, too, may be present; one study found microhemorrhages on T2*-weighted imaging in >90% of tumors.⁹ Compressive edema of the brachium pontis, cerebellum, and/or brainstem is observed in 5% of cases.¹⁰ In the past decade, there has been growing interest in changes of fundal fluid and labyrinthine fluid signal in association with vestibular schwannomas.¹¹ Previous studies have demonstrated that increased FLAIR signal and decreased T2 signal occur in many cases and correlate with the development of progressive hearing loss.¹² There are also studies examining early and delayed gadolinium signal in the fundus and

Received May 15, 2023; accepted after revision August 3.

From the Departments of Radiology (J.C.B., J.I.L.) and Otolaryngology-Head and Neck Surgery (M.L.C.), Mayo Clinic, Rochester, Minnesota.

Please address correspondence to John C. Benson, Department of Radiology, Mayo Clinic, 200 1st St SW, Rochester, MN 55905; e-mail: Benson.john3@mayo.edu

<http://dx.doi.org/10.3174/ajnr.A7979>

labyrinth.¹³ The mechanisms behind this phenomenon remain unknown but may include buildup of endogenous proteins and metabolites (“biologic pollution”) from trapped CSF, increased permeability of the blood-labyrinthine barrier, activation of an inflammatory cascade, and/or tumor “shedding” or “leaking.”

Anecdotally, peritumoral signal has been observed around VSs on postcontrast 3D FLAIR imaging. The incidence and characterization of this finding have not yet been assessed in a systematic study, however, and the mechanism for this finding also remains unexplored. This study sought to further define this radiologic finding by evaluating peritumoral FLAIR signal in a cohort of patients with treatment-naïve sporadic vestibular schwannomas.

MATERIALS AND METHODS

Patient Selection

This study was performed following approval by the local institutional review board. A retrospective radiologic review was completed of a random sample of 20 patients diagnosed with a presumed vestibular schwannoma (based on imaging findings). The size of the patient cohort ($n = 20$) was chosen as a number deemed sufficient to achieve the goal of this study: to confirm anecdotal evidence of the presence of peritumoral post-FLAIR signal in patients with treatment-naïve VSs to establish it as a potential biomarker for future study. Neither the size nor the appearance of VSs was used as an inclusion criterion for the studied cohort.

All patients underwent dedicated IAC MR imaging between October 11, 2017, and November 23, 2021. No examinations required exclusion for artifacts such as motion. In addition, to assess precontrast FLAIR images in the same patient cohort, we completed a review of MR imaging performed before the included study with precontrast FLAIR images performed through the IAC. As noted in the results, precontrast FLAIR images were available for evaluation in 16/20 patients. Demographic information was assessed using a review of the electronic medical record.

MR Imaging Parameters

A 3T Siemens scanner (Magnetom Prisma) with multichannel phased array coils was used to image all patients (with either a 32- or a 64-channel head coil). Dedicated IAC imaging was performed using axial 3D T1 sampling perfection with application-optimized contrasts by using different flip angle evolution (SPACE sequence; Siemens) (TR = 600 ms, TE = 32 ms, data matrix = 192×192 , acquisition time = 4 minutes 0 second), axial 3D T2 SPACE (TR = 1300 ms, TE = 184 ms, data matrix = 320×320 , acquisition time = 3 minutes 55 seconds), axial 3D fat-saturated postcontrast T1 SPACE (TR = 600 ms, TE = 32 ms, data matrix = 192×192 , acquisition time = 4 minutes 0 second), and axial 3D postcontrast FLAIR (TR = 5000 ms, TE = 379 ms, data matrix = 192×192 , acquisition time = 4 minutes 29 seconds). The FOV for each sequence was 150 mm.

Imaging Review

A blinded retrospective review was performed by 2 Certificate of Added Qualification–certified neuroradiologists (J.C.B., J.I.L.). All VSs were evaluated for laterality, location, and size. Location

was categorized as being completely within the IAC, completely in the CPA, or in both the IAC and CPA. Size was based on the largest single axial dimension. For VSs located in both the IAC and CPA, only the largest dimension of the CPA component was measured, in accordance with reporting guidelines.¹⁴ Fundal cleft size was assessed by measuring between the lateral-most aspect of the VS and the IAC fundus. Obstruction of the fundus from the remainder of the IAC and/or CPA by the VS was assessed using axial T2 SPACE images; obstruction was defined as a region of complete occlusion of the IAC by the tumor.

The presence or absence of peritumoral signal on FLAIR images (here referred to as the peritumoral “halo”) was assessed by fusing and overlaying axial postcontrast FLAIR images with both axial fat-saturated postcontrast T1W1 and axial T2 SPACE images using the institutional PACS system. The difference in tumor size between these images was used to measure the thickness of the halo, if present. Image subtraction was also used to better define the differences between both the tumor margins and the halo. Only 1 MR imaging examination was assessed per patient; changes across serial examinations were not evaluated as part of this study.

In addition, the maximum signal intensity of specified areas was measured on postcontrast FLAIR images: the ipsilateral fundus (if a cleft was present) and the ipsilateral cochlea. Free-form ROIs were used for each measurement to precisely outline these anatomic structures. The contralateral fundus and cochlea were also measured to provide internal reference. The ipsilateral and contralateral signal intensities were compared using the statistical methods outlined below. Each measurement was obtained using our institution’s PACS system.

Any interobserver disagreements regarding the presence or absence of a halo or abnormal signal in the fundus were resolved by consensus. For the measured thickness of the halo, the average measurement between observers was used.

Statistical Analyses

Means and SDs were calculated for all continuous variables. Statistical calculations were performed using BlueSky Statistics software (<https://www.blueskystatistics.com/>). The Student t test was used to calculate statistical differences among continuous variables. The Cohen κ was used to calculate interobserver agreement with categorical variables, and Bland-Altman plots were used to calculate interobserver agreement with continuous variables. The threshold for statistical significance was set at $P < .05$.

RESULTS

Twenty patients were included; 11/20 (55.0%) were women, and the average age was 69.5 years (SD, 10.2). Eleven VSs (55.0%) were located on the right; 16/20 tumors (80.0%) were located entirely within the IAC, while 4 exhibited CPA extension of the tumor. None of the observed VSs were isolated to the CPA. The average size of all tumors was 5.8 (SD, 3.1) mm (Table). A fundal cleft was present in 16 patients (80.0%). The average size of the fundal cleft, when present, was 2.3 (SD, 1.8) mm. Of the patients with a fundal cleft, the VS obstructed the fundus in 5 patients (38.5%). In patients with a nonobstructed fundus, the tumor was either smaller than the diameter of the IAC or the tumor was centered along one wall of the IAC with a patent CSF cleft on the other side.

Hyperintense peritumoral signal (halo) was present on postcontrast FLAIR images in 90% of patients. In general, the observed halo was nonuniform in thickness around the circumference of the tumors (Fig 1). The halo itself was difficult to confirm as being outside the tumor borders without the use of overlay/fused images in the PACS system. The average maximum thickness of the halo in the axial plane was 1.0 (SD, 0.2) mm. Interobserver agreement regarding the presence or absence of a halo was nearly perfect ($\kappa = 0.95$), while the interobserver correlation coefficient of the halo thickness was 0.38.

Regarding the 2 tumors that lacked a halo, one was 14 mm and the other was 8 mm. The 14-mm tumor lacked a fundal cleft, potentially affecting the ability to visualize a peritumoral halo. Both fusion and subtraction on the 8-mm tumor were somewhat difficult, possibly affecting the ability to visualize a halo in that lesion.

Sixteen patients had prior MR imaging with precontrast FLAIR imaging of the IAC available for review (the average time between

prior and current MR imaging was 62.9 [SD, 24.6] months). None of these studies (0%) showed abnormal peritumoral signal on precontrast FLAIR images.

Of patients with a fundal cleft, hyperintense signal was observed in the fundus on postcontrast FLAIR images in 13 of 16 (82.3%). Ten of those 13 patients had tumoral obstruction of the fundus. Interobserver agreement for observation of a fundal cleft signal was moderate ($\kappa = 0.46$). On average, the maximum signal in the ipsilateral fundus on postcontrast FLAIR images was 205.9 (SD, 110.2), which was significantly greater than the average signal in the contralateral fundus (121.6 [SD, 27.8]) ($P = .02$).

Pathologically increased FLAIR signal was subjectively observed in the ipsilateral cochleae structures in 10 patients (50.0%) (Fig 2). Of those 10 patients, the fundal cleft was obstructed in 6. On postcontrast FLAIR images, the mean measured maximum intracochlear signal (167.8 [SD, 104.5]), on average, was greater than contralateral cochlear signal (113.4 [SD, 40.1]) ($P = .04$).

Sizes and incidence of multiple imaging findings

Findings	
Size	
Vestibular schwannoma (mean) (mm)	5.8 (SD, 3.1)
Fundal cleft (mean) (mm)	2.3 (SD, 1.8)
Peritumoral halo (mean) (mm)	1.0 (SD, 0.2)
Incidence	
Peritumoral halo (No.) (%)	18/20 (90%)
Hyperintense signal in fundal cleft (No.) (%)	13/16 (82.3%)
Elevated intracochlear signal on postcontrast FLAIR (No.) (%)	10/20 (50.0%)

DISCUSSION

This study assessed the incidence of peritumoral signal on postcontrast FLAIR images in a cohort of patients with vestibular schwannoma. We found that a previously undescribed peritumoral halo was observed in a high proportion of cases, typically measuring about 1 mm in maximum

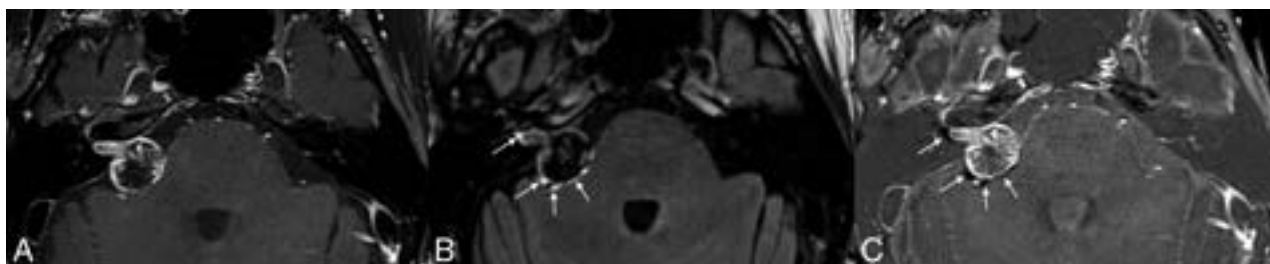


FIG 1. Example of a peritumoral halo on postcontrast FLAIR images. Postcontrast fat-saturated T1-weighted image (A) demonstrates a vestibular schwannoma extending from the right IAC into the CPA. Postcontrast FLAIR image (B) shows a discontinuous peritumoral hyperintense halo that extends beyond the tumor margins on postcontrast T1-weighted images (arrows). Corresponding subtraction FLAIR image (C) confirms the presence of a halo (arrows).

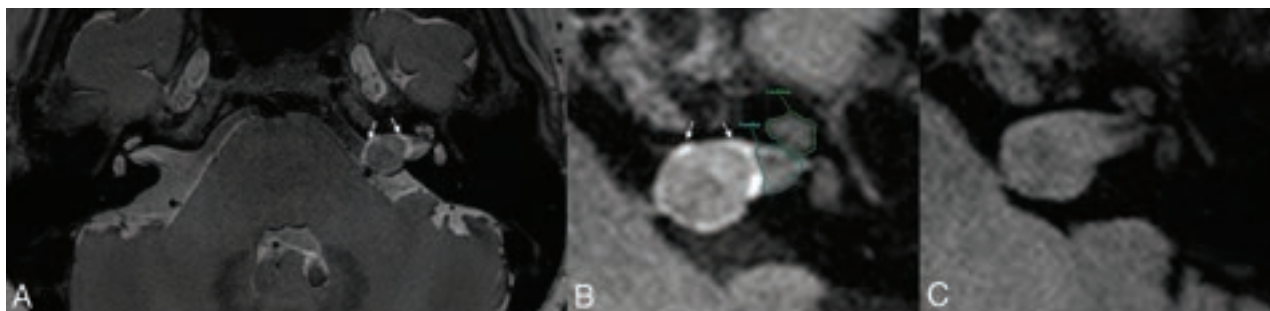


FIG 2. Multiple findings of possible gadolinium leakage from a vestibular schwannoma. Fused axial T2 SPACE and postcontrast T2 FLAIR image (A) shows a vestibular schwannoma centered in the left porus acusticus with a peritumoral halo (arrows) that extends past the solid tumor borders. Hyperintense signal was observed in the ipsilateral IAC fundus and asymmetrically elevated signal in the ipsilateral cochlea on the postcontrast T2 FLAIR image (B). Noncontrast FLAIR image on a prior study (C) does not show abnormal peritumoral signal, indicating that this finding should not represent peritumoral edema.

thickness. This halo is difficult to detect without the use of coregistered/fused images or image subtraction. Asymmetrically elevated postcontrast FLAIR signal was also present in both the fundus and ipsilateral cochlea, confirming the findings of prior studies.

The most logical explanation for the observed findings is that the hyperintense halo represents extratumoral leakage of gadolinium, because it was not demonstrated on any precontrast FLAIR images. The presence of halos on postcontrast FLAIR images, but not postcontrast T1WI, could be due to ≥ 1 factor. It is possible that leakage is better visualized on FLAIR images because of superior sensitivity to the T1-shortening effect of gadolinium.^{15,16} It is also possible that the FLAIR images better depicted gadolinium leakage because they were acquired later in the imaging protocol, allowing interval leaching of contrast into the peritumoral space.

Because the halos were nonuniform in appearance and not circumferential peritumoral rings, we hypothesize that the early gadolinium leakage becomes trapped in a peritumoral space created by the enveloping arachnoid or within cysts and/or adhesions. As seen during microsurgical tumor dissection, VSs are surrounded by a thin, irregular, arachnoid layer that is variably adherent to the tumor surface. Additionally, peritumoral cysts have been well-documented previously,¹⁷ and vestibular schwannomas specifically are known to exist in an inflammatory microenvironment.¹⁸ Thus gadolinium may have become caught in the veil of the peritumoral arachnoid or peritumoral cysts before leaking further into the adjacent CSF. Elevated signal in the ipsilateral fundus and cochlea suggests that additional leakage does occur, possibly in a more diluted manner than the halo effect.

It is not clear why the leaked contrast seems to be locally transported and/or resorbed in the ipsilateral cochlea, even in cases in which the IAC is not occluded. It is possible that even partial occlusion of the IAC (or underlying anatomic constraints of the IAC structure) leads to somewhat stagnant flow of leaked gadolinium into the adjacent CSF, which then, preferentially, is concentrated in the adjacent cochlea rather than diffusing more broadly into the CSF. Altogether, the precise pathophysiology of the observed findings is likely even more complex than what we have outlined, and it is possible that the halo sign is not specific to VS. Further studies using larger cohorts and a greater variety of tumors should help to develop a more comprehensive model.

The concept of VS contrast extravasation is not new. VSs have previously been hypothesized to leak gadolinium into the adjacent structures, most notably the ipsilateral cochlea. Yamazaki et al,¹⁹ for example, found that ipsilateral cochleae have higher signal intensity on postcontrast FLAIR. In that study, the authors demonstrated a convincing increase in intracochlear signal intensity on postcontrast FLAIR images compared with precontrast images, and gadolinium leakage into the IAC fundus. This increased ipsilateral cochlear signal on postcontrast FLAIR signal has been assessed elsewhere.²⁰ The authors contended that the intracochlear signal was suggestive of increased permeability of the blood-labyrinthine barrier in the setting of a VS.¹⁹ Most interesting, Fig 3 in the article of Yamazaki et al also seems to show a peritumoral halo on postcontrast FLAIR images, though neither fused nor subtraction images were shown. The finding was not discussed by the authors.

Pathologically elevated signal is known to be present in the ipsilateral cochlea of patients with VSs on precontrast FLAIR imaging. Unlike the cochlear enhancement observed by Yamazaki et al,¹⁹ this intrinsic T2/FLAIR hyperintensity in the cochlea is attributed to elevated well-documented protein in the perilymph.^{21,22} Historically, labyrinthine taps to test for proteinaceous fluid were used as a diagnostic test for VSs.²³ Although the cause of increased perilymphatic protein remains unknown, the mechanism favored by many authors is the blockage of neuroaxonal protein transportation, related to compression of the vestibulocochlear nerve.²¹

Conceivably, the detection of elevated protein seen on FLAIR and gadolinium in the fundus and labyrinth may be related to the tumor obstructing the IAC and creating a relatively stagnant small reservoir of CSF. In this setting, low levels of leached protein or contrast from the tumor may accumulate to sufficient levels to be detected. This possibility may explain the relative gradient seen in FLAIR and gadolinium, with higher concentrations often occurring in the fundus adjacent to the tumor compared with the labyrinth. In contrast, any tumor or contrast leaked medially into the relatively large CPA cistern will naturally dilute and recirculate in the ~ 125 mL of CSF in the remaining subarachnoid space, which turns over approximately 3 times daily. Prior literature has identified elevated levels of circulating CSF protein in individuals with small vestibular schwannomas and associated communicating hydrocephalus, implicating protein sludging as a mechanism for reduced CSF resorption.²⁴

Finally, a very recent publication proposed that peritumoral “rim” enhancement on FLAIR could help differentiate meningiomas from malignant dural-based tumors.²⁵ The rim of peritumoral contrast shown in that study is very similar to the halo we describe in the current findings. The authors of that study opined that the observed findings were due to the dual blood supply (meningeal and pial) of meningiomas. This supply led to relative intratumoral signal loss on FLAIR (due to T2 shortening in the setting of high gadolinium concentrations) and relatively hyperintense signal in the pial-supplied tumor capsule. The results of the current study, however, seem to refute the findings of this recent publication. Most important, vestibular schwannomas are nonencapsulated tumors and do not have a dual blood supply.²⁶ Also, as previously stated, the halos observed in the current study are irregular and do not appear to conform to a smooth capsule. Last, enhancement of a tumor capsule would not explain the more diffuse gradient that seems to occur in our cohort, in which gadolinium appears to accumulate in the fundus and cochlea. Together, we believe that these findings are evidence of gadolinium leakage, as stated above.

The results of this current study leave many questions unanswered, and future studies will be needed for further exploration. For example, the incidence and/or appearance of peritumoral halos may vary among patients with sporadic VSs compared with VSs related to neurofibromatosis type 2, or in untreated-versus-treated tumors. It may also be useful to determine whether the presence of a peritumoral halo predicts the stability or growth of a tumor. At present, these topics are outside of the scope of the current study.

This study has several limitations, including being a retrospective review. The goal of defining the radiologic finding of a

peritumoral halo is somewhat limited in scope, and the results of this study are preliminary and based on a small cohort. Additional studies using large sample sizes will be needed to confirm these findings, as well as to assess the clinical relevance. In addition, many of the VSs analyzed in this study were relatively small. Future studies will also be needed to address whether tumor size has an impact on the thickness and/or appearance of a peritumoral halo. Also, the precontrast FLAIR images that were available for review were part of separate MR imaging examinations. Thus, no precontrast FLAIR imaging was available for direct comparison with postcontrast FLAIR in this study. Finally, the proposed mechanisms for the observed peritumoral halo are hypothetical and based on limited prior studies. It is possible that the peritumoral signal on postcontrast FLAIR images simply represents an imaging artifact. Also, because the average halo thickness was around 1 mm, the precision of the reported measurements may be questionable, given the pixel size on standard PACS calipers.

CONCLUSIONS

A peritumoral halo on postcontrast FLAIR images was observed in nearly all vestibular schwannomas in our cohort of 20 patients, usually measuring about 1 mm in thickness. Although the pathophysiologic mechanism of this finding is uncertain, we hypothesize that this halo may represent local leakage of gadolinium into the peritumoral space.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

- Lin EP, Crane BT. **The management and imaging of vestibular schwannomas.** *AJNR Am J Neuroradiol* 2017;38:2034–43 CrossRef Medline
- Marinelli JP, Beeler CJ, Carlson ML, et al. **Global incidence of sporadic vestibular schwannoma: a systematic review.** *Otolaryngol Head Neck Surg* 2022;167:209–14 CrossRef Medline
- Fukuda M, Oishi M, Kawaguchi T, et al. **Etiopathological factors related to hydrocephalus associated with vestibular schwannoma.** *Neurosurgery* 2007;61:1186–92; discussion 1192–93 CrossRef Medline
- Lewis D, Roncaroli F, Agushi E, et al. **Inflammation and vascular permeability correlate with growth in sporadic vestibular schwannoma.** *Neuro Oncol* 2019;21:314–25 CrossRef Medline
- Gurewitz J, Schnurman Z, Nakamura A, et al. **Hearing loss and volumetric growth rate in untreated vestibular schwannoma.** *J Neurosurg* 2022;136:768–75 CrossRef Medline
- Luryi AL, Babu S, Bojrab DI, et al. **Progression of hearing loss in observed non-growing vestibular schwannoma.** *Otol Neurotol* 2022;43:e767–72 CrossRef Medline
- Connor SE. **Imaging of the vestibular schwannoma: diagnosis, monitoring, and treatment planning.** *Neuroimaging Clin N Am* 2021;31:451–71 CrossRef Medline
- Bonneville F, Savatovsky J, Chiras J. **Imaging of cerebellopontine angle lesions: an update, Part 1: enhancing extra-axial lesions.** *Eur Radiol* 2007;17:2472–78 CrossRef Medline
- Thamburaj K, Radhakrishnan VV, Thomas B, et al. **Intratumoral microhemorrhages on T2*-weighted gradient-echo imaging helps differentiate vestibular schwannoma from meningioma.** *AJNR Am J Neuroradiol* 2008;29:552–57 CrossRef Medline
- Giordano M, Gerganov V, Metwali H, et al. **Imaging features and classification of peritumoral edema in vestibular schwannoma.** *Neuroradiol J* 2020;33:169–73 CrossRef Medline
- Welby JP, Benson JC, Lohse CM, et al. **Increased labyrinthine T1 postgadolinium signal intensity is associated with the degree of ipsilateral sensorineural hearing loss in patients with sporadic vestibular schwannoma.** *AJNR Am J Neuroradiol* 2023;44:317–22 CrossRef Medline
- Cass ND, Fan Y, Lindquist NR, et al. **Automated whole cochlear T2 signal demonstrates weak correlation with hearing loss in observed vestibular schwannoma.** *Audiol Neurotol* 2023;11 CrossRef Medline
- Bowen AJ, Carlson ML, Lane JL. **Inner ear enhancement with delayed 3D-FLAIR MRI imaging in vestibular schwannoma.** *Otol Neurotol* 2020;41:1274–79 CrossRef Medline
- Committee on Hearing and Equilibrium guidelines for the evaluation of hearing preservation in acoustic neuroma (vestibular schwannoma); American Academy of Otolaryngology-Head and Neck Surgery Foundation, INC.** *Otolaryngol Head Neck Surg* 1995;113:179–80 CrossRef Medline
- Mathews VP, Caldemeyer KS, Lowe MJ, et al. **Brain: gadolinium-enhanced fast fluid-attenuated inversion-recovery MR imaging.** *Radiology* 1999;211:257–63 CrossRef Medline
- Mahale A, Choudhary S, Ullal S, et al. **Postcontrast fluid-attenuated inversion recovery (FLAIR) sequence MR imaging in detecting intracranial pathology.** *Radiol Res Pract* 2020;2020:8853597 CrossRef Medline
- Herde RF, Hoang N, Tran DK, et al. **Peritumoral cysts associated with pituitary macroadenoma.** *J Neurosurg* 2015;123:789–93 CrossRef Medline
- Hannan CJ, Lewis D, O'Leary C, et al. **The inflammatory microenvironment in vestibular schwannoma.** *Neurooncol Adv* 2020;2:vdaa023 CrossRef Medline
- Yamazaki M, Naganawa S, Kawai H, et al. **Increased signal intensity of the cochlea on pre- and post-contrast enhanced 3D-FLAIR in patients with vestibular schwannoma.** *Neuroradiology* 2009;51:855–63 CrossRef Medline
- Naganawa S, Komada T, Fukatsu H, et al. **Observation of contrast enhancement in the cochlear fluid space of healthy subjects using a 3D-FLAIR sequence at 3 Tesla.** *Eur Radiol* 2006;16:733–37 CrossRef Medline
- Kim DY, Lee JH, Goh MJ, et al. **Clinical significance of an increased cochlear 3D fluid-attenuated inversion recovery signal intensity on an MR imaging examination in patients with acoustic neuroma.** *AJNR Am J Neuroradiol* 2014;35:1825–29 CrossRef Medline
- Bhadelia RA, Tedesco KL, Hwang S, et al. **Increased cochlear fluid-attenuated inversion recovery signal in patients with vestibular schwannoma.** *AJNR Am J Neuroradiol* 2008;29:720–23 CrossRef Medline
- Silverstein H. **Labyrinthine tap as a diagnostic test for acoustic neuroma.** *Otolaryngol Clin North Am* 1973;6:229–44
- Al Hinai Q, Zeitouni A, Sirhan D, et al. **Communicating hydrocephalus and vestibular schwannomas: etiology, treatment, and long-term follow-up.** *J Neurol Surg B Skull Base* 2013;74:68–74 CrossRef Medline
- Panyaping T, Pimpichet M, Tunlayadechanont P, et al. **Usefulness of a rim-enhancing pattern on the contrast-enhanced 3D-FLAIR sequence and MRI characteristics for distinguishing meningioma and malignant dural-based tumor.** *AJNR Am J Neuroradiol* 2023;44:247–53 CrossRef Medline
- Kuo TC, Jackler RK, Wong K, et al. **Are acoustic neuromas encapsulated tumors?** *Otolaryngol Head Neck Surg* 1997;117:606–09 CrossRef Medline

Differentiation between Chondrosarcoma and Synovial Chondromatosis of the Temporomandibular Joint Using CT and MR Imaging

B.G. Jang, K.H. Huh, H.G. Yeom, J.H. Kang, J.E. Kim, H.J. Yoon, W.J. Yi, M.S. Heo, and S.S. Lee



ABSTRACT

BACKGROUND AND PURPOSE: Chondrosarcoma and synovial chondromatosis of the temporomandibular joint share overlapping clinical and histopathologic features. We aimed to identify CT and MR imaging features to differentiate chondrosarcoma from synovial chondromatosis of the temporomandibular joint.

MATERIALS AND METHODS: The CT and MR images of 12 and 35 patients with histopathologically confirmed chondrosarcoma and synovial chondromatosis of the temporomandibular joint, respectively, were retrospectively reviewed. Imaging features including lesion size, center, enhancement, destruction/sclerosis of surrounding bone, infiltration into the tendon of the lateral pterygoid muscle, calcification, periosteal reaction, and osteophyte formation were assessed. A comparison between chondrosarcoma and synovial chondromatosis was performed with a Student *t* test for quantitative variables and the Fisher exact test or linear-by-linear association test for qualitative variables. Receiver operating characteristic analysis was performed to determine the diagnostic performance for differentiation of chondrosarcoma and synovial chondromatosis based on a composite score obtained by assigning 1 point for each of 9 imaging features.

RESULTS: High-risk imaging features for chondrosarcoma were the following: lesion centered on the mandibular condyle, destruction of the mandibular condyle, no destruction/sclerosis of the articular eminence/glenoid fossa, infiltration into the tendon of the lateral pterygoid muscle, absent or stippled calcification, periosteal reaction, internal enhancement, and size of ≥ 30.5 mm. The best cutoff value to discriminate chondrosarcoma from synovial chondromatosis was the presence of any 4 of these high-risk imaging features, with an area under the curve of 0.986 and an accuracy of 95.8%.

CONCLUSIONS: CT and MR imaging features can distinguish chondrosarcoma from synovial chondromatosis of the temporomandibular joint with improved diagnostic performance when a subcombination of 9 imaging features is used.

ABBREVIATIONS: AUC = area under the curve; CBCT = conebeam CT; CS = chondrosarcoma; LPM = lateral pterygoid muscle; MDCT = multidetector row CT; NPV = negative predictive value; PPV = positive predictive value; ROC = receiver operating characteristic; SC = synovial chondromatosis; TMJ = temporomandibular joint

Chondrosarcoma (CS) and synovial chondromatosis (SC) are tumor or tumorlike lesions of the temporomandibular joint (TMJ) that are characterized by cartilaginous neoplasia or metaplasia with variable calcification associated with a mass that causes bony changes of the mandibular condyle and/or the articular eminence/glenoid fossa complex.¹⁻⁷ CS is a malignant neoplasm


generating a variably calcified cartilage matrix, which accounts for ~20%–27% of all primary malignant bone tumors.^{1,8} CS usually arises in the pelvis, humerus, and/or femur, with only 12% originating in the head and neck region.^{2,8} Even in the head and neck region, CS is rare in the TMJ, with only about 49 cases reported through 2020.² SC is a benign tumorlike arthropathy characterized by formation of nodular cartilage in the synovial joints.^{3,5} These cartilaginous nodules can detach from the synovium and produce intra-articular loose bodies⁶ that are nourished by synovial fluid and can be calcified.⁴ SC predominantly manifests in the knee, hip, and elbow and is uncommon in the TMJ.⁵ Approximately 400 cases of TMJ involvement have been reported through 2021.⁷

Differentiating CS from SC is essential because they require different therapeutic approaches. The criterion standard for CS treatment is surgical resection.⁹ Resection for CS must be as wide as possible, and a large healthy tissue margin of ≥ 2 cm seems to

Received February 21, 2023; accepted after revision August 3.

From the Department of Oral and Maxillofacial Radiology and Dental Research Institute (B.G.J., K.H.H., J.E.K., H.J.Y., W.J.Y., M.S.H., S.S.L.), School of Dentistry, Seoul National University, Seoul, Korea; Department of Oral and Maxillofacial Radiology and Wonkwang Dental Research Institute (H.G.Y.), School of Dentistry, Wonkwang University, Iksan, Korea; and Department of Oral and Maxillofacial Radiology (J.H.K.), Seoul National University Dental Hospital, Seoul, Korea.

Please address correspondence to K.H. Huh, DDS, Department of Oral and Maxillofacial Radiology, School of Dentistry, Seoul National University, 101 Daehak-ro, Jongno-gu, Seoul 03080, Korea; e-mail: future3@snu.ac.kr

 Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A7980>

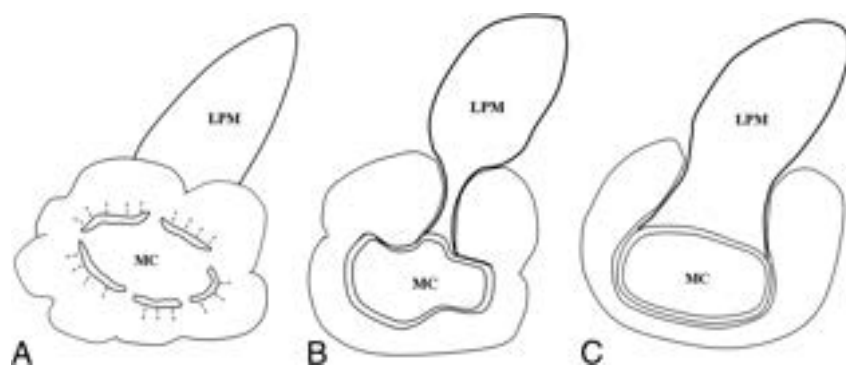


FIG 1. Schematic drawings of the degree of infiltration into the tendon of the LPM by a mass in the pterygoid fovea. Complete and partial infiltration is shown in A and B, respectively. For such cases, a score of 1 is assigned. No infiltration with an intact LPM attachment is shown in C, in which case a score of zero is assigned. MC indicates mandibular condyle.

positively affect prognosis.^{9,10} In contrast, treatment of SC, if indicated, generally includes open joint surgery to remove all affected synovium and loose bodies but usually conserves the mandibular condyle. If the lesion causes damage to the mandibular condyle, then high condylectomy can be considered.¹¹

Clinicians, radiologists, and even pathologists may have difficulty differentiating CS from SC of the TMJ.¹²⁻¹⁵ Presenting symptoms such as pain, trismus, and/or swelling are similar,^{2,3} and their imaging features have not been systematically compared in previous studies. Moreover, because low-grade CS and SC share some histopathologic features such as binucleated chondrocytes and considerable cellularity, problems have been raised with a definitive histopathologic diagnosis.¹⁶⁻¹⁹ Histopathologic discrimination between CS and SC is especially difficult in cases of secondary CS, which develops from a pre-existing cartilaginous tumorlike arthropathy such as SC, compared with primary CS, which arises *de novo*.¹⁴ Thus, differentiation between CS and SC based on imaging is very important. Recently, certain imaging features of CS, such as outward growth from the mandibular condyle and infiltration into the tendon of the lateral pterygoid muscle (LPM) attachment, were reported, and the possibility of differentiation from other lesions of the TMJ was suggested.² Comparisons using consistent imaging parameters for CS and SC of the TMJ have not, however, been reported.

The aim of this study was to assess the diagnostic performance of a series of specific CT and MR imaging parameters for differentiation of CS from SC of the TMJ.

MATERIALS AND METHODS

This retrospective study was approved by the institutional review board of Seoul National University Dental Hospital (ERI19009), and informed consent was waived.

The histopathologic database of our institution between January 2001 and October 2020 was searched for patients who had a proved diagnosis of CS or SC of the TMJ. Patients with secondary SC due to degenerative change were excluded.

Clinical information was obtained from the electronic medical record. Demographic characteristics, affected side, and chief symptom were analyzed.

Image Acquisition

For all patients with CS, both multidetector row CT (MDCT) and MR imaging were performed. For patients with SC, MDCT, conebeam CT (CBCT), and MR imaging were performed on 27, 13, and 24 patients, respectively. All patients with SC underwent either an MDCT or MR imaging examination for soft-tissue evaluation and MDCT or CBCT examination for hard-tissue evaluation. All patients with CS and 12 patients with SC underwent either MDCT and/or MR imaging with contrast media.

Contrast-enhanced CT was performed using an MDCT scanner (Somatom Sensation 10; Siemens)

from the orbit to the bottom of the sternum. The scan parameters used were 150 mAs, 120 kV, and 1- to 2-mm section collimation. After scanning the scout image, contrast media (1.5 mL/kg, iopromide, Ultravist 370; Bayer HealthCare Pharmaceuticals) was injected intravenously at a rate of 2.0–3.0 mL/s. A 3T MR imaging system (Magnetom Skyra; Siemens) and a 1.5T MR imaging system (Signa HDxt; GE Healthcare) were used on 20 and 16 of 47 patients, respectively. Thirty-two- and 16-channel phased array coils were used for the 3T and 1.5T scanners, respectively. Section thickness of 4–6 mm, 320 × 240 or 320 × 192 matrix size, and 19 × 19 or 22 × 22 cm FOV were used. Axial and coronal non-fat-suppressed or fat-suppressed T2-weighted fast spin-echo (TR, 3000–5600 ms; TE, 60–110 ms); axial and coronal non-fat-suppressed T1-weighted spin-echo (TR, 500–600 ms; TE, 10–15 ms); axial, coronal, and sagittal gadolinium-enhanced fat-suppressed T1-weighted spin-echo (TR, 500–700 ms; TE, 9–15 ms) images were acquired. For contrast-enhanced fat-suppressed T1-weighted sequences, a contrast agent (0.1 mmol/kg, gadopentetate dimeglumine, Magnevist; Schering) was injected intravenously via a power injector with a flow rate of 1.5 mL/s. CBCT imaging was performed using a DINNOVA3 scanner (HDXWILL), with an FOV of 20 × 19 cm for men and 20 × 14 cm for women, 9 mAs, 100 kV (peak), and isotropic voxels of 0.3 mm.

Analysis of Imaging Features

Imaging features assessed were lesion center, destruction or sclerosis of the mandibular condyle or articular eminence/glenoid fossa, infiltration into the tendon of the LPM in the pterygoid fovea, calcification, periosteal reaction, presence or absence of osteophytes, enhancement pattern, and lesion size. Each imaging feature was reviewed through a multiplanar assessment of axial, coronal, and sagittal images.

The lesion center was assessed as mandibular condyle or joint space origin of the lesion on MR imaging and/or MDCT images. The presence of bone destruction or sclerosis was assessed on CBCT or MDCT images. Infiltration into the tendon of the LPM was defined as replacement of the LPM attachment by a mass in the pterygoid fovea (Fig 1) and was assessed as present or absent. According to the World Health Organization classification scheme and previous studies,^{7,20,21} the pattern of calcification was

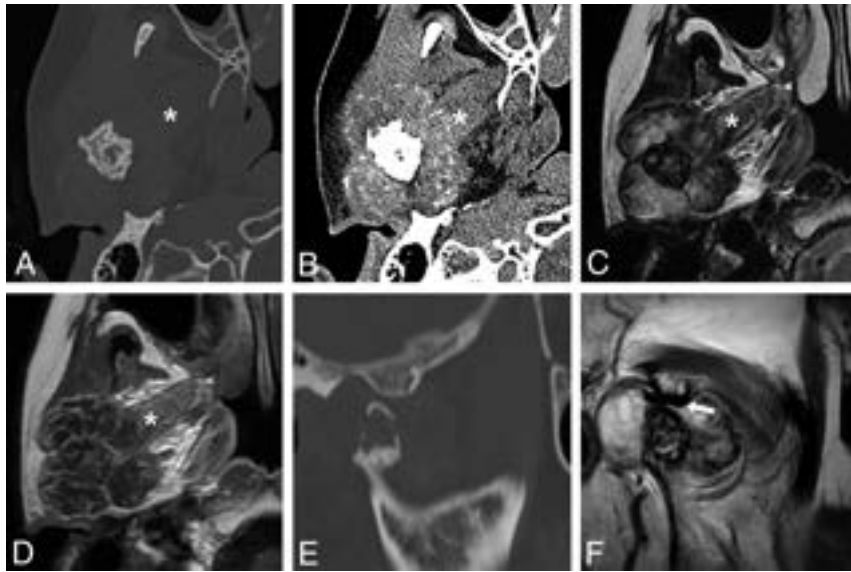


FIG 2. CS of the right TMJ in 54-year-old woman. Axial MDCT with bone and soft-tissue windows (A and B). T2-weighted MR (C) and contrast-enhanced T1-weighted MR (D) images demonstrate a large mass infiltrating the tendon of the lateral pterygoid muscle (*asterisk*) in the pterygoid fovea and show septalike internal enhancement. Stippled calcification and severe destruction of the mandibular condyle with a spiculate periosteal reaction are revealed (A and B). Note the lesion center on the mandibular condyle. Sagittal MDCT (E) and T2-weighted MR images (F) show the eroded mandibular condyle and the intact articular eminence. Note the disc (*arrow*) located between them. This mass showed all 9 high-risk imaging features for CS and received a composite score of 9.

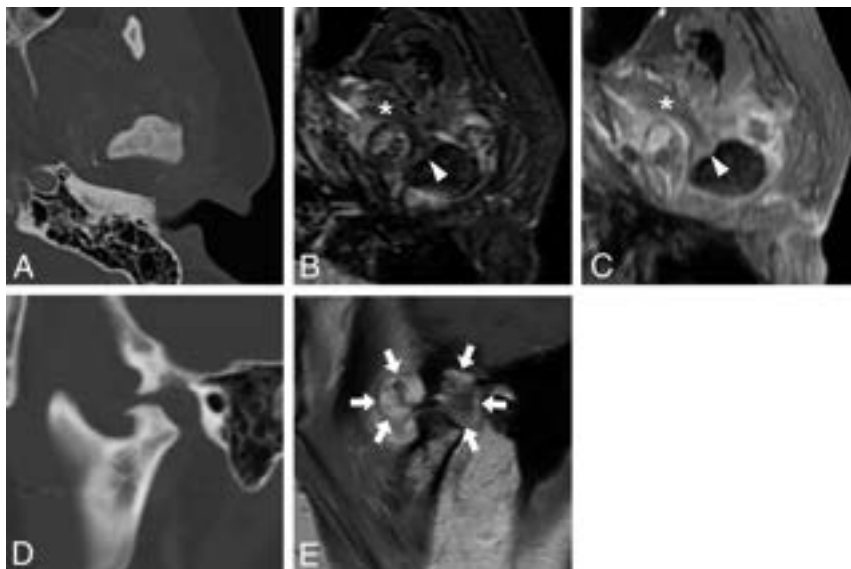


FIG 3. CS of the left TMJ in 44-year-old woman. Axial MDCT (A), fat-suppressed T2-weighted MR (B), and contrast-enhanced T1-weighted MR (C) images reveal a mass surrounding the mandibular condyle. Sagittal MDCT (D) and contrast-enhanced T1-weighted MR images (E) demonstrate the mass center in the joint space. Lesion centered on the superior joint space (*arrows*), destruction and sclerosis of the articular eminence/glenoid fossa, absence of periosteal reaction, and the relatively small lesion size favor SC. On the other hand, destruction of the mandibular condyle, infiltration into the tendon (*arrowhead*) of the lateral pterygoid muscle (*asterisk*), stippled calcification, and internal enhancement favor CS. A histopathologic examination after mass resection resulted in a diagnosis of CS. This TMJ mass revealed 4 high-risk imaging features and a composite score of 4. It was not easy to differentiate CS from SC by imaging features.

classified on MDCT or CBCT images as absence of calcification, stippled type (punctate) (Figs 2 and 3), flocculent type (irregularly shaped), ring-and-arc type (curvilinear, comma-shaped, or annular) with or without a stippled and flocculent type (Fig 4), or popcorn type (amorphous calcifications often with a ring-and-arc type). The presence of a periosteal reaction and/or osteophytes of the mandibular condyle was assessed on MDCT or CBCT images. “Osteophyte” was defined as a beaklike exophytic structure emanating from the cortical surface of the mandibular condyle. The enhancement pattern was assessed as peripheral enhancement or internal enhancement, including septalike, heterogeneous solid, and focal solid. The presence of internal enhancement was evaluated only on contrast-enhanced MDCT or MR images. Lesion size on MR or MDCT images was quantitatively assessed as the longest diameter (X) on an axial plane, the longest diameter (Y) perpendicular to the longest diameter on the axial plane, and the longest diameter (Z) in the cranio-caudal direction on a coronal or sagittal plane. The lesion size was calculated by dividing the sum of X, Y, and Z by 3. To evaluate whether the shape of the lesion was a long saddle, analogous to the shape of the synovium, which is the origin of SC, or a sphere, analogous to the shape of the mandibular condyle, which is the origin of most of CSs, we calculated the ratio of X to Z (X/Z).

Two oral and maxillofacial radiologists with 20 and 15 years of experience, respectively, analyzed all CT and MR imaging scans for imaging analyses while blinded to the histopathologic and clinical information. In addition, they were blinded to the purpose of this study as well. In cases of disagreement between the 2 radiologists, discrepancies were resolved by consensus.

Statistical Analysis

Interobserver agreement was assessed by calculation of the Cohen κ coefficient for qualitative variables and the intraclass correlation coefficient for quantitative variables. The strength of the Cohen κ coefficient was considered as follows: >0.80 , almost perfect

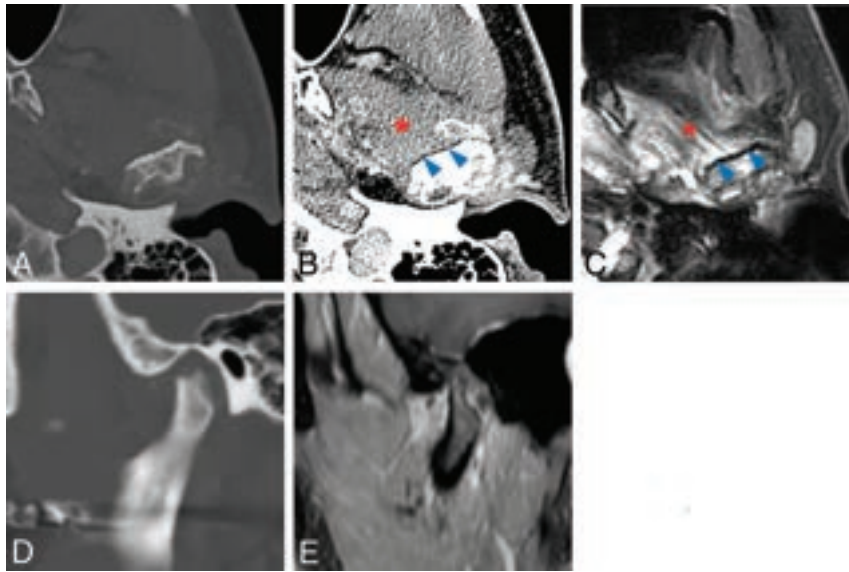


FIG 4. SC of the left TMJ in 52-year-old man. Axial MDCT (A and B) and fat-suppressed T2-weighted MR images (C) show a mass surrounding and eroding the mandibular condyle. Sagittal MDCT (D) and contrast-enhanced T1-weighted MR images (E) reveal the mass to be centered on the joint space. Ring-and-arc calcification, intact attachment of tendon (blue arrowheads) of the lateral pterygoid muscle (red asterisk), lesion center on the joint space, and a relatively small lesion size favor an imaging diagnosis of SC. The result of a preoperative incisional biopsy was CS, but the final histopathologic diagnosis was changed to SC after mass resection. This TMJ mass showed 3 high-risk imaging features and received a composite score of 3.

Table 1: Clinical features of CS and SC^a

Independent Variables	Total (n = 47)	CS (n = 12)	SC (n = 35)	P Value
Age ^b (mean) (yr)	49.3 (SD, 13.5)	50.3 (SD, 15.6)	48.9 (SD, 13.2)	.756
Sex ^c				.659
Male	7 (14.9)	1 (8.3)	6 (17.1)	
Female	40 (85.1)	11 (91.7)	29 (82.9)	
Side ^d				.679
Right	25 (53.2)	7 (58.3)	18 (51.4)	
Left	22 (46.8)	5 (41.7)	17 (48.6)	
Chief complaint				
Swelling ^d	11 (23.4)	5 (41.7)	6 (17.1)	.118
Pain ^c	43 (91.5)	11 (91.7)	32 (91.4)	1.000
Trismus ^c	20 (42.6)	8 (66.7)	12 (34.3)	.089

^a Values are the number of patients with percentage in parentheses or the mean \pm SD.

^b Student *t* test.

^c Fisher exact test.

^d χ^2 test.

reliability; and 0.60–0.80, substantial reliability.²² The strength of the intraclass correlation coefficient was considered as follows: >0.90, excellent reliability; and 0.75–0.90, good reliability.²³ Comparisons of each clinical and imaging feature between the CS and SC groups were performed using the Fisher exact test, χ^2 test, or a linear-by-linear association test for qualitative variables and a Student *t* test for quantitative variables, all at a .05 significance level.

For statistically significant imaging variables, receiver operating characteristic (ROC) analysis with determination of the Youden index was performed to identify the best discriminating value to differentiate between the 2 diseases. Sensitivity, specificity, accuracy, positive predictive value (PPV), negative predictive value (NPV), and area under the curve (AUC) for differentiation were calculated using the best discriminating value. An AUC of

0.5–0.7 was defined as poor discrimination; 0.7–0.8, as acceptable; 0.8–0.9, as excellent; and >0.9 as outstanding.²⁴

Finally, for statistically significant variables, a score of 1 or 0 was assigned to each on the basis of the cutoff value. A point of 1 indicates a high-risk feature for CS, while zero indicates a favorable feature for SC. A composite score for differentiating CS from SC was created by summing the points of all variables. ROC analysis with determination of the Youden index was performed to evaluate the best cutoff value of the composite score to differentiate CS from SC.

SPSS Statistics 23 (IBM) was used for all statistical analyses.

RESULTS

Analysis of Clinical Features

No statistically significant difference was found in age, male-to-female ratio, right-to-left ratio, or chief symptom of swelling, pain, and/or trismus between CS and SC (Table 1).

Analysis of Imaging Features

All imaging features of CS and SC are summarized in Table 2. The interobserver agreement between the 2 readers was almost perfect and excellent for imaging variables, respectively ($P < .001$).

Imaging features that were statistically significant for differentiating CS from SC were as follows: 1) lesion center ($P < .001$), 2) destruction of the mandibular condyle ($P = .037$), 3) destruction of the articular eminence/glenoid fossa ($P = .012$), 4) sclerosis of the articular eminence/glenoid fossa ($P < .001$), 5) infiltration into the tendon of the

LPM ($P < .001$), 6) pattern of calcification ($P = .008$), 7) periosteal reaction ($P = .001$), 8) internal enhancement ($P < .001$), and 9) lesion size ($P < .001$).

The presence of calcification showed no statistical significance. When a calcification pattern was considered as a dimensional continuum from the absence of calcification to popcorn calcification, however, a significant difference in pattern was observed by the linear-by-linear association test ($P = .008$).

Diagnostic Performance of Each Imaging Feature and Composite Scoring Model

The discriminating values for each imaging feature to differentiate CS from SC were as follows: 1) “Mandibular condyle” for lesion center, 2) “Presence” for destruction of the mandibular condyle, 3) “Absence” for destruction of the articular eminence/glenoid

Table 2: Imaging features for CS and SC^a

Independent Variables	Total (n = 47)	CS (n = 12)	SC (n = 35)	P Value	Interobserver Agreement
Lesion center				<.001 ^{b,c}	.832 ^f
Joint space	40 (85.1)	5 (41.7)	35 (100)		
Mandibular condyle	7 (14.9)	7 (58.3)	0 (0)		
Destruction of the mandibular condyle	31 (66.0)	11 (91.7)	20 (57.1)	.037 ^{b,c}	.905 ^f
Sclerosis of the mandibular condyle	23 (48.9)	8 (66.7)	15 (42.9)	.193 ^c	.872 ^f
Destruction of the articular eminence/glenoid fossa	31 (66.0)	4 (33.3)	27 (77.1)	.012 ^{b,c}	.856 ^f
Sclerosis of the articular eminence/glenoid fossa	29 (61.7)	2 (16.7)	27 (77.1)	<.001 ^{b,c}	.861 ^f
Infiltration into the tendon of the LPM	17 (36.2)	12 (100.0)	5 (14.3)	<.001 ^{b,c}	.908 ^f
Calcification	37 (78.7)	8 (66.7)	29 (82.9)	.251 ^c	.828 ^f
Pattern of calcification				.008 ^{b,d}	.835 ^f
Absence	10 (21.3)	4 (33.3)	6 (17.1)		
Stippled	9 (19.1)	6 (50.0)	3 (8.6)		
Flocculent	4 (8.5)	0 (0.0)	4 (11.4)		
Ring-and-arc	9 (19.1)	0 (0.0)	9 (25.7)		
Popcorn	15 (31.9)	2 (16.7)	13 (37.1)		
Periosteal reaction	12 (25.5)	8 (66.7)	4 (11.4)	.001 ^{b,c}	.827 ^f
Osteophyte	10 (21.3)	2 (16.7)	8 (22.9)	1.000 ^c	.873 ^f
Peripheral enhancement	19 (79.2)	11 (91.7)	8 (66.7)	.317 ^c	.864 ^f
Internal enhancement	15 (62.5)	12 (100.0)	3 (25.0)	<.001 ^{b,c}	.830 ^f
Lesion size	28.5 (SD, 7.4)	37.1 (SD, 8.2)	25.6 (SD, 4.2)	<.001 ^{b,e}	.932 ^g
X/Z ratio	1.6 (SD, 0.3)	1.5 (SD, 0.2)	1.7 (SD, 0.3)	.067 ^e	.928 ^g

^a Values are the number of patients with percentages in parentheses or the mean \pm SD.

^b Statistically significant.

^c Fisher exact test.

^d Linear-by-linear association test.

^e Student *t* test.

^f Cohen κ coefficient.

^g Intraclass correlation coefficient.

fossa, 4) “Absence” for sclerosis of the articular eminence/glenoid fossa, 5) “Presence” for infiltration into the tendon of LPM, 6) “Stippled calcification” for pattern of calcification, 7) “Presence” for periosteal reaction, 8) “Presence” for internal enhancement, and 9) “30.5 mm” for lesion size (Fig 5).

Sensitivity, specificity, accuracy, PPV, NPV, and AUC for each imaging feature for differentiating CS from SC are listed in Table 3. The maximum value of sensitivity (100%) was obtained for infiltration into the tendon of the LPM and internal enhancement, while the maximum value of specificity (100%) was revealed for the lesion center. The single imaging feature that showed outstanding performance (AUC = 0.929; 95% CI, 0.855–1.000) was infiltration into the tendon of the LPM.

ROC analysis for the composite score based on the presence of each high-risk feature for CS is presented in Fig 6, and a scoring rubric for each imaging feature is provided in the Online Supplemental Data. The best cutoff value to distinguish CS from SC was observed for the presence of ≥ 4 high-risk features (Youden index = 0.917; AUC = 0.986; 95% CI, 0.950–1.000). This cutoff value showed 100.0%, 91.7%, 95.8%, 92.3%, and 100.0% values for sensitivity, specificity, accuracy, PPV, and NPV, respectively.

DISCUSSION

The present study investigated CT and MR imaging features to differentiate between CS and SC of the TMJ. The imaging distinction between these 2 entities is important because a preoperative histopathologic examination can yield low rates of a correct differential diagnosis between CS and SC.^{2,14} The present study

analyzed 9 imaging features to differentiate between CS and SC. The single imaging feature with the highest performance for the differential diagnosis was infiltration into the tendon of the LPM, followed by lesion size, internal enhancement, and sclerosis of the articular eminence/glenoid fossa. If at least 1 of 4 imaging features showing the 4 highest AUCs is observed, CS should be included in the differential diagnosis. Moreover, with a combination of the 9 imaging features, the diagnostic performance is improved.

Infiltration into the tendon of the LPM, an imaging feature for CS, was first reported in a previous study² and demonstrated the highest AUC and sensitivity (100%) for the prediction of CS. The lesion center on the mandibular condyle and resultant infiltration into the tendon of the LPM in CS could be explained by the conventional intramedullary type of most cases of CS.⁸ All SC in the present study showed a lesion center in the joint space, which can be understood by the joint synovium as the origin of the SC.²⁵ We think that benign tumorlike arthropathies arising in the synovium rarely infiltrate the tendon of the LPM; thus, this imaging feature might be a strong discriminator between other benign and malignant tumors of the TMJ, though further study is required.

The lesion size of CS was significantly larger than that of SC, with a cutoff value of ≥ 30.5 mm. According to the Milgram classification of SC, which divides SC into onset phase I (intra-synovial involvement), transitional phase II (intrasynovial involvement and free bodies), and resolution phase III (multiple free bodies), the onset phase I represents an inflammatory intra-synovial process.²⁵ CS with a malignant nature can show more rapid growth and a resultant larger size compared with SC. In

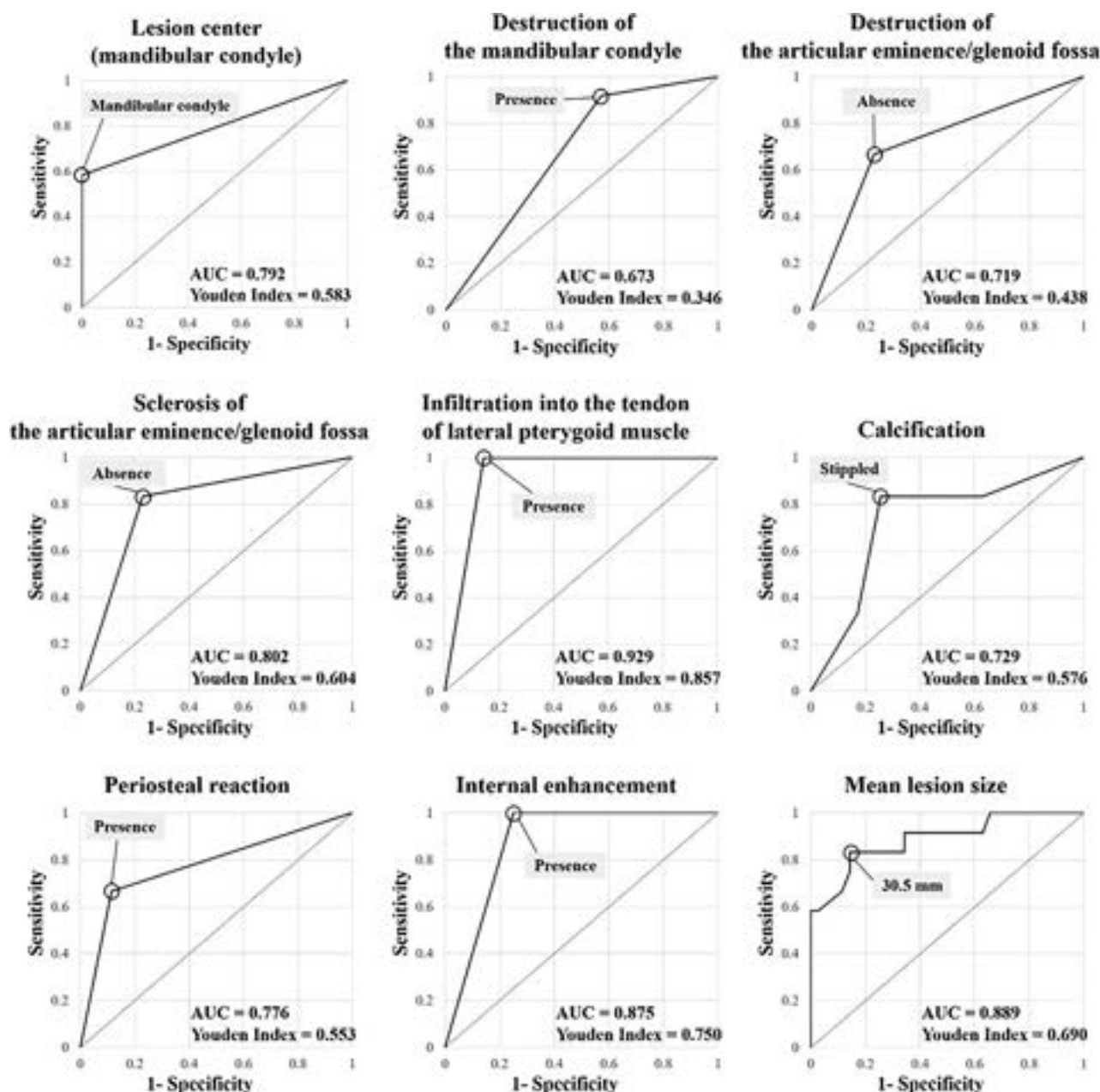


FIG 5. ROC curves for qualitative and quantitative variables show that the best discriminating values for differentiating CS from SC are the following: Mandibular condyle for lesion center, Presence for destruction of the mandibular condyle, Absence for destruction of the articular eminence/glenoid fossa, Absence for sclerosis of the articular eminence/glenoid fossa, Presence for infiltration into the tendon of lateral pterygoid muscle, Stippled pattern of calcification, Presence for periosteal reaction, Presence for internal enhancement, and 30.5 mm for lesion size. Of the 9 variables, except for pattern of calcification and mean lesion size, the remaining 7 variables were dichotomous qualitative variables with discriminating values determined at 1 of the 2 characteristics. On the other hand, mean lesion size was a quantitative variable, and the discriminating value was determined at the point of 30.5 mm, which revealed the maximum Youden index among multiple points. Last, the pattern of calcification, the 5 patterns, namely absence, stippled, flocculent, ring-and-arc, and popcorn, were considered as a dimensional continuum in ascending order and converted into an ordinal quantitative variable. The ROC curve was then constructed, and the discriminating value was derived at the point of stippled calcification, which indicates that CS showed a higher prevalence in absence or stippled calcification, and SC, in the other calcification patterns.

the body, CSs tend to be large tumors, and most exceed 4 cm in maximal lesion size.^{8,26} The average lesion size of 25 cases of SC anywhere in the whole body, including large joints such as shoulder and knee, was 3.9×1.7 cm.²⁷ The larger lesion size in CS than in SC is consistent with findings in previous studies.^{8,26,27}

The presence of internal enhancement was significantly more frequent in CS than in SC and showed high sensitivity (100%) for CS in the present study. Heterogeneous solid and focal solid enhancement in CS might represent malignant cells rather than cartilaginous tissue.² On the other hand, septalike enhancement in cartilaginous tumors or cartilaginous tumorlike conditions

Table 3: Diagnostic performance of each qualitative and quantitative parameter that showed a statistically significant difference for differentiating CS from SC^a

Independent Variables	Prevalence	Sensitivity (%)	Specificity (%)	Accuracy (%)	PPV (%)	NPV (%)	AUC (95% CI)
Lesion center (mandibular condyle)	CS 7/12 SC 0/35	58.3	100.0	89.4	100.0	87.5	0.792 (0.611–0.972)
Destruction of the mandibular condyle (presence)	CS 11/12 SC 20/35	91.7	42.9	55.3	35.5	93.8	0.673 (0.510–0.835)
Destruction of articular eminence/glenoid fossa (absence)	CS 8/12 SC 8/35	66.7	77.1	74.5	50.0	87.1	0.719 (0.542–0.896)
Sclerosis of the articular eminence/glenoid fossa (absence)	CS 10/12 SC 8/35	83.3	77.1	78.7	55.6	93.1	0.802 (0.654–0.950)
Infiltration into the tendon of the LPM (presence)	CS 12/12 SC 5/35	100.0	85.7	89.4	70.6	100.0	0.929 (0.855–1.000)
Calcification (absence or stippled)	CS 10/12 SC 9/35	83.3	74.3	76.6	52.6	92.9	0.729 (0.556–0.901)
Periosteal reaction (presence)	CS 8/12 SC 4/35	66.7	88.6	83.0	66.7	88.6	0.776 (0.604–0.948)
Internal enhancement (presence)	CS 12/12 SC 3/12	100.0	75.0	87.5	80.0	100.0	0.875 (0.719–1.000)
Mean lesion size (≥ 30.5 mm)	CS 10/12 SC 5/35	83.3	85.7	85.1	66.7	93.8	0.889 (0.773–1.000)

^a The items in parentheses correspond to the characteristics of CS, and 1 point was assigned if relevant imaging features in the parentheses were present.

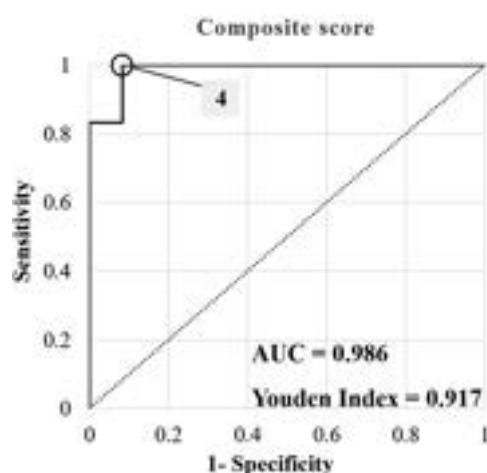


FIG 6. The ROC curve of the composite score for distinguishing CS from SC. The score was assessed by assigning 1 point for each of the following high-risk imaging features for chondrosarcoma: lesion center on the mandibular condyle, destruction of the mandibular condyle, no destruction of the articular eminence/glenoid fossa, no sclerosis of the articular eminence/glenoid fossa, infiltration into the tendon of the lateral pterygoid muscle, absence or stippled calcification, periosteal reaction, internal enhancement, and ≥ 30.5 mm lesion size. The ROC analysis demonstrated that the best cutoff value for differentiating CS from SC was +4 points (Youden index = 0.917, AUC = 0.986; 95% CI, 0.950–1.000).

such as SC are thought to be fibrovascular tissue caused by a lobulated growth pattern.^{27–29} Thus, various patterns of internal enhancement might allow differentiation of CS from other benign tumors.

Destruction of the mandibular condyle and no destruction/sclerosis of the articular eminence/glenoid fossa were significantly more commonly observed in CS than in SC. These significant differences could be explained by the lesion center on the mandibular condyle in most CS, while SC has a lesion center in the joint space, especially the superior joint space.^{7,30,31} The articular disc can protect the articular eminence/glenoid fossa from the mass of CS arising from the mandibular condyle and might protect the mandibular condyle from the mass of SC if it is centered at the superior joint space.

The present study has some limitations. First, this was a retrospective study with a small number of patients, so the possibility of overfitting cannot be avoided. A prospective multicenter study with a larger number of patients is needed to validate the diagnostic performance of the imaging features in the present study. Second, multivariate regression analysis is desired to determine the most impactful independent variables and how those variables interact with each other. In the present study, though the number of imaging parameters was large, and the number of patients was relatively small. This feature prevented regression analysis, and 1 or 0 point was assigned to each imaging variable for calculating the composite score. Third, MR imaging is better than CT for evaluating infiltration into the tendon of the LPM, but not all the patients with SC underwent MR imaging. For 11 of 35 patients, the infiltration into the tendon of the LPM was assessed on MDCT images, and calcifications throughout the SC lesions made it possible to evaluate the extent of the lesion relative to the tendon of the LPM. Fourth, the type of imaging performed was not consistent between patients with CS and those with SC. The fact that MR imaging and contrast-enhanced images were obtained for all patients with CS but not for patients with SC might have influenced the evaluation of the readers. In other words, absence of MR imaging or a contrast image could indicate a low suspicion for CS.

CONCLUSIONS

On the basis of our small case series, we suggest that CT and MR imaging features can differentiate CS and SC. Single imaging features with the highest performance for a differential diagnosis were infiltration into the tendon of the LPM, followed by lesion size and internal enhancement. A combination of imaging features showed very good performance for the differential diagnosis, so that a TMJ mass with ≥ 4 high-risk imaging features can be considered likely CS. Further validation with additional data sets will be informative.

ACKNOWLEDGMENTS

This study was submitted as a doctoral dissertation at Seoul National University and has not been published elsewhere.

REFERENCES

- Thorkildsen J, Taksdal I, Bjerkehagen B, et al. **Chondrosarcoma in Norway 1990–2013; an epidemiological and prognostic observational study of a complete national cohort.** *Acta Oncol* 2019;58:273–82 CrossRef Medline
- Jang BG, Huh KH, Kang JH, et al. **Imaging features of chondrosarcoma of the temporomandibular joint: report of nine cases and literature review.** *Clin Radiol* 2020;75:878.e1–e12 CrossRef Medline
- von Lindern JJ, Theuerkauf I, Niederhagen B, et al. **Synovial chondromatosis of the temporomandibular joint: clinical, diagnostic, and histomorphologic findings.** *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2002;94:31–38 CrossRef Medline
- Murphey MD, Vidal JA, Fanburg-Smith JC, et al. **Imaging of synovial chondromatosis with radiologic-pathologic correlation.** *Radiographics* 2007;27:1465–88 CrossRef Medline
- Meng J, Guo C, Yi B, et al. **Clinical and radiologic findings of synovial chondromatosis affecting the temporomandibular joint.** *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2010;109:441–48 CrossRef Medline
- Testaverde L, Perrone A, Caporali L, et al. **CT and MR findings in synovial chondromatosis of the temporo-mandibular joint: our experience and review of literature.** *Eur J Radiol* 2011;78:414–18 CrossRef Medline
- Jang BG, Huh KH, Kang JH, et al. **Imaging features of synovial chondromatosis of the temporomandibular joint: a report of 34 cases.** *Clin Radiol* 2021;76: 627.e1–e11 CrossRef Medline
- Murphey MD, Walker EA, Wilson AJ, et al. **From the archives of the AFIP: imaging of primary chondrosarcoma—radiologic-pathologic correlation.** *Radiographics* 2003;23:1245–78 CrossRef Medline
- Slimani F, Iro MS. **Chondrosarcoma of the temporomandibular joint: a case report and review of the literature.** *EC Dental Science* 2019;18:1495–501
- Mostafapour SP, Futran ND. **Tumors and tumorous masses presenting as temporomandibular joint syndrome.** *Otolaryngol Head Neck Surg* 2000;123:459–64 CrossRef Medline
- Huh JK, Park JY, Lee S, et al. **Synovial chondromatosis of the temporomandibular joint with condylar extension.** *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2006;101:e83–88 CrossRef Medline
- Merrill RG, Yih WY, Shamloo J. **Synovial chondrosarcoma of the temporomandibular joint: a case report.** *J Oral Maxillofac Surg* 1997;55:1312–16 CrossRef Medline
- Ichikawa T, Miyauchi M, Nikai H, et al. **Synovial chondrosarcoma arising in the temporomandibular joint.** *J Oral Maxillofac Surg* 1998;56:890–94 CrossRef Medline
- Coleman H, Chandraratnam E, Morgan G, et al. **Synovial chondrosarcoma arising in synovial chondromatosis of the temporomandibular joint.** *Head Neck Pathol* 2013;7:304–09 CrossRef Medline
- Ye ZX, Yang C, Chen MJ, et al. **Digital resection and reconstruction of TMJ synovial chondrosarcoma involving the skull base: report of a case.** *Int J Clin Exp Med* 2015;8:11589–93 Medline
- Morris MR, Clark SK, Porter BA, et al. **Chondrosarcoma of the temporomandibular joint: case report.** *Head Neck Surg* 1987;10:113–17 CrossRef Medline
- Warner BF, Luna MA, Robert Newland T. **Temporomandibular joint neoplasms and pseudotumors.** *Adv Anat Pathol* 2000;7:365–81 CrossRef Medline
- Angiero F, Vinci R, Sidoni A, et al. **Mesenchymal chondrosarcoma of the left coronoid process: report of a unique case with clinical, histopathologic, and immunohistochemical findings, and a review of the literature.** *Quintessence Int* 2007;38:349–55 Medline
- Oliveira RC, Marques KD, Mendonça AR, et al. **Chondrosarcoma of the temporomandibular joint: a case report in a child.** *J Orofac Pain* 2009;23:275–81 Medline
- Davies AM, Pettersson H. **The WHO manual of diagnostic imaging: radiographic anatomy and interpretation of the musculoskeletal system.** *World Health Organization* 2002;174
- Subhawong TK, Fishman EK, Swart JE, et al. **Soft-tissue masses and masslike conditions: what does CT add to diagnosis and management?** *AJR Am J Roentgenol* 2010;194:1559–67 CrossRef Medline
- Landis JR, Koch GG. **The measurement of observer agreement for categorical data.** *Biometrics* 1977;33:159–74
- Koo TK, Li MY. **A guideline of selecting and reporting intraclass correlation coefficients for reliability research.** *J Chiropr Med* 2016;15:155–63 CrossRef Medline
- Hosmer DW, Lemeshow S. *Applied Logistic Regression.* John Wiley and Sons 2000;160–64
- Milgram JW. **Synovial osteochondromatosis: a histopathological study of thirty cases.** *J Bone Joint Surg Am* 1977;59:792–801
- Kendell SD, Collins MS, Adkins MC, et al. **Radiographic differentiation of enchondroma from low-grade chondrosarcoma in the fibula.** *Skeletal Radiol* 2004;33:458–66 CrossRef Medline
- Walker EA, Murphey MD, Fetsch JF. **Imaging characteristics of tenosynovial and bursal chondromatosis.** *Skeletal Radiol* 2011;40:317–25 CrossRef Medline
- Aoki J, Sone S, Fujioka F, et al. **MR of enchondroma and chondrosarcoma: rings and arcs of Gd-DTPA enhancement.** *J Comput Assist Tomogr* 1991;15:1011–16 CrossRef Medline
- Geirnaerdt MJ, Bloem JL, Eulerink F, et al. **Cartilaginous tumors: correlation of gadolinium-enhanced MR imaging and histopathologic findings.** *Radiology* 1993;186:813–17 CrossRef Medline
- Sato J, Notani KI, Goto J, et al. **Synovial chondromatosis of the temporomandibular joint accompanied by loose bodies in both the superior and inferior joint compartments: case report.** *Int J Oral Maxillofac Surg* 2010;39:86–88 CrossRef Medline
- Chen MJ, Yang C, Cai XY, et al. **Synovial chondromatosis in the inferior compartment of the temporomandibular joint: different stages with different treatments.** *J Oral Maxillofac Surg* 2012;70:e32–38 CrossRef Medline

Performance of Neck Imaging Reporting and Data System (NI-RADS) for Diagnosis of Recurrence of Head and Neck Squamous Cell Carcinoma: A Systematic Review and Meta-analysis

Akira Baba, Ryo Kurokawa, Mariko Kurokawa, Takafumi Yanagisawa, and Ashok Srinivasan



ABSTRACT

BACKGROUND: The Neck Imaging Reporting and Data System (NI-RADS) is a reporting template used in head and neck cancer posttreatment follow-up imaging.

PURPOSE: Our aim was to evaluate the pooled detection rates of the recurrence of head and neck squamous cell carcinoma based on each NI-RADS category and to compare the diagnostic accuracy between NI-RADS 2 and 3 cutoffs.

DATA SOURCES: The MEDLINE, Scopus, and EMBASE databases were searched.

STUDY SELECTION: This systematic review identified 7 studies with a total of 694 patients (1233 lesions) that were eligible for the meta-analysis.

DATA ANALYSIS: The meta-analysis of pooled recurrence detection rate estimates for each NI-RADS category and the diagnostic accuracy of recurrence with NI-RADS 3 or 2 as the cutoff was performed.

DATA SYNTHESIS: The estimated recurrence rates in each category for primary lesions were 74.4% for NI-RADS 3, 29.0% for NI-RADS 2, and 4.2% for NI-RADS 1. The estimated recurrence rates in each category for cervical lymph nodes were 73.3% for NI-RADS 3, 14.3% for NI-RADS 2, and 3.5% for NI-RADS 1. The area under the curve of the summary receiver operating characteristic for recurrence detection with NI-RADS 3 as the cutoff was 0.887 and 0.983, respectively, higher than 0.869 and 0.919 for the primary sites and cervical lymph nodes, respectively, with NI-RADS 2 as the cutoff.

LIMITATIONS: Given the heterogeneity of the data of the studies, the conclusions should be interpreted with caution.

CONCLUSIONS: This meta-analysis revealed estimated recurrence rates for each NI-RADS category for primary lesions and cervical lymph nodes and showed that NI-RADS 3 has a high diagnostic performance for detecting recurrence.

ABBREVIATIONS: AUC = area under the curve; CE-CT = contrast-enhanced CT; CE-MRI = contrast-enhanced MRI; DOR = diagnostic odds ratio; NI-RADS = Neck Imaging Reporting and Data System; sROC = summary receiver operating characteristic

Follow-up imaging after head and neck cancer treatment is used for the assessment of the treatment response and the detection of recurrence. Recurrences may involve the primary site and/or cervical lymph nodes, and early detection of such recurrences may facilitate subsequent salvage therapy.^{1,2} Posttreatment follow-up imaging of head and neck cancer is often challenging, however,


because of the anatomic complexity of the head and neck region, complex resection and reconstruction operations, and the post-treatment effects of radiation and chemotherapy that mimic recurrent disease. These factors affect radiologists' interpretations, rendering them nonuniform and potentially suboptimal.^{3,4}

The Neck Imaging Reporting and Data System (NI-RADS) is a head and neck cancer posttreatment follow-up imaging reporting template that was proposed by the American College of Radiology in 2016 to standardize imaging interpretation and communication between clinicians and radiologists.⁵ NI-RADS provides standardized terminology, report structure, and evaluation categories to convey the degree of suspicion of recurrence in the interpretation of imaging studies. The NI-RADS lexicon established for the evaluation of both posttreatment primary sites and cervical lymph nodes has 4 categories (category 1 [no evidence of recurrence], category 2 [low suspicion], category 3 [high

Received March 31, 2023; accepted after revision August 12.

From the Division of Neuroradiology (A.B., R.K., M.K., A.S.), Department of Radiology, University of Michigan, Ann Arbor, Michigan; Departments of Radiology (A.B.) and Urology (T.Y.), The Jikei University School of Medicine, Tokyo, Japan; and Department of Radiology (R.K., M.K.), The University of Tokyo, Tokyo, Japan.

Please address correspondence to Akira Baba, MD, PhD, Department of Radiology, The Jikei University School of Medicine, 3-25-8 Nishi-Shimbashi, Minato-ku, Tokyo 105-8461, Japan; e-mail: akirababa0120@gmail.com

 Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A7922>

suspicion], and category 4 [definitive recurrence]). NI-RADS 1–3 provide linked recommendations for clinical management along with an estimate of the degree of suspicion for recurrent head and neck cancer.

Most previously published studies of the diagnostic performance of NI-RADS have been limited by small sample sizes. Therefore, the purpose of this systematic article was to summarize the existing data and estimate the detection rate for recurrent head and neck squamous cell carcinoma for each NI-RADS category and to compare diagnostic test accuracy estimates using NI-RADS 3 versus 2 cutoffs for detecting recurrent lesions.

MATERIALS AND METHODS

Study Selection

This study was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement.⁶ We searched the Cochrane database and confirmed that there were no reviews/meta-analyses similar to the present research design. On November 10, 2022, MEDLINE via PubMed, Scopus, and EMBASE databases were screened using the following search terms, without any language or date limits: “Neck Imaging Reporting and Data System” or “NI-RADS.”

Inclusion criteria for this evaluation were as follows:

- Data on the number of lesions in each NI-RADS category and the number of proven primary site or cervical lymph node recurrences
- Data including either primary sites or lymph nodes in NI-RADS 1, 2, or 3
- Data with contrast-enhanced CT (CE-CT), contrast-enhanced MRI (CE-MRI), PET/CT, or PET/MRI
- Data with the pathology of squamous cell carcinoma only
- Original studies that investigated human findings
- In cases of duplicate publications, the highest quality or most recent publication was selected
- Written in English

The exclusion criteria were as follows:

- Studies published before 2016
- Studies without an identified imaging period from treatment
- The full text was unavailable
- Studies with incomplete data
- Review, case reports, and systematic review/meta-analysis articles
- Books and conference proceedings only that lacked an associated peer-reviewed full-fledged publication

Data Extraction

Two board-certified radiologists with 13 and 9 years of experience, respectively, in head and neck radiology reviewed the full text of the eligible studies and extracted the following information from the included studies by consensus: first author's name, study region, publication year, study period, study design, number of patients, age, sex, tumor subsite, pathology, treatment method, type of imaging technique, vendor and model of equipment used, imaging period from therapy, reference standard, and recurrent and nonrecurrent lesions for each NI-RADS

category. Any disagreements were resolved by consensus agreement of the evaluators.

Quality and Risk Assessment

The Newcastle-Ottawa Scale was used to assess the quality of the included studies in accordance with the Cochrane Handbook for Systematic Reviews of Interventions for included nonrandomized studies.^{7,8} The scale rates the following 3 factors: selection (0–4 points), comparability (0–2 points), and exposure (0–3 points), with total scores ranging from 0 (lowest) to 9 (highest). Studies with scores of >6 were identified as “high-quality” choices.

Data Analyses

Proportional meta-analyses were performed using a random effects model or a common effects model to determine the estimated prevalence of recurrent disease for each NI-RADS category. The following data were available for this analysis: 2 articles for which all data from NI-RADS 1–3 of the primary lesions and cervical lymph nodes were available,^{9,10} 1 article with only NI-RADS 3 data available for the primary lesions and cervical lymph nodes,¹¹ 1 article with only NI-RADS 2 and NI-RADS 3 data points available for the primary lesion,¹² 1 article with only primary lesion data from NI-RADS 1–3 due to an additional modification of the NI-RADS assessment for the cervical lymph node,¹³ 1 article with only primary lesion data from NI-RADS 1–3 available,¹⁴ and 1 article with only cervical lymph node data from NI-RADS 1–3 available due to case overlap of primary lesion cases with another article.¹⁵ Forest plots were used to assess and summarize the data. Heterogeneity among the outcomes of the studies included in this article was evaluated using the I^2 statistic.

Significant heterogeneity was indicated by a ratio $>50\%$ in I^2 statistics. Publication bias was assessed using funnel plots. In 4 articles for which all case number data for recurrence and nonrecurrence from each of NI-RADS 1–3 at the primary lesions were available^{9,10,13,14} or in 3 articles for which all case number data for recurrence and nonrecurrence from each of NI-RADS 1–3 at the cervical lymph nodes were available,^{9,10,15} we divided the data for the diagnostic accuracy analysis into groups of NI-RADS 3 and NI-RADS 1/2 when NI-RADS 3 was used as the cutoff, and into groups of NI-RADS 2/3 and NI-RADS 1 when NI-RADS 2 was used as the cutoff. Data were pooled using random or fixed effects models to summarize the estimates of sensitivity, specificity, and diagnostic odds ratio (DOR). Bivariate models were used to construct summary receiver operator characteristic (sROC) curves and calculate the area under the curve (AUC). All statistical analyses were performed using R, Version 4.2.2 (<http://www.r-project.org/>).

RESULTS

Study Selection and Characteristics

Our initial search identified 257 records, and 185 remained after removing duplicates and/or conference proceedings and book chapters. In the next screening, 85 articles published before 2016, non-English language reports, review articles, case reports, and systematic reviews/meta-analyses were excluded. After applying the inclusion/exclusion criteria, we identified 7 articles with 694 patients (1233 lesions) for this review (Fig 1).^{9–15}

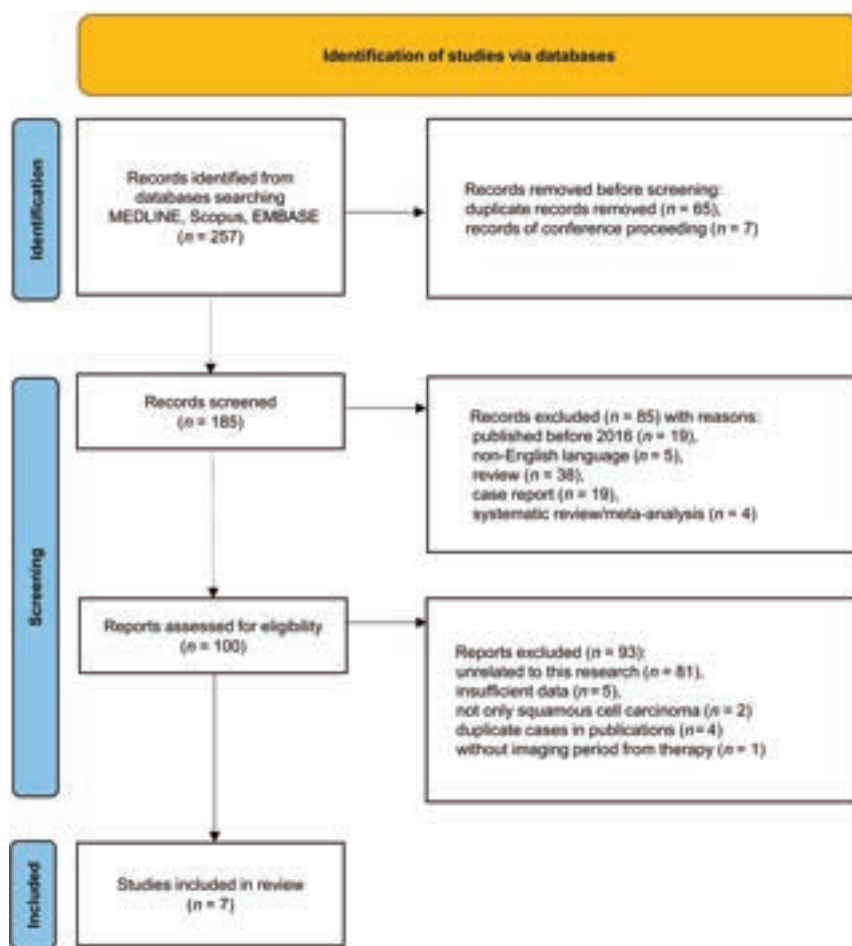


FIG 1. The PRISMA 2020 flow chart for the article-selection process. After applying the selection criteria, we identified 7 articles for the systematic review and the meta-analysis.

The data extracted from the 7 studies are outlined in the Online Supplemental Data. All were published between 2019 and 2022, with 2 studies from North America, 2 from South Asia, 2 from Africa, and 1 from Europe. The studies had a median Newcastle-Ottawa Scale score of 4 (range, 4–5). The 6 studies for which information regarding the sex of individual participants was available included 419 men and 191 women (male/female ratio = 2.2:1). The studies had a mean age range of 49–63.4 years and a median age range of 59–62 years. The primary tumor subsites in the studies included the nasopharynx, oropharynx, hypopharynx, larynx, oral cavity, sinonasal cavity, and salivary gland. The 6 studies for which treatment methods were available included radiation therapy, chemoradiotherapy, surgery, surgery plus radiation therapy, and surgery plus chemoradiotherapy. The imaging modalities in the studies were CE-CT, CE-MRI, PET/CT, or PET/CT with CE-CT. Imaging was performed >1.5–3 months after completion of therapy.

In the reference standard, histology and follow-up were used in all articles to determine recurrent lesions, whereas only follow-up was used in 3 articles to determine nonrecurrent lesions. In NI-RADS at the primary site, 6 articles were available for NI-RADS 3; five, for NI-RADS 2; and 4, for NI-RADS 1, for a total of 710 lesions evaluated. In NI-RADS for cervical lymph nodes, data

for 4 articles were available for NI-RADS 3, three for NI-RADS 2, and 3 for NI-RADS 1, for a total of 523 lesions evaluated. For diagnostic performance evaluation using NI-RADS 3 or NI-RADS 2 as a cutoff, 4 articles were available for NI-RADS at the primary sites and 3 articles were available for NI-RADS at the lymph nodes. Seven studies reported 270 recurrent lesions and 963 nonrecurrent lesions at the primary sites and cervical lymph nodes.

Meta-analysis of Summarized Recurrence Rate Estimates for Each NI-RADS Category

A forest plot of the summarized estimates of the prevalence for each NI-RADS category of recurrent lesions at the primary sites or cervical lymph nodes is shown in Figs 2 and 3. At the primary site, the estimated recurrence rate for NI-RADS 3 was 74.4% (95% CI, 59.2%–85.4%; $I^2 = 74\%$); for NI-RADS 2, it was 29.0% (95% CI, 21.2%–38.2%; $I^2 = 47\%$); and for NI-RADS 1, it was 4.2% (95% CI, 2.6%–6.7%; $I^2 = 0\%$). For cervical lymph nodes, the estimated recurrence rates for NI-RADS 3, NI-RADS 2, and NI-RADS 1 were 73.3% (95% CI, 63.0%–81.5%; $I^2 = 38\%$), 14.3% (95% CI, 6.1%–30.0%; $I^2 = 0\%$), and 3.5% (95% CI, 2.1%–5.8%; $I^2 =$

0%), respectively. Funnel plots of these results are shown in the Online Supplemental Data.

Meta-analysis of Diagnostic Test Accuracy of Recurrence with NI-RADS 3 or 2 as the Cutoff

Primary Lesion. A forest plot of the summarized estimates of sensitivity, specificity, and DOR of recurrence detection with NI-RADS 3 or NI-RADS 2 as the cutoff in the primary lesion is shown in the Online Supplemental Data, and the sROC for diagnostic performance is shown in Fig 4. With NI-RADS 3 as the cutoff, the estimated sensitivity for recurrence was 60.6% (95% CI, 39.5%–78.4%), the estimated specificity was 94.6% (95% CI, 84.6%–98.3%), the estimated DOR was 26.6 (95% CI, 13.5–52.4), and the AUC in sROC was 0.887. With NI-RADS 2 as the cutoff, the estimated sensitivity for recurrence was 81.8% (95% CI, 54.5%–94.4%), the estimated specificity was 76.6% (95% CI, 29.5%–96.2%), the estimated DOR was 18.9 (95% CI, 9.4–37.9), and the AUC in sROC was 0.869.

Lymph Nodes. A forest plot of the summarized estimates of sensitivity, specificity, and the DOR of recurrence detection with NI-RADS 3 or NI-RADS 2 as the cutoff for cervical lymph nodes is shown in the Online Supplemental Data, and the sROC for diagnostic performance is shown in Fig 5. With NI-RADS 3 as the

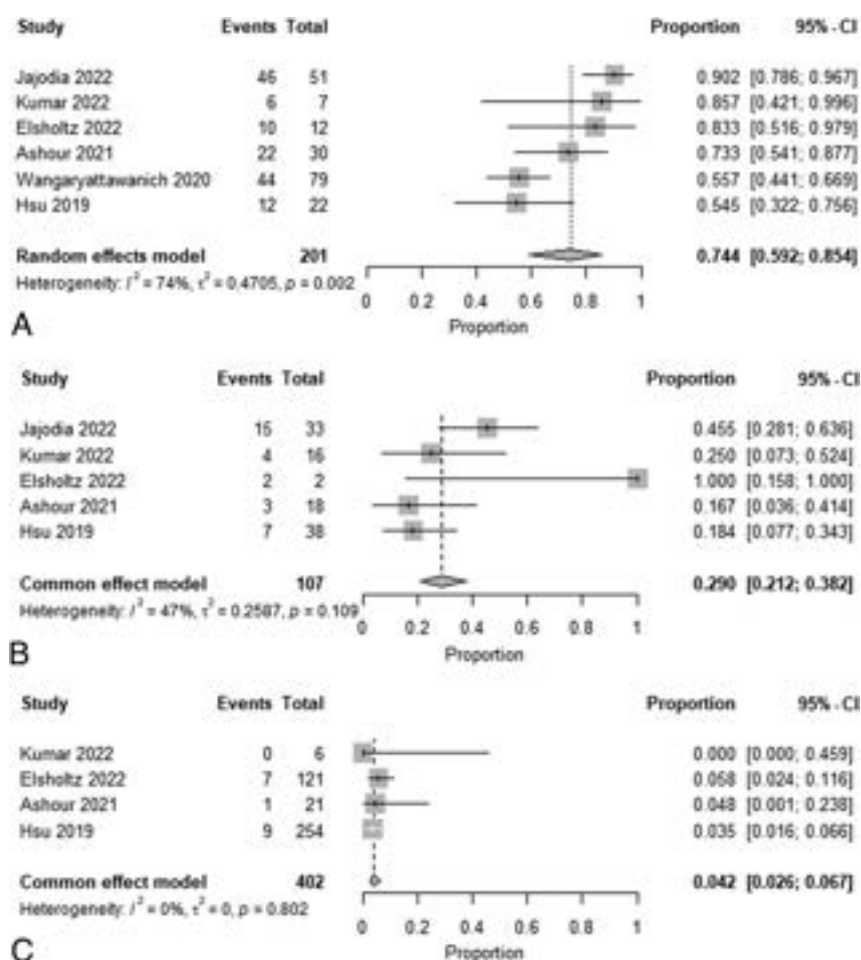


FIG 2. Forest plot of summary estimates of the prevalence of recurrence for each NI-RADS category at the primary site (A, NI-RADS 3. B, NI-RADS 2. C, NI-RADS 1).

cutoff, the estimated sensitivity for recurrence was 66.9% (95% CI, 28.9%–90.9%), the estimated specificity was 98.6% (95% CI, 96.7%–99.4%), the estimated DOR was 96.0 (95% CI, 31.1–296.2), and the AUC in sROC was 0.983. With NI-RADS 2 as the cutoff, the estimated sensitivity for recurrence was 75.7% (95% CI, 38.5%–94.0%), the estimated specificity was 89.3% (95% CI, 78.8%–95.0%), the estimated DOR was 21.3 (95% CI, 9.9–45.9), and the AUC in the sROC was 0.919.

DISCUSSION

This systematic review and meta-analysis used 7 studies reporting 694 patients and 1233 lesions to calculate the summarized estimated detection rates of head and neck cancer recurrence for each NI-RADS category and to compare the diagnostic accuracy of NI-RADS 2 and 3 cutoffs to define the optimal cutoff value. The estimated recurrence rates in the categories of primary lesions and cervical lymph nodes were 73.3%–74.4%, 14.3%–29.0%, and 3.5%–4.2% for NI-RADS 3, NI-RADS 2, and NI-RADS 1, respectively. Furthermore, the summarized estimates of specificity, the DOR, and the AUC of the sROC for recurrence detection were higher when NI-RADS 3 was used as a cutoff in primary lesions or cervical lymph nodes than when NI-RADS 2 was used.

The current lexicon in NI-RADS is based on standardized report templates specific to CE-CT and PET/CE-CT, though NI-RADS can also be used for the interpretation of CE-MRI or PET/MR imaging.^{5,16} Therefore, in addition to CE-CT and PET/CT, there have been a number of articles regarding NI-RADS using CE-MRI and PET/MR imaging.^{10,12,14,15,17–21} The inclusion of T2-weighted images, DWI, and ADC findings improves the diagnostic performance of NI-RADS.¹² Other important MR imaging findings include quantitative values such as ADC values and dynamic contrast-enhanced MR imaging parameters, which are reported to be useful in differentiating recurrent head and neck cancer and posttreatment effects.^{22,23} Further reporting on the utility of these qualitative and quantitative MR imaging findings for NI-RADS incorporation or the establishment of a revised NI-RADS lexicon that includes these findings and assessment parameters may be warranted in the future but is beyond the scope of this work.

Each NI-RADS category has a different set of clinical recommendations that have been proposed as follows: NI-RADS 2: clinical evaluation of the concerning mucosal region, relatively close follow-up (~ 3 months) or FDG-PET; and NI-RADS 3 recommends tissue correlation.⁵ The results of the current study showed that the estimated specificity, estimated DOR, and AUC for head and neck cancer recurrence for NI-RADS 3 as a cutoff were higher than those for NI-RADS 2 for both primary lesions and cervical lymph nodes, with an estimated recurrence rate in NI-RADS 3 as high as 74.5%–74.6%. These results support using NI-RADS 3 as a cutoff for recurrent lesion detection and proceeding to the linked clinical recommendation for NI-RADS 3 of biopsy, an invasive procedure. However, the estimated recurrence rates of 12.2%–29.8% for NI-RADS 2 in the results of this study are a relatively cautionary prevalence that cannot be considered safe. Therefore, close follow-up and direct testing as recommended by the current lexicon for NI-RADS 2 is validated by this analysis.

The utility of liquid biopsy for monitoring recurrence in head and neck cancer is promising^{24,25} and may have the potential to be incorporated into future NIRADS clinical recommendations.

Our study does have certain limitations, most notably the limited number of available studies. The Newcastle-Ottawa Scale score for all studies was low (4–5), suggesting a high risk of bias. Variations in the type of imaging technique, the location of head and neck cancers, time from treatment to imaging studies, and the reference method among the included studies could have

influenced the heterogeneous outcomes. One study that included only NI-RADS 3 cases¹¹ and another study that included only primary NI-RADS 2 and NI-RADS 3 cases¹² may have disproportionately affected the results of the meta-analysis of proportions. Some funnel plots exhibited asymmetry indicating potential publication bias. Due to the limited number of included studies, however, a precise evaluation of publication bias was not feasible. Such

heterogeneity and bias among studies must be considered a potential limitation when assessing the significance of this analysis.

The NI-RADS assessments and outcome differences may also have been influenced by the treatment method. In the NI-RADS, only local lesions could be classified into NI-RADS 2a and 2b.⁵ Of the articles included in this study, the number of recurrences and nonrecurrences in NI-RADS 2a and 2b could be extracted for only 2 articles.^{10,13} Due to the limited number of available articles, this study could not perform a meta-analysis of the NI-RADS 2a/2b. Further studies are warranted to comprehensively evaluate the diagnostic performance of the NI-RADS 2a/2b cutoff. Last, although random effects models were used in some of the tests to treat heterogeneity across studies, our conclusions should still be interpreted with caution because the underlying studies on this topic were not strongly based.²⁶ Appropriately designed prospective large-scale trials are needed to validate the results of this study.

CONCLUSIONS

This meta-analysis demonstrated that NI-RADS 3 has a high diagnostic performance for detecting clinically significant recurrence and confirmed that a NI-RADS category 3 is the optimal cut-off value as a clinical recommendation linked to tissue sampling. Given that NI-RADS 2 lesions also have relatively high estimated recurrence rates, careful follow-up is mandatory in this group of patients. Considering the limitations,

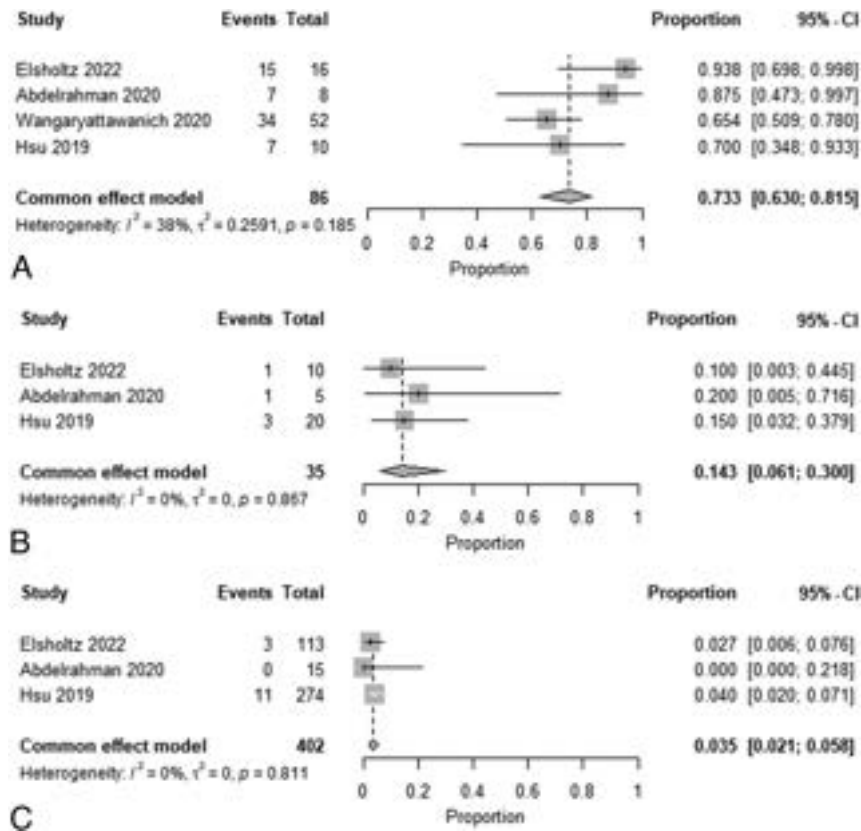


FIG 3. Forest plot of summary estimates of the prevalence of recurrence for each NI-RADS category at the neck node (A, NI-RADS 3. B, NI-RADS 2. C, NI-RADS 1).

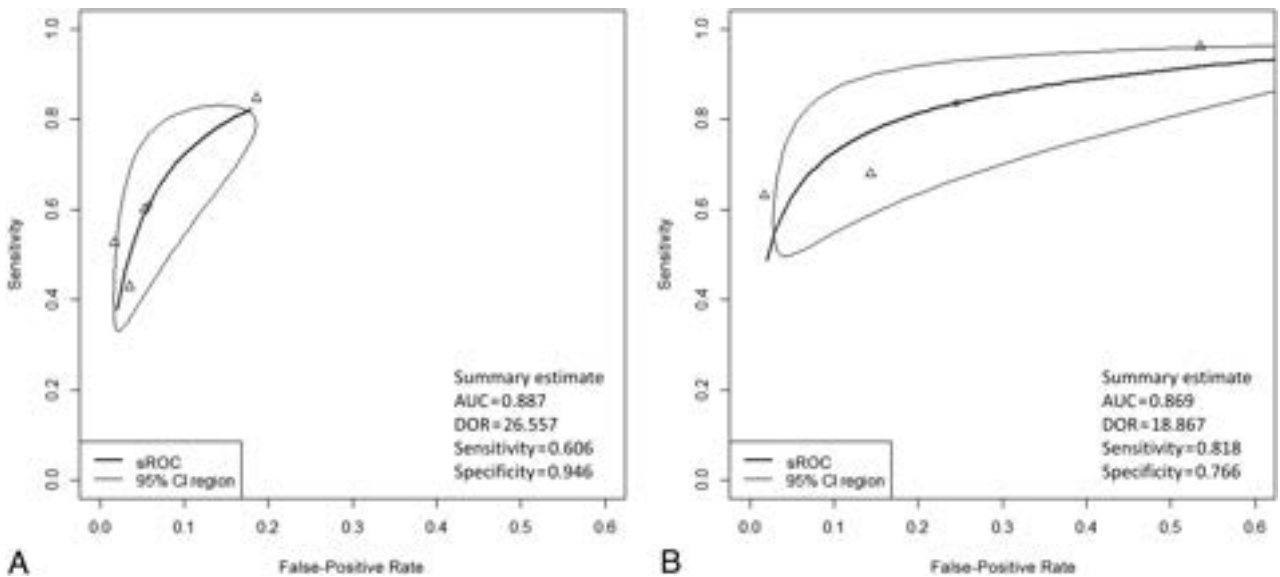


FIG 4. sROC for diagnostic performance in the primary site. A, sROC with NI-RADS 3 as the cutoff. B, sROC with NI-RADS 2 as the cutoff.

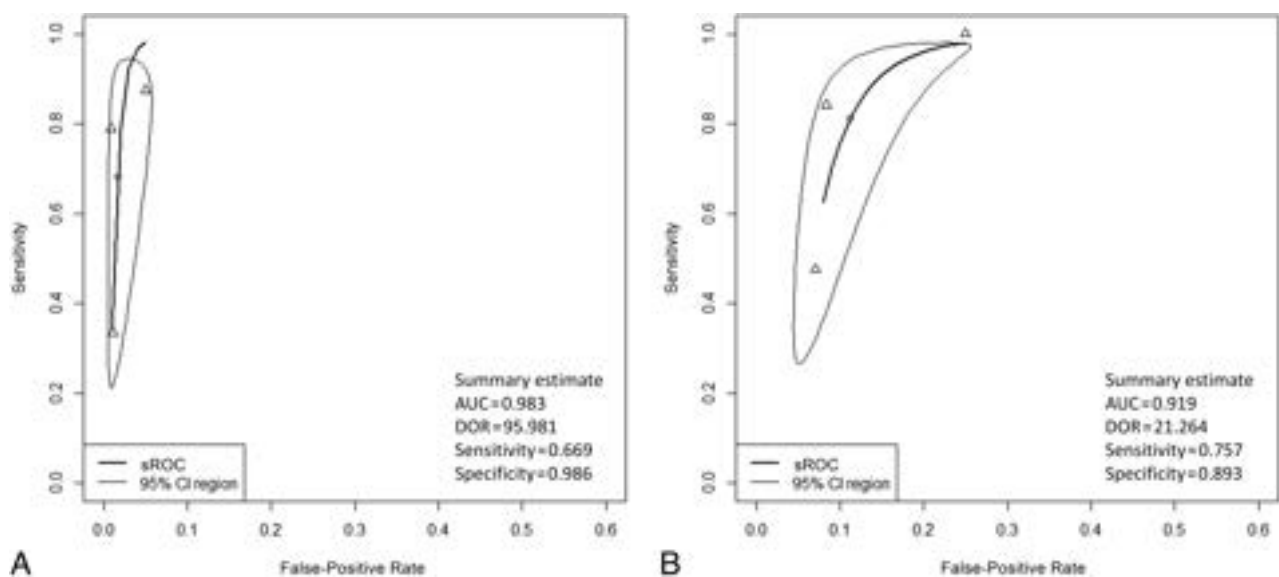


FIG 5. sROC for diagnostic performance in the lymph node. *A*, sROC with NI-RADS 3 as the cutoff. *B*, sROC with NI-RADS 2 as the cutoff.

including procedural and methodologic heterogeneity among the eligible studies, further investigation and validation of this study are needed.

ACKNOWLEDGMENTS

We gratefully acknowledge Mr Mark MacEachern (Taubman Health Sciences Library, University of Michigan) for his advice and Dr Roberto Rivera-de Choudens (Division of Neuroradiology, Department of Radiology, University of Michigan) for his English proofreading.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

- Wong LY, Wei WI, Lam LK, et al. **Salvage of recurrent head and neck squamous cell carcinoma after primary curative surgery.** *Head Neck* 2003;25:953–59 CrossRef Medline
- van Weert S, Leemans CR. **Salvage surgery in head and neck cancer.** *Oral Dis* 2021;27:117–24 CrossRef Medline
- Baba A, Ojiri H, Ikeda K, et al. **Essentials on oncological imaging: postoperative computed tomography and magnetic resonance imaging of oral tongue cancer.** *Can Assoc Radiol J* 2018;69:458–67 CrossRef Medline
- Strauss SB, Aiken AH, Lantos JE, et al. **Best practices: application of NI-RADS for posttreatment surveillance imaging of head and neck cancer.** *AJR Am J Roentgenol* 2021;216:1438–51 CrossRef Medline
- Aiken AH, Rath TJ, Anzai Y, et al. **ACR Neck Imaging Reporting and Data Systems (NI-RADS): a white paper of the ACR NI-RADS Committee.** *J Am Coll Radiology* 2018;15:1097–108 CrossRef Medline
- Page MJ, McKenzie JE, Bossuyt PM, et al. **The PRISMA 2020 statement: an updated guideline for reporting systematic reviews.** *BMJ* 2021;372:n71 CrossRef Medline
- Stang A. **Critical evaluation of the Newcastle-Ottawa Scale for the assessment of the quality of nonrandomized studies in meta-analyses.** *Eur J Epidemiol* 2010;25:603–05 CrossRef Medline
- Deeks JJ, Dinnes J, D'Amico R, et al; European Carotid Surgery Trial Collaborative Group. **Evaluating non-randomised intervention studies.** *Health Technol Assess* 2003;7:iii–x. 1–173 CrossRef Medline
- Hsu D, Chokshi FH, Hudgins PA, et al. **Predictive value of first post-treatment imaging using standardized reporting in head and neck cancer.** *Otolaryngol Head Neck Surg* 2019;161:978–85 CrossRef Medline
- Elsholtz FH, Ro SR, Shnayien S, et al. **Impact of double reading on NI-RADS diagnostic accuracy in reporting oral squamous cell carcinoma surveillance imaging: a single-center study.** *Dentomaxillofac Radiol* 2022;51:1–8 CrossRef Medline
- Wangyattawanich P, Branstetter BF, Ly JD, et al. **Positive predictive value of neck imaging reporting and data system categories 3 and 4 posttreatment FDG-PET/CT in head and neck squamous cell carcinoma.** *AJNR Am J Neuroradiol* 2020;41:1070–75 CrossRef Medline
- Jajodia A, Mandal G, Yadav V, et al. **Adding MR diffusion imaging and T2 signal intensity to Neck Imaging Reporting and Data System categories 2 and 3 in primary sites of postsurgical oral cavity carcinoma provides incremental diagnostic value.** *AJNR Am J Neuroradiol* 2022;43:1018–23 CrossRef Medline
- Kumar I, Reza SO, Choudhary S, et al. **Performance of NI-RADS on CECT alone to predict recurrent head and neck squamous cell carcinoma after chemoradiotherapy: added value of RECIST 1.1.** *Indian J Radiol Imaging* 2022;32:151–58 CrossRef Medline
- Ashour MM, Darwish EA, Fahiem RM, et al. **MRI posttreatment surveillance for head and neck squamous cell carcinoma: proposed MR NI-RADS criteria.** *AJNR Am J Neuroradiol* 2021;42:1123–29 CrossRef Medline
- Abdelrahman AS, Ashour MM, Abdelaziz TT. **Predictive value of Neck Imaging Reporting and Data System (NIRADS) in CECT/CEMRI of laryngeal and oral cavity squamous cell carcinoma.** *Egypt J Radiol Nucl Med* 2020;51:241 CrossRef
- Aiken AH, Hudgins PA. **Neck Imaging Reporting and Data System.** *Magn Reson Imaging Clin N Am* 2018;26:51–62 CrossRef Medline
- Lee J, Kaht D, Ali S, et al. **Performance of the Neck Imaging Reporting and Data System as applied by general neuroradiologists to predict recurrence of head and neck cancers.** *Head Neck* 2022;44:2257–64 CrossRef Medline
- Abdelaziz TT, Abdel Razk AA, Ashour MM et al. **Interreader reproducibility of the Neck Imaging Reporting and Data system (NI-RADS) lexicon for the detection of residual/recurrent disease in**

- treated head and neck squamous cell carcinoma (HNSCC). *Cancer Imaging* 2020;20:1–10 CrossRef Medline
19. Dinkelborg P, Ro SR, Shnayien S, et al. **Retrospective evaluation of NI-RADS for detecting postsurgical recurrence of oral squamous cell carcinoma on surveillance CT or MRI.** *AJR Am J Roentgenol* 2021;217:198–206 CrossRef Medline
 20. Elsholtz FH, Erxleben C, Bauknecht HC, et al. **Reliability of NI-RADS criteria in the interpretation of contrast-enhanced magnetic resonance imaging considering the potential role of diffusion-weighted imaging.** *Eur Radiol* 2021;31:6295–304 CrossRef Medline
 21. Patel LD, Bridgham K, Ciriello J, et al. **PET/MR imaging in evaluating treatment failure of head and neck malignancies: a Neck Imaging Reporting and Data System-based study.** *AJNR Am J Neuroradiol* 2022;43:435–41 CrossRef Medline
 22. Baba A, Kurokawa R, Kurokawa M, et al. **Apparent diffusion coefficient for differentiation between benign post-treatment changes and recurrence in head and neck cancer: a systematic review and meta-analysis.** *AJNR Am J Neuroradiol* 2022;43:442–47 CrossRef Medline
 23. Baba A, Kurokawa R, Kurokawa M, et al. **Dynamic contrast-enhanced MRI parameters and normalized ADC values could aid differentiation of skull base osteomyelitis from nasopharyngeal cancer.** *AJNR Am J Neuroradiol* 2022;44:74–78 CrossRef Medline
 24. Mishra V, Singh A, Chen X, et al. **Application of liquid biopsy as multi-functional biomarkers in head and neck cancer.** *Br J Cancer* 2022;126:361–70 CrossRef Medline
 25. Flach S, Howarth K, Hackinger S, et al. **Liquid BIOPsy for MiNimal RESidual DiSease Detection in Head and Neck Squamous Cell Carcinoma (LIONESS): a personalised circulating tumour DNA analysis in head and neck squamous cell carcinoma.** *Br J Cancer* 2022;126:1186–95 CrossRef Medline
 26. Ioannidis JP. **The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses.** *Milbank Q* 2016;94:485–514 CrossRef Medline

Automatic Localization of the Pons and Vermis on Fetal Brain MR Imaging Using a U-Net Deep Learning Model

Farzan Vahedifard, Xuchu Liu, Jubril O. Adepoju, Shiqiao Zhao, H. Asher Ai, Kranthi K. Marathu, Mark Supanich, Sharon E. Byrd, and Jie Deng



ABSTRACT

BACKGROUND AND PURPOSE: An MRI of the fetus can enhance the identification of perinatal developmental disorders, which improves the accuracy of ultrasound. Manual MRI measurements require training, time, and intra-variability concerns. Pediatric neuroradiologists are also in short supply. Our purpose was developing a deep learning model and pipeline for automatically identifying anatomic landmarks on the pons and vermis in fetal brain MR imaging and suggesting suitable images for measuring the pons and vermis.

MATERIALS AND METHODS: We retrospectively used 55 pregnant patients who underwent fetal brain MR imaging with a HASTE protocol. Pediatric neuroradiologists selected them for landmark annotation on sagittal single-shot T2-weighted images, and the clinically reliable method was used as the criterion standard for the measurement of the pons and vermis. A U-Net-based deep learning model was developed to automatically identify fetal brain anatomic landmarks, including the 2 anterior-posterior landmarks of the pons and 2 anterior-posterior and 2 superior-inferior landmarks of the vermis. Four-fold cross-validation was performed to test the accuracy of the model using randomly divided and sorted gestational age-divided data sets. A confidence score of model prediction was generated for each testing case.

RESULTS: Overall, 85% of the testing results showed a $\geq 90\%$ confidence, with a mean error of < 2.22 mm, providing overall better estimation results with fewer errors and higher confidence scores. The anterior and posterior pons and anterior vermis showed better estimation (which means fewer errors in landmark localization) and accuracy and a higher confidence level than other landmarks. We also developed a graphic user interface for clinical use.

CONCLUSIONS: This deep learning-facilitated pipeline practically shortens the time spent on selecting good-quality fetal brain images and performing anatomic measurements for radiologists.

ABBREVIATIONS: AI = artificial intelligence; AP = anterior-posterior; DL = deep learning; GA = gestational age; SI = superior-inferior

CNS abnormalities are relatively common in fetuses, ranging from 0.1% to 0.2% in live births and 3% to 6% in stillbirths.¹ A diagnosis of fetal brain abnormalities at an early stage is essential. Fetal sonography is considered the criterion standard of anatomic

measurements. MR imaging is often performed when sonography is inconclusive to provide additional information for assessing fetal anatomy during all phases of gestation.² MR imaging provides superior soft-tissue contrast and spatial resolution for differentiating highly variable fetal brain tissue.³ Fetal MR imaging combined with fetal sonography increases confidence in the early detection of perinatal disorders of development.

Manual measurements have several disadvantages, including clinicians' training requirements, time commitment, and inter- and intraobserver variability.⁴ Radiologists must choose the highest quality image series without motion artifacts or missing anatomy and then identify and measure various anatomic structures.⁵ Small measurement errors may result in misdiagnosis and misguided pregnancy management.⁵ Accurate measurements of fetal brain anatomy are critical to differentiate hypoplastic, absent, or malformed brains from normal brain structures.⁶ Measurement errors can have significant consequences in clinical practice because they can lead to misdiagnosis and misguided pregnancy management.

Received February 6, 2023; accepted after revision August 2.

From the Department of Diagnostic Radiology and Nuclear Medicine (F.V., X.L., J.O.A., K.K.M., S.E.B.), Rush Medical College, Chicago, Illinois; Department of Biostatistics (S.Z.), Yale School of Public Health, New Haven, Connecticut; Division for Diagnostic Medical Physics (H.A.A., M.S.), Department of Radiology and Nuclear Medicine, Rush University Medical Center, Chicago, Illinois; and Department of Radiation Oncology (J.D.), Division of Medical Physics and Engineering, University of Texas Southwestern Medical Center, Dallas, Texas.

The funding of this project is from the Colonel Robert R. McCormick Professorship of Diagnostic Imaging fund at Rush University Medical Center, and the Swim Across America Pilot Project Grant from Rush University Medical Center.

Please address correspondence to Jie Deng, PhD, Department of Radiation Oncology, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390-9303; e-mail: jie.deng@utsouthwestern.edu



Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A7978>

Fetal MR imaging interpretation also requires specialized training. However, there is a shortage of pediatric neuroradiologists, resulting in limited availability.

During the past decade, artificial intelligence (AI) algorithms, specifically deep learning (DL), have significantly advanced image-recognition tasks.⁷ The machine learning approaches have the potential to aid in the early detection of these issues, thereby enhancing the diagnostic and follow-up processes.

By means of AI, processing of fetal brain MR imaging has investigated models that automatically predict specific landmarks and segmentation. Various AI models (primarily convolutional neural networks and U-Net) were used.^{8,9} Some models achieved an accuracy of $\geq 95\%$. AI could aid in the pre- and postprocessing¹⁰ and reconstruction,¹¹ predicting gestational age (GA) (with an accuracy of 1 week),¹² fetal brain extraction,¹³ and fetal brain segmentation.^{11,14} AI could also help in GA prediction, fetal motion detection, motion tracking, pose estimation,¹⁵ and super-resolution reconstruction.¹¹ Recently, several publications developed AI models for automatic fetal brain anatomic measurements in a biparietal diameter,¹⁶ which were derived from identified landmarks after several preprocessing steps such as computation of an ROI, reference section selection, segmentation, midsagittal line and fetal brain orientation, and, finally, measurements.

In this work, we focused on identifying 2 anterior-posterior (AP) landmarks of the pons and 2 AP and 2 superior-inferior (SI) landmarks of the vermis. All the landmarks on each structure were predicted simultaneously using U-Net multisegmentation features. We exploited U-Net¹⁷ to determine imaging features surrounding a landmark point and calculated the probability of any image pixel being the defined landmark point.

In addition, the image pixel with the highest probability within the output Gaussian distribution mask was predicted as the landmark point by the U-Net model. Finally, we developed a tool that could be extended to clinical use on the basis of the prediction model to help radiologists select the best image series for interpretation and perform fetal brain anatomic measurements more efficiently.

MATERIALS AND METHODS

Database

This retrospective study, approved by the institutional review board with a waiver of consent, included 55 fetal MR imaging studies at different GAs. The studies were selected from a database of pregnant women who underwent routine clinical fetal screening at Rush University Medical Center, Chicago, Illinois, between 2007 and 2020. All the selected studies confirmed normal fetal brain development based on radiology reports. Expert pediatric neuroradiologists performed image-quality screening and landmark annotation on the exported sagittal T2-weighted HASTE images. Of these 55 patients, some patients had >1 image series, so the total number of image series was 100. We added image series for data augmentation, increasing the data set.

Six landmarks, including AP landmarks on the pons and AP/SI landmarks on the vermis (drawn by the radiologist), served as the ground truth. In addition, manual biometric measurements on the pons and vermis were performed according to the standard clinical recommendations.¹⁸

MR Imaging Protocol

Fetal MR images were obtained at our institution using Siemens 1.5T MR imaging scanners (Siemens, Erlangen, Germany), without sedation. Single-shot HASTE images were acquired in the axial, coronal, and sagittal planes with the following parameters: TR = 1400 ms, TE = 120 ms, FOV = 230×230 mm², and section thickness/gap = 3/0 mm, under free breathing. The fetal age range was 20–39 weeks, and the cases did not involve twins or significant maternal risk factors.

Image Preprocessing

All images were resized to 512×512 and augmented through rotating, flipping, adding Gaussian noise, motion blurring, median blurring, contrast-limited adaptive histogram equalization, sharpening, embossing, random brightness contrast adjustment, and random hue saturation adjustment using an open-source library (albumentations.ai; <https://albumentations.ai/>) with default parameter settings.

The study used original MR images without super-resolution reconstruction or other quality-selection preprocessing. A preselection process ensured suitable images for the study, focusing on landmark visibility and differentiation from neighboring structures. Exclusion criteria included motion artifacts, which hindered landmark identification and blurred anatomic borders. Consistency was maintained using midsagittal planes for pons and vermis measurements, avoiding oblique planes. Clinicians selected the data set for pons and vermis annotations, which underwent independent verification by radiologists and AI engineers for accuracy and consistency.

An innovative aspect of our research lies in the use of U-Net for landmark predictions. Rather than using the conventional binary segmentation output of 0/1, we modified the final output layer to generate a distribution map indicating the probability of landmark locations. This novel approach allowed us to extract valuable information from the U-Net model and precisely predict the positions of the landmarks, further enhancing the significance of our article.

Model Performance Validation

The U-Net model was used to fit the Gaussian distribution function. After the radiologist provided specific landmarks, the AI model calculated the distribution probability of these landmarks on the image by collecting the image features.

Model performance evaluation focuses on the probability relationship between image features and landmarks. Given an unsuitable image for labeling the vermis, AI will try to determine where the most likely landmark point will be. However, confidence in this point may be low due to differences in the features learned by image and AI.

With this probability distribution, we feed all the images into the pipeline and filter out the most suitable images for the physician's annotation.

U-Net Model as a Transforming Function

The encoding path of the U-Net model incorporates convolutional and max-pooling layers for feature extraction and dimension reduction, while the decoding path uses up-sampling and

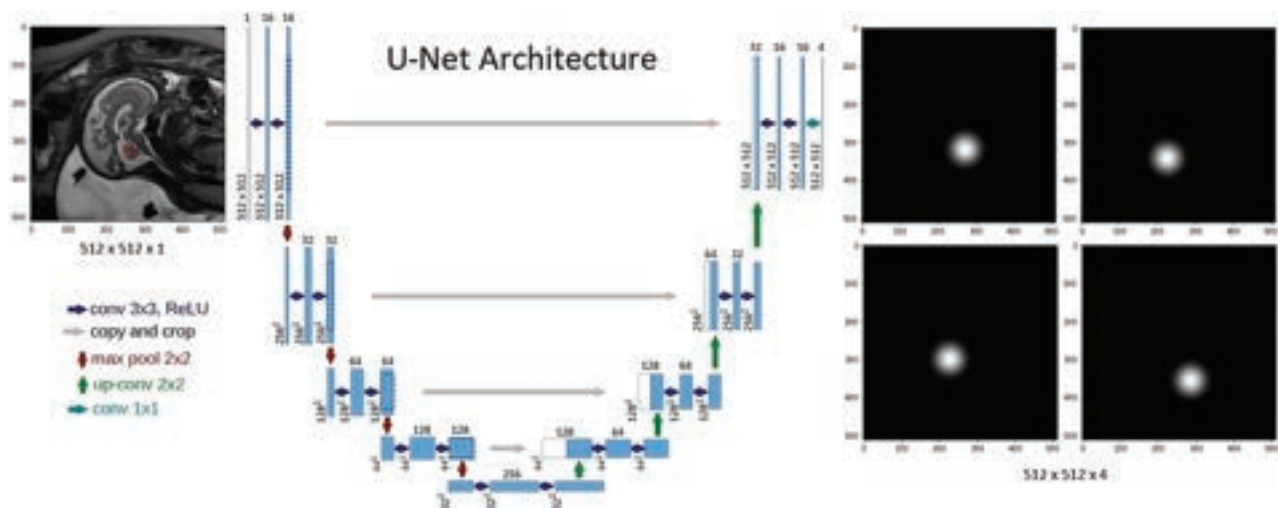


FIG 1. The architecture of the U-Net model for landmark detection. The U-Net converts a sagittal fetal brain image into gray-scale masks in which the vertex represents the locations of different landmark points. Up-conv indicates up-sampling operation; conv, convolutional layer; ReLU, Rectified Linear Activation.

concatenation for restoring spatial resolution and creating segmentation maps.

A U-Net model was built to transform an input MR image into a gray-scale image mask with its vertex representing the location of the predicted landmark point (Fig 1).

The transforming function $f(x, y)$ can be expressed as the impact of an arbitrary point (x, y) within the mask on the predicted landmark point (x_0, y_0) (Equation 1).

$$\text{Equation 1 } f(x, y) = f\left(\sqrt{(x - x_0)^2 + (y - y_0)^2}\right) = f(r),$$

where r represents the radial distance. The gray-scale mask is a rotationally symmetric function (radius of a circle: R) with a Gaussian distribution centered on the annotated landmark point (Fig 2), where R is proportional to the fronto-occipital radius of the fetal brain (Fig 3). We chose the Gaussian distribution function $f(r; \mu, \sigma)$ as $f(r)$ and simplified it by setting the mean and $SD\sigma = \frac{R/2}{\sqrt{2\pi}}$ as in Equation 2:

$$\text{Equation 2 } f(r; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(r-\mu)^2}{2\sigma^2}} = \frac{2}{R} e^{-\frac{4\pi^2 r^2}{R^2}}.$$

In practical applications, we removed the coefficient $\frac{2}{R}$ and replaced it in Equation 1.

$$\text{Equation 3 } f(r) = e^{-\frac{4\pi^2 r^2}{R^2}}, \quad r \in [0, R]; \quad f(r) = 0, \quad r > R, \\ \frac{r}{R} \in [0, 1].$$

We customized the U-Net for landmark-prediction reliability by implementing a Gaussian output. This step permits AI to consider landmark-location uncertainty. The probability of a predicted point being the desired landmark was represented using this Gaussian function. The confidence score for each landmark was based on this probability. The model was trained to detect multiple landmarks simultaneously by generating separate masks representing the Gaussian probability of each landmark.

We balanced model complexity and computational efficiency by optimizing the hyperparameters of the model. We used a 3×3 kernel size to capture local contextual information efficiently. The channel depth gradually doubled after each max-pooling operation to learn more complex representations at different levels. The number of layers was chosen considering the task complexity and available computational resources, finding a suitable balance for the model.

Our study demonstrated that splitting the DL model outputs into separate models for pons and vermis landmarks resulted in improved accuracy. This approach allowed fine-tuned adjustments and enhanced detection of each landmark. By focusing on specific imaging features, the individual models improved the identification of anatomic structures. Compared with predicting all landmarks at once with a single model, this approach achieved superior performance and increased accuracy in detecting landmarks in fetal brain MR imaging.

Model Training

In the training process, a normalized weighted binary mean squared error loss function was used to compensate for the data imbalance. Other model parameters included batch size = 15 for both training and validation and epoch = 100 with an early stop after 10 epochs of loss increasing.

We used the Adam optimizer to update model weights efficiently, the EarlyStopping strategy to prevent overfitting, and the ModelCheckpoint callback to save weights of the best-performing model, ensuring reproducibility. The entire training process on the graphics processing unit Nvidia GeForce RTX 2080 Ti 11 Gb (NVIDIA GeForce RTX 2080 Ti 11Gb) took <2 hours.

K-Fold cross-validation

Two 4-fold cross-validation methods were implemented for model training and testing (Fig 4). In the first method, the data set was divided into 4 groups by a sorted range of GAs (ie, 20–26 weeks, 27–29 weeks, 30–33 weeks, and 34–39 weeks) (Fig 4A). In the second method, the data set was randomly divided with

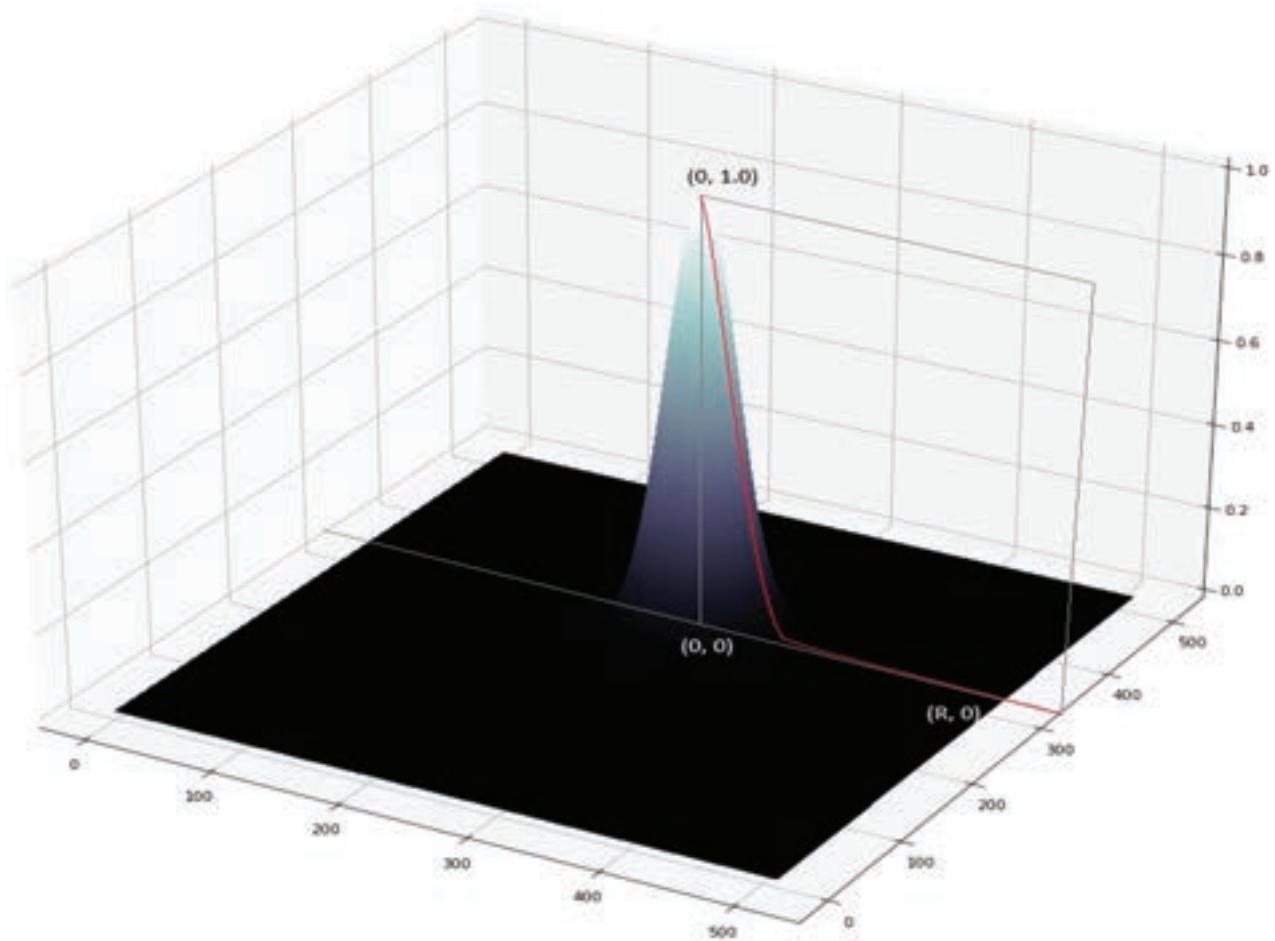


FIG 2. The cut surface of a 3D rotationally symmetric Gaussian distribution function with the radius (R), and the top represents the landmark point position.

mixed GAs without overlapping among all groups (Fig 4B). In each cross-validation fold, 3 groups of patient images were used as the training data set, and the other group, as the testing data set.

RESULTS

Model Performance

The randomly divided mixed GA method outperformed the sorted GA-divided method by providing smaller prediction errors ($P < .001$) with higher confidence scores ($P < .001$).

In the first cross-validation method, in which the data set was divided by sorted GA weeks, the prediction error distribution with associated confidence scores for 6 predicted landmarks in 100 image series (ie, the number of total landmarks: $6 \times 100 = 600$) is shown in Fig 5 (A, scatterplot; B, contour line plot). Among all 600 predicated landmark locations, 73% (440/600) showed a confidence score of $>90\%$, with a mean prediction error of <2.25 mm. Among the 6 landmarks, the anterior vermis and anterior and posterior pons were best predicted with fewer errors and higher confidence scores (Online Supplemental Data). The anterior and posterior pons and anterior vermis had significantly fewer errors compared with the posterior vermis ($P < .01$, $P < .01$, and $P < .001$, respectively). Additionally, the anterior pons had a significantly higher confidence score than the superior/

inferior/posterior vermis ($P < .05$, $P < .05$, and $P < .01$, respectively), while the posterior pons had a significantly higher confidence score than the superior/posterior vermis ($P < .05$ and $P < .01$, respectively).

In the second cross-validation method, in which the data set was randomly divided with mixed GA weeks, the prediction error distribution with an associated confidence score for the 600 predicted landmarks is shown in Fig 6 (A, scatterplot; B, contour line plot). Among all 600 predicated landmarks, 85% (511/600) showed a confidence score of $>90\%$, with a mean prediction error of <2.22 mm. Among the 6 landmarks, the posterior pons was the best-predicted landmark, with the smallest error and highest confidence score (Online Supplemental Data). The posterior pons had significantly lower error compared with all other landmarks ($P < .05$ for the anterior pons and anterior vermis, and $P < .001$ for the superior/inferior/posterior vermis). Additionally, the posterior pons had significantly higher confidence scores than the superior, inferior, and posterior vermis ($P < .05$, $P < .05$, and $P < .01$, respectively).

Automatic Landmark Detection

We evaluated the differences between manual landmark localization performed by a radiologist and an expert pediatric neuroradiologist, as shown in the Table. The variations between their

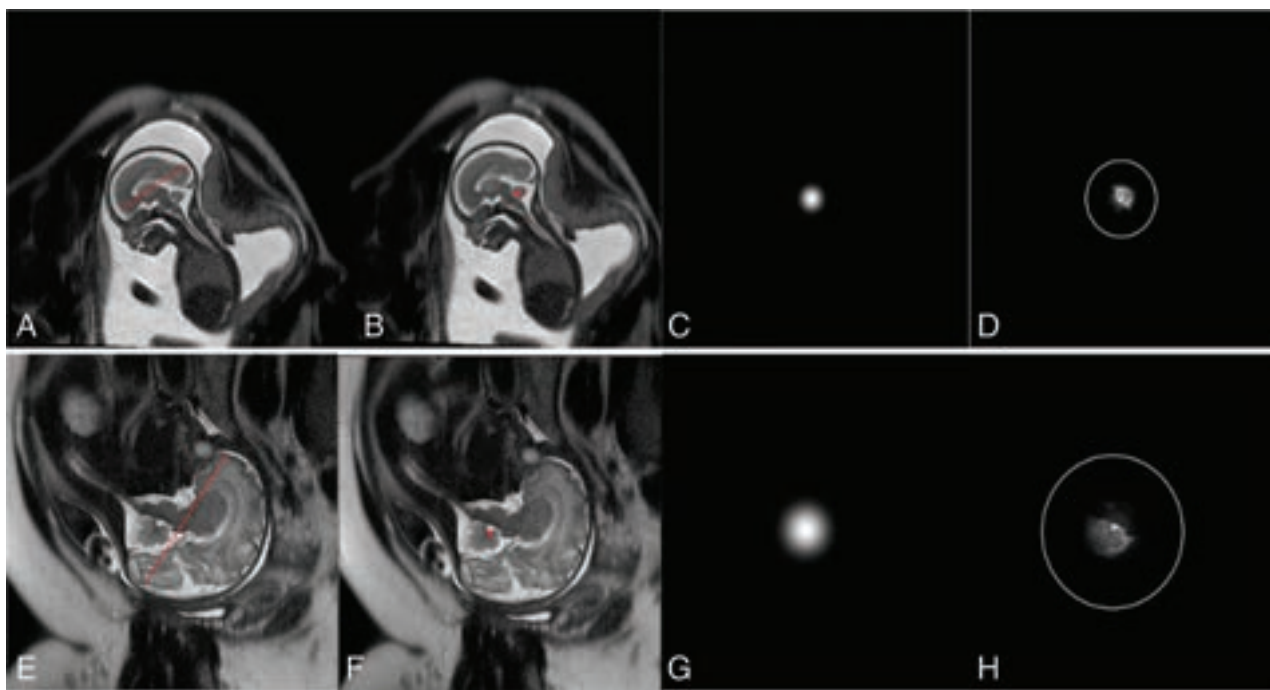


FIG 3. Fronto-occipital radius of the fetal brain determines the radius of Gaussian distribution function in a patient at GA week 20 (A–D) and another at GA week 33 (E–H). The left 2 columns are MR images with a fronto-occipital radius (A and E) and annotated landmark points on the vermis (B and F). The third column (C and G) shows the image mask with the Gaussian distribution used in model training. The fourth column shows the image area surrounding the landmark (D and H) determined by the Gaussian distribution function, with an added *white circle* indicating the radius range.

manual measurements ranged from a mean of 0.42 (SD, 0.59) mm for vermis1 to 1.87 (SD, 1.81) mm for vermis2. These disparities emphasize the presence of interrater variability and the possibility of measurement inconsistencies with manual assessments.

By automating landmark identification and using DL capabilities, our AI model consistently provides more reliable and consistent measurements.

Figure 7 shows examples of model-predicted landmark locations with biometric measurements between each pair demonstrated and compared with manual detections performed by the radiologist.

Despite variations in image quality and white noise levels, the AI system maintained accuracy, highlighting its robustness and adaptability. Figure 7 shows the precision and confidence of our AI model in 1 case. The left picture has a pixel spacing of 0.41, resulting in a relatively low resolution. The middle and right images have a pixel spacing of 0.35, which means better resolution. However, the right image has significantly more white noise points. The Online Supplemental Data show confidence and measurements in these 3 different series. The predictions of AI had an error rate below 0.5 mm and provided confidence scores for each prediction, enhancing trust in its outputs. (The AI provided confidence scores for each prediction and selected the optimal predictions with an error rate below 0.5 mm, thereby enhancing the trust of its outputs). Therefore, this AI model offers a reliable, accurate, and consistent tool for measurements.

Distance Measurement. The Online Supplemental Data illustrate the “distance measurements” of the pons and vermis, comparing

manual and AI measurements and the corresponding errors. The error-to-total measurement ratio is reasonable. For instance, when the confidence threshold is set above 90%, the pons distance from the average is 10.81 mm, with an error of 1.12 mm.

Statistical Analysis

Data were transformed to achieve a normal distribution before we conducted statistical testing. The prediction errors of the 6 landmarks and confidence scores were compared using paired *t* tests across the two 4-fold cross-validation methods. The Tukey Honest Significant Difference test was used to compare and rank the prediction errors and confidence scores for each landmark among all 6 landmarks. All statistical analyses were performed using R Studio (<http://rstudio.org/download/desktop>). A significance level of $P < .05$ was used.

Clinical Pipeline with Graphic User Interface

An interactive tool was developed on the basis of our developed model. This graphic user interface was created to help radiologists identify the landmarks of the pons and vermis and obtain biometric measurements on fetal MR imaging more efficiently (Fig 8). This graphic user interface will also provide confidence for each suggested image. To accept, reject, or modify the landmark prediction, the clinicians can judge according to the AI-provided confidence.

DISCUSSION

We proposed a novel U-Net DL model that automatically detects anatomical landmarks on the “pons” and “vermis” in fetal brain

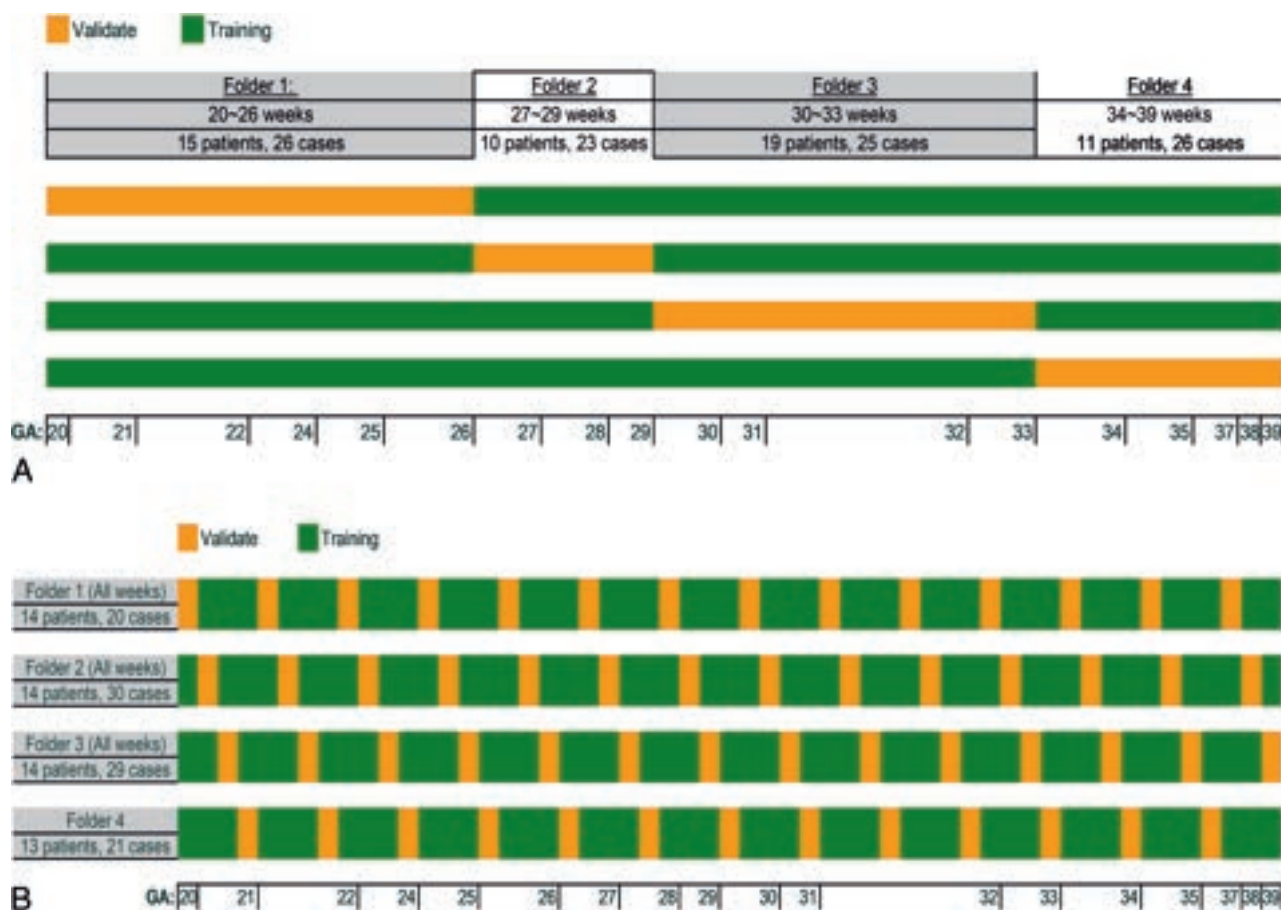


FIG 4. Two 4-fold cross-validation methods for DL-model training and testing. Method 1 divided the data sets by sorting the ranges of GA weeks (A). Method 2 divided the data sets randomly with mixed GA weeks (B).

MR imaging. Using these landmark locations, we predicted important fetal biometric parameters, including vermis diameter and height, and pons diameter. The critical component of our model was using U-Net as a transformational function to generate a gray-scale image mask with a Gaussian distribution by extracting the imaging features adjacent to the center of the mask. Although we only tested this model for landmarks on the pons and vermis with promising results, it may also be applied to detect landmarks on other brain structures with asymmetric image features.

This study initially used a 2-stage anisotropic 3D U-Net to detect fetal brain ROIs. A reference image section on which the landmarks were identified using a Fetal Measurement by Landmarks was used. A Gaussian Mixture Model estimated the landmark measurement reliability. Compared with fetal MR imaging radiologists, the model yielded a 95% confidence interval agreement of 3.70 mm for cerebral biparietal diameter, 2.20 mm for biparietal bone diameter, and 2.40 mm for transcerebral diameter. Our study is the first to automatically detect landmarks on the pons and vermis on fetal MR imaging using a U-Net as a function with very few training parameters, reducing the computational complexity and shortening the training time.

The accuracy of our model depends on learning imaging features around landmarks. Given the varying fetal brain size and appearance at different GAs, we set the mask size proportional to

the fronto-occipital diameter of each brain for adequate feature extraction. Among the 6 landmarks, 2 on the pons and 4 on the vermis, the anterior-posterior pons and anterior vermis had better accuracy, possibly due to distinctive adjacent image features like the fourth ventricle, aiding model learning.

We built a pipeline for automatic batch processing of multiple image series for landmark prediction. It selected the reference section on the basis of the highest confidence score and skipped poor-quality images. The reference section was presented to the radiologists for review, and they were alerted to manually adjust any landmark prediction with a confidence score of <0.8 .

In model training and validation, we implemented 3 schemes for 4-fold cross-validation (dividing folds by GA weeks versus dividing folds by mixed weeks). The model trained and validated with mixed GA weeks provides overall better accuracy compared with the sorted GA week-divided approach, suggesting that a large number of imaging features extracted from fetal MR images with various white/gray matter contrast and anatomic details during fetal brain development are essential to include in model training.

In similar studies, Dovjak et al¹⁹ conducted manual studies on cerebellar vermian lobulation and vermis/brainstem-specific markers using prenatal MR images. Their research improved hindbrain malformation classification and provided insights into vermian growth patterns. In contrast, our study introduces a DL model for automated identification and measurement of

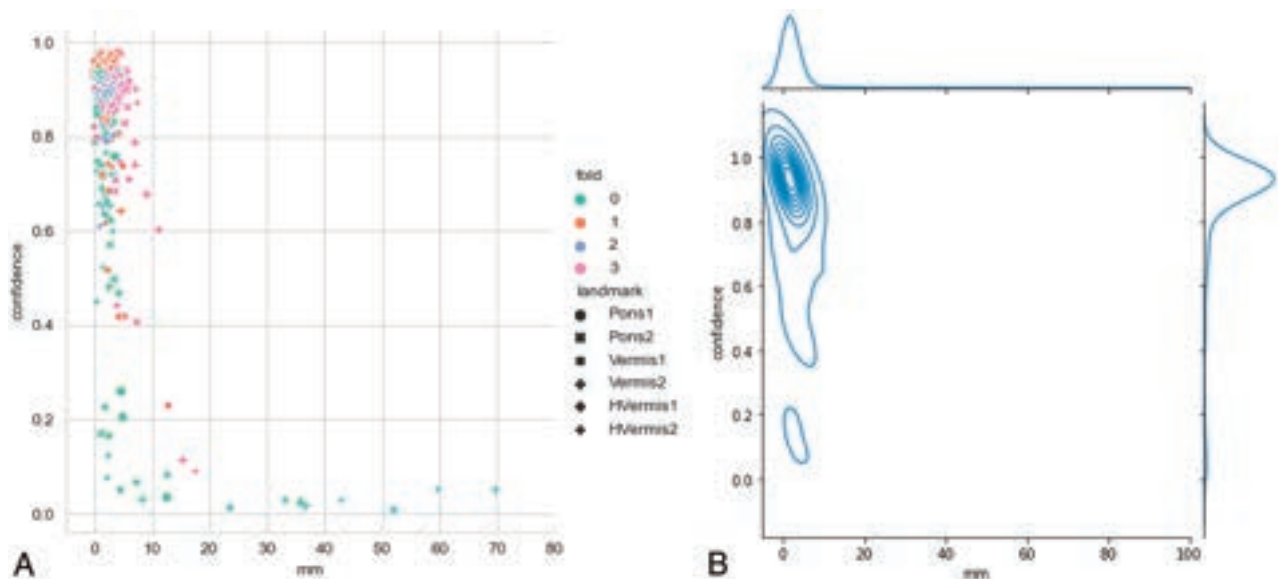


FIG 5. The scatterplot (A) and contour line plot (B) of the prediction error distribution with associated confidence scores in the sorted GA week validation method. The x-axis represents the distance (millimeters) between the predicted landmark and the ground truth, and the y-axis represents the confidence score of the prediction.

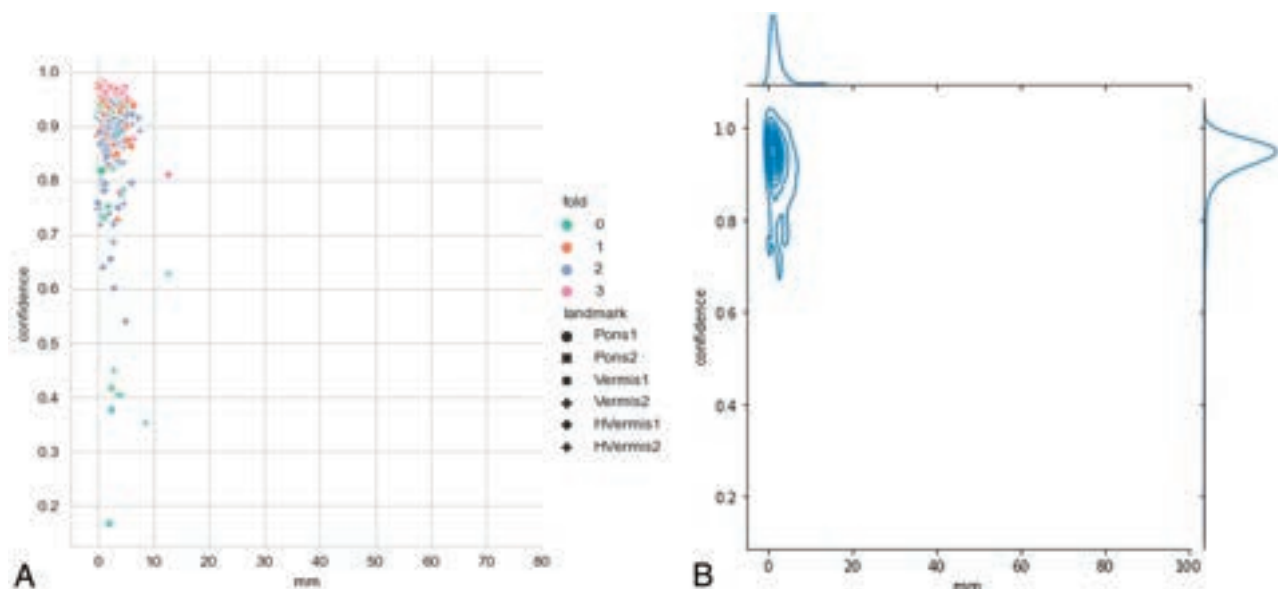


FIG 6. The scatterplot (A) and contour line plot (B) of the distribution of prediction error with associated confidence scores in the randomly mixed GA week validation method. The x-axis represents the distance (millimeters) between the predicted landmark and the ground truth, and the y-axis represents the confidence score of the prediction.

pons and vermis landmarks in fetal MR imaging. Our approach enhances efficiency, reduces errors, and offers confidence scores for predictions.

Advantages of Our Model

1) The benefit of automated landmark localization: The U-Net algorithm offers the advantage of automated landmark localization and provides associated confidence scores. This feature allows us to process all sagittal sequences of a patient and automatically identify the top 5 fetal brain MR image series with the highest confidence for physicians to choose from or manually adjust.

2) Time efficiency of the U-Net algorithm: Although the initial implementation and training of the U-Net algorithm require time and resources, its application on new fetal brain MR images is quick and efficient. The model can automatically identify landmarks and provide measurements without manual intervention. On our hardware (GeForce RTX 2080 Ti 12G), the average processing time for our U-Net model to predict 6 markers for an image is as low as 0.23 seconds, which is negligible compared with the time required by a physician for manual screening and measurements.

Compared manual landmark localization conducted by a radiologist and an expert pediatric neuroradiologist

	Pons1	Pons2	Vermis1	Vermis2	HVermis1	HVermis2
Mean (mm)	1.41	0.79	0.42	1.87	1.28	1.51
SD (mm)	1.12	0.76	0.59	1.81	1.38	1.68

Note:—Pons1 indicates Anterior landmark of Pons; Pons2, Posterior landmark of Pons; Vermis1, Anterior landmark of Vermis; Vermis2, Posterior landmark of Vermis; HVermis1, Superior landmark of Vermis (height of vermis); HVermis2, Inferior landmark of Vermis (height of vermis).

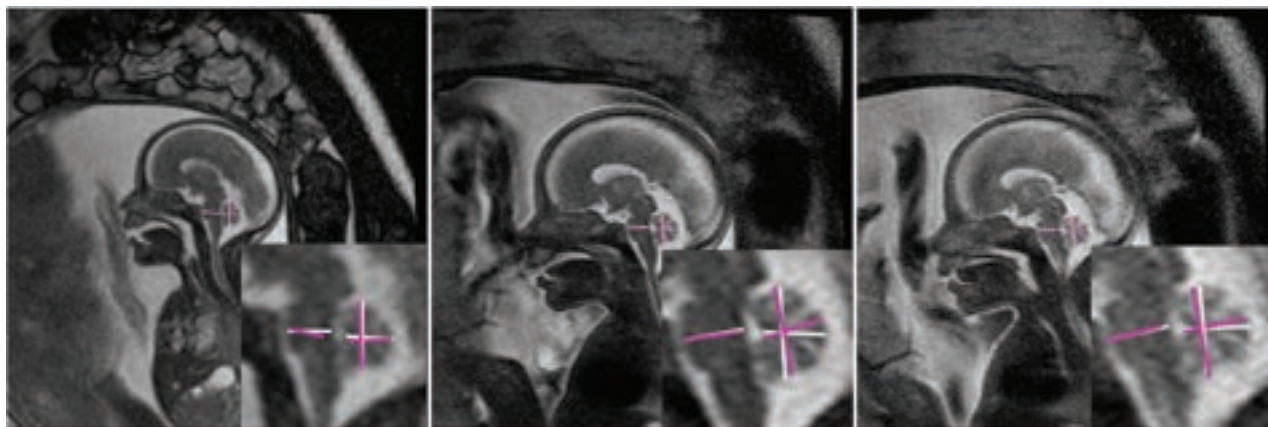


FIG 7. Representative image examples of model-predicted landmark locations with biometric measurements. (The white line is manual annotations by radiologists, and the purple line is DL-model predicted measurements). Three patients (A, 22 weeks; B, 22 weeks; C, 27 weeks) with accurate model-predicted landmarks compared with radiologists.

- 3) Consistency and reduced interobserver variability: The U-Net algorithm offers consistent landmark identification across different images and cases, reducing interobserver variability. This standardized approach leads to more reliable and reproducible measurements.
- 4) Accommodation of different resolutions and image quality: Our U-Net model robustly handles variations in resolution and evaluates image quality. It assigns confidence scores to each image slice, prioritizing high-quality images for accurate measurements and increased diagnostic confidence, resulting in errors of < 0.5 mm, regardless of resolution and image-quality differences.
- 5) Potential of the U-Net algorithm: The model demonstrates promising accuracy levels, potentially matching or exceeding human expert annotations. Further validation and optimization can enhance its reliability for posterior fossa biometry quantification.
- 6) Choosing linear measurements: Our study prioritized using linear measurements, specifically the diameter of the pons and vermis, due to their well-established clinical relevance and diagnostic utility. These measurements have been widely adopted in clinical practice and have demonstrated their effectiveness in detecting various brain abnormalities, including vermian hypoplasia, Dandy-Walker malformation, and pontocerebellar hypoplasia. Moreover, monitoring changes in the pons and vermis offers valuable insights into the neurologic development of the fetus, identifying potential issues and evaluating posterior fossa lesions.

The Run Algorithm on a Public Data Set

For an additional test, we ran the algorithm on the larger, publicly available Fetal Tissue Annotation and Segmentation Challenge

(<https://feta.grand-challenge.org/>), resulting in good accuracy. Please see Online Supplemental Data, more codes and examples are provided.

Limitations

This study had some limitations: First, the sample size in each GA week range was small, possibly leading to insufficient model training. We used several strategies to enhance the accuracy and generalizability of our AI model to address the limitation of a limited number of cases. The MR imaging selection was conducted by experienced medical professionals, ensuring a high-quality training, validation, and testing data set. We used “transfer learning,” enabling the model to identify distinctive features across a wider image range, thereby increasing its applicability in diverse clinical scenarios. Despite the size of our data set, the promising performance of the model in this pilot study suggests adaptability across different institutional settings. It provides landmark coordinates and confidence values, giving clinicians flexibility in MR image selection.

Second, this study did not assess fetal brain biometry in pathologic cases because of a paucity of cases with abnormal pons and vermis structures across different GA weeks. The study was limited to a single scanner platform and could potentially be restricted to a single cohort due to the limitation in the cohort. As a result, the generalizability of this study and its utility on large-scale data platforms may be limited. We recognize the need for further validation with larger, diverse data sets to ascertain the robustness and generalizability of our model in varied clinical environments.

Ensemble learning is a valuable option for future studies if additional data are available. This technique involves using all models obtained through 4-fold cross-validation and selecting

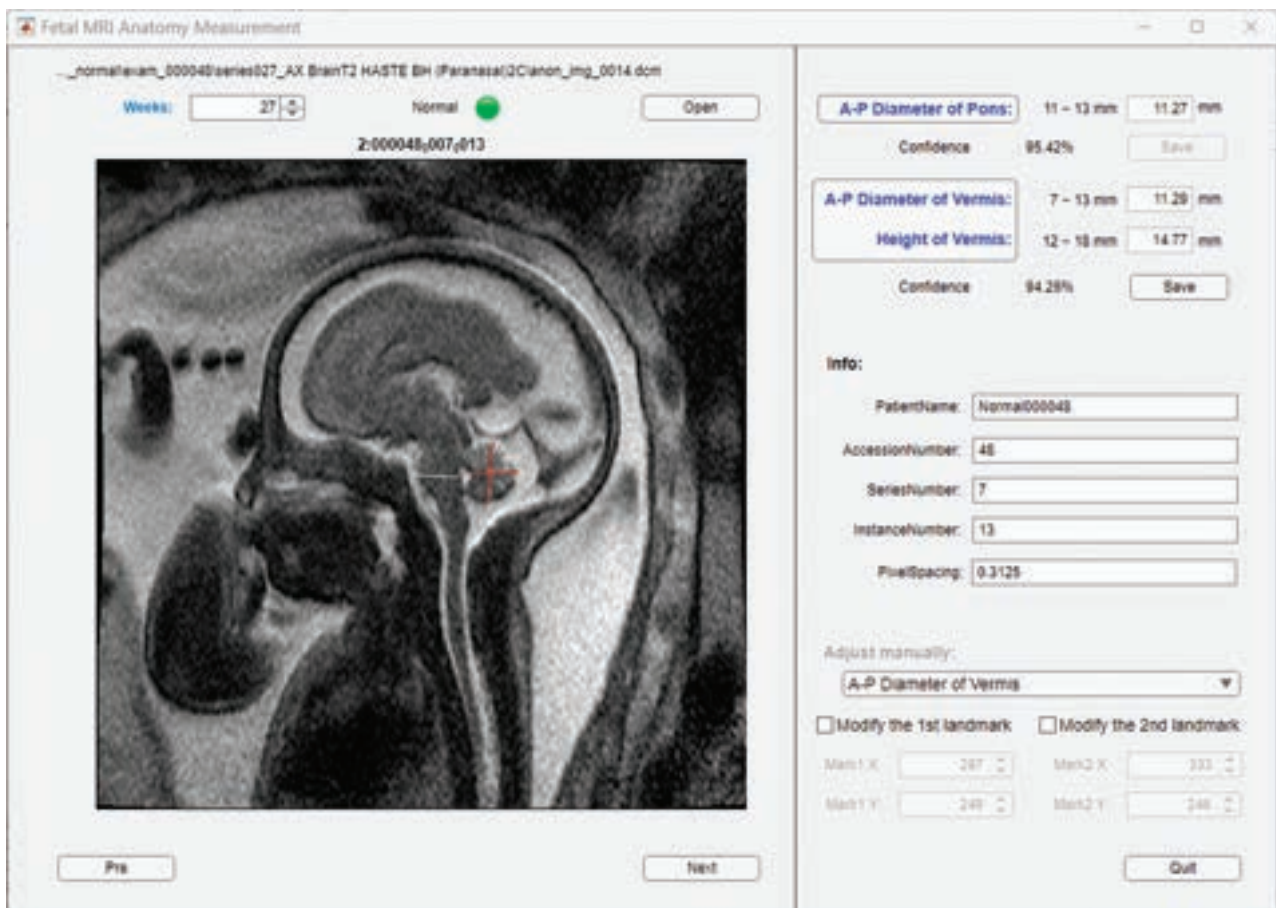


FIG 8. An interactive tool integrating DL-model-based prediction of landmarks. It helps the radiologist quickly locate, confirm, or adjust the landmarks on the automatically selected image section. After the landmark locations are confirmed, the distance between 2 related landmarks on a particular brain structure is calculated. The white line is the “Pons” AP diameter, and red crossed lines are “Vermis” AP diameter and height.

the landmark with the highest confidence, though it comes at the cost of increased processing time.

CONCLUSIONS

A U-Net model was developed to detect AP landmarks on the pons and AP/SI landmarks on the vermis. Our pipeline includes image series screening and selection, landmark prediction that improves radiologists’ efficiency and time in identifying landmarks, performing anatomic measurements, and screening high-quality images. Using a U-Net-based DL model, we achieved a mean error of <2.22 mm and a $\geq 90\%$ confidence score in 85% of the testing cases, resulting in improved estimation accuracy with reduced errors.

While manual measurements by radiologists often yield robust results, our AI model brings value by significantly reducing interrater variability and measurement errors. It accurately identifies high-confidence landmarks and optimizes image selection, even in instances in which blurred margins due to motion artifacts are present.

We also established a pipeline, graphic user interface, consisting of imaging selection and landmark prediction, followed by an interactive second check tool to help radiologists quickly locate, confirm, or adjust the landmarks on the autoselected image slices.

Using U-Net as a transformational function, our model accurately extracts imaging features around landmarks, particularly for the anterior and posterior pons and anterior vermis.

We validated this algorithm on a public data set, demonstrating good accuracy. The AI model addresses interrater variability, reduces measurement errors, and saves time, presenting advantages over manual measurements. Implementing AI-driven automation enables comprehensive and efficient fetal brain MR imaging assessment, potentially enhancing radiologists’ efficiency and diagnostic accuracy and improving patient outcomes in fetal brain MR imaging analysis.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

1. Cater SW, Boyd BK, Ghathe SV. **Abnormalities of the fetal central nervous system: prenatal US diagnosis with postnatal correlation.** *Radiographics* 2020;40:1458–72 CrossRef Medline
2. Hibbeln JF, Shors SM, Byrd SE. **MRI: is there a role in obstetrics?** *Clin Obstet Gynecol* 2012;55:352–66 CrossRef Medline
3. Shi Y, Xue Y, Chen C, et al. **Association of gestational age with MRI-based biometrics of brain development in fetuses.** *BMC Med Imaging* 2020;20:125 CrossRef Medline

4. Joskowicz L, Cohen D, Caplan N, et al. **Inter-observer variability of manual contour delineation of structures in CT.** *Eur Radiol* 2019;29:1391–99 CrossRef Medline
5. Warrander LK, Ingram E, Heazell AE, et al. **Evaluating the accuracy and precision of sonographic fetal weight estimation models in extremely early-onset fetal growth restriction.** *Acta Obstet Gynecol Scand* 2020;99:364–73 CrossRef
6. Rajagopalan V, Deoni S, Panigrahy A, et al. **Is fetal MRI ready for neuro-imaging prime time? An examination of progress and remaining areas for development.** *Dev Cogn Neurosci* 2021;51:100999 CrossRef Medline
7. Hosny A, Parmar C, Quackenbush J, et al. **Artificial intelligence in radiology.** *Nat Rev Cancer* 2018;18:500–10 CrossRef Medline
8. Vahedifard F, Adepoju JO, Supanich M, et al. **Review of deep learning and artificial intelligence models in fetal brain magnetic resonance imaging.** *World J Clin Cases* 2023;11:3725–35 CrossRef Medline
9. Vahedifard F, Ai HA, Supanich MP, et al. **Automatic ventriculomegaly detection in fetal brain MRI: A step-by-step deep learning model for novel 2D-3D linear measurements.** *Diagnostics* 2023;13:2355 CrossRef Medline
10. Chen Z, Pawar K, Ekanayake M, et al. **Deep learning for image enhancement and correction in magnetic resonance imaging: state-of-the-art and challenges.** *J Digit Imaging* 2023;36:204–30 CrossRef Medline
11. Ebner M, Wang G, Li W, et al. **An automated framework for localization, segmentation and super-resolution reconstruction of fetal brain MRI.** *Neuroimage* 2020;206:116324 CrossRef Medline
12. Shen L, Zheng J, Lee EH, et al. **Attention-guided deep learning for gestational age prediction using fetal brain MRI.** *Sci Rep* 2022;12:1408 CrossRef Medline
13. Išgum I, Benders MJ, Avants B, et al. **Evaluation of automatic neonatal brain segmentation algorithms: the NeoBrainS12 challenge.** *Med Image Anal* 2015;20:135–51 CrossRef Medline
14. Makropoulos A, Counsell SJ, Rueckert D. **A review on automatic fetal and neonatal brain MRI segmentation.** *Neuroimage* 2018;170:231–48 CrossRef Medline
15. Gagoski B, Xu J, Wighton P, et al. **Automated detection and reacquisition of motion-degraded images in fetal HASTE imaging at 3 T.** *Magn Reson Med* 2022;87:1914–22 CrossRef Medline
16. Avidris N, Yehuda B, Ben-Zvi O, et al. **Automatic linear measurements of the fetal brain on MRI with deep neural networks.** *Int J Comput Assist Radiol Surg* 2021;16:1481–92 CrossRef Medline
17. Ronneberger O, Fischer P, Brox T, eds. **U-Net: convolutional networks for biomedical image segmentation.** In: Navab N, Hornegger J, Wells WM, et al, eds. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015.* Springer Link; 2015
18. Garel C, Delezoide AL. *MRI of the Fetal Brain: Normal Development and Cerebral Pathologies.* Springer-Verlag; 2004
19. Dovjak GO, Diogo MC, Brugger PC, et al. **Quantitative fetal magnetic resonance imaging assessment of cystic posterior fossa malformations.** *Ultrasound Obstet Gynecol* 2020;56:78–85 CrossRef Medline

Dandy-Walker Phenotype with Brainstem Involvement: 2 Distinct Subgroups with Different Prognosis

C.A.P.F. Alves,¹ J. Sidpra,² A. Manteghinejad,³ S. Sudhakar,⁴ F.V. Massey,⁵ K.A. Aldinger,⁶ P. Haldipur,⁷ L.T. Lucato,⁸ S.F. Ferracioli,⁹ S.R. Teixeira,¹⁰ Ö. Öztekin,¹¹ D. Bhattacharya,¹² A. Taranath,¹³ S.P. Prabhu,¹⁴ D.M. Mirsky,¹⁵ S. Andronikou,¹⁶ K.J. Millen,¹⁷ A.J. Barkovich,¹⁸ E. Boltshauser,¹⁹ W.B. Dobyns,²⁰ M.J. Barkovich,²¹ M.T. Whitehead, and K. Mankad²²



ABSTRACT

BACKGROUND AND PURPOSE: Although cardinal imaging features for the diagnostic criteria of the Dandy-Walker phenotype have been recently defined, there is a large range of unreported malformations among these patients. The brainstem, in particular, deserves careful attention because malformations in this region have potentially important implications for clinical outcomes. In this article, we offer detailed information on the association of brainstem dysgenesis in a large, multicentric cohort of patients with the Dandy-Walker phenotype, defining different subtypes of involvement and their potential clinical impact.

MATERIALS AND METHODS: In this established multicenter cohort of 329 patients with the Dandy-Walker phenotype, we include and retrospectively review the MR imaging studies and clinical records of 73 subjects with additional brainstem malformations. Detailed evaluation of the different patterns of brainstem involvement and their potential clinical implications, along with comparisons between posterior fossa measurements for the diagnosis of the Dandy-Walker phenotype, was performed among the different subgroups of patients with brainstem involvement.

RESULTS: There were 2 major forms of brainstem involvement in patients with Dandy-Walker phenotype including the following: 1) the mild form with anteroposterior disproportions of the brainstem structures “only” (57/73; 78%), most frequently with pontine hypoplasia (44/57; 77%), and 2) the severe form with patients with tegmental dysplasia with folding, bumps, and/or clefts (16/73; 22%). Patients with severe forms of brainstem malformation had significantly increased rates of massive ventriculomegaly, additional malformations involving the corpus callosum and gray matter, and interhemispheric cysts. Clinically, patients with the severe form had significantly increased rates of bulbar dysfunction, seizures, and mortality.

CONCLUSIONS: Additional brainstem malformations in patients with the Dandy-Walker phenotype can be divided into 2 major subgroups: mild and severe. The severe form, though less prevalent, has characteristic imaging features, including tegmental folding, bumps, and clefts, and is directly associated with a more severe clinical presentation and increased mortality.

ABBREVIATION: DW = Dandy-Walker phenotype

Cystic malformations of the posterior fossa manifest with varied imaging patterns and diagnoses and have a broad range of etiologies and clinical outcomes.¹ A classic disorder captured in this group is the historically termed Dandy-Walker malformation,

now Dandy-Walker phenotype (DW). Several causes have been implicated in the pathogenesis of DW, including chromosomal abnormalities and genetic variants,^{2–7} disruptive/acquired events in fetal life such as prenatal exposure to viruses, drugs, and

Received May 15, 2023; accepted after revision July 18.

From the Division of Neuroradiology (C.A.P.F.A., A.M., S.R.T., S.A., M.T.W.), Department of Radiology, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; Unit of Neuroradiology (J.S., S.S., K.M.), Great Ormond Street Hospital for Children, National Health Service Foundation Trust, London, United Kingdom; Developmental Biology & Cancer Section (J.S., K.M.), University College London Great Ormond Street Institute of Child Health, London, United Kingdom; Unit of Functional Neurosurgery (F.V.M.), National Hospital for Neurology & Neurosurgery, London, UK; Center for Integrative Brain Research (K.A.A., P.H., K.J.M.), Seattle Children's Research Institute, Seattle, Washington; Departments of Pediatrics and Neurology (K.A.A., P.H., K.J.M.), University of Washington, Seattle, Washington; Department of Radiology, Division of Neuroradiology (L.T.L., S.F.F.), Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil; Department of Neuroradiology (Ö.Ö.), Bakırçay University, Çi li Education and Research Hospital, İzmir, Turkey; Department of Neuroradiology (D.B.), Royal Victoria

Hospital, Belfast, UK; Department of Medical Imaging (A.T.), Women's and Children's Hospital, Adelaide, South Australia, Australia; Department of Radiology, Neuroradiology Division (S.P.P.), Boston Children's Hospital, Boston, Massachusetts; Department of Radiology, Neuroradiology Division (D.M.M.), Children's Hospital Colorado, Aurora, Colorado; Department of Neuroradiology (A.J.B., M.J.B.), University of California, San Francisco, San Francisco, California; Department of Pediatric Neurology (E.B.), University Children's Hospital, Zürich, Switzerland; and Department of Genetics and Metabolism (W.B.D.), University of Minnesota, Minneapolis, Minnesota. Please address correspondence to Cesar Augusto Alves, MD, PhD, Division of Neuroradiology, Department of Radiology, Children's Hospital of Philadelphia, 324 South 34th St, Philadelphia, PA 19104; e-mail: alvesc@chop.edu; @cesaralvesneuro

Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A7967>

maternal ill-health^{8,9} as well as fetal sporadic posthemorrhagic cerebellar events.^{10,11} However, there have also been recent modifications in formerly well-established associations. For example, the previously reported association between congenital melanocytic nevus syndrome and DW¹² has been re-evaluated in a recent case report and literature review that emphasizes a potential misleading connection between these 2 disorders.¹³ Nevertheless, it is important to note that mesenchymal-based gene action has been shown to play a role in the development of the Dandy-Walker phenotype in some circumstances.¹⁴

Although cardinal imaging features for the diagnostic criteria of DW have recently been defined by our group,¹⁵ there is a large variability of additional unreported malformations in children with DW. DW encompasses a broad range of many potential associated abnormalities, most of them recognized in the supratentorial brain, such as agenesis/dysgenesis of the corpus callosum, hydrocephalus, and abnormalities of cortical migration, among others.^{16,17}

Contrary to the frequent description of supratentorial findings, brainstem involvement in patients with DW is far less frequently cited in the literature. Nevertheless, the brainstem is frequently involved,¹⁵ and unsurprisingly so, given that its development is closely related to that of the cerebellum.¹⁸⁻²⁰ Furthermore, when affected, the brainstem is likely to have a major impact on the brainstem and cerebellar circuitry essential for normal function.²¹ Proper brainstem function is essential for autonomic nervous system functions (ie, breathing, heart rate). Thus, brainstem involvement in DW requires critical attention due to its potential role in the pathogenesis and clinical outcome of these patients.²⁰

In this article, we present detailed information on the brainstem abnormalities found in patients with DW, including the variability of imaging patterns, their potential implications for human neurodevelopment, and their clinical impact.

MATERIALS AND METHODS

This retrospective multicenter, multinational study focuses on a subset group of patients with DW with brainstem malformations. These patients were selected from a previously published cohort of 329 patients who were confirmed to have DW on the basis of the modern diagnostic radiologic criteria established by Whitehead et al.¹⁵ Among the 329 patients, 220 showed some degree of dysmorphism. Of these, 147 (67%) had extrinsic mass effect caused by the surrounding structures, which provided a better explanation for their brainstem abnormality, for example the compression of the brainstem against the clivus, most commonly due to hydrocephalus. Since no definitive signs of primary brainstem dysgenesis were observed in this particular subset of patients, these patients were excluded. For our analysis, we included and retrospectively reviewed the MR imaging studies and clinical records of a cohort of 73 patients with DW and brainstem malformations. We provide a detailed evaluation of the different brainstem malformative findings, including the following:

1. The brainstem region involved (midbrain/tectum, pons, and/or medulla oblongata)
2. Features of hypoplasia or hyperplasia of each of the brainstem segments and their ratios to categorize potential developmental anterior-posterior patterning defects²⁴

3. Malformation of the brainstem with abnormalities resulting in deformity of normal brainstem architecture (abnormal folding, bumps, and clefts) and disturbances of axonal orientation and/or guidance

4. Additional supratentorial malformations, such as abnormalities of cortical migration, commissural dysgenesis/agenesis, and the presence of cerebral interhemispheric cysts.

Qualitative analyses and quantitative measures including size and morphology of the posterior fossa and choroid plexus, tectal choroidea complex location, and torcular location were performed following our previous standardized approach.¹⁵ Neuroimaging evidence of prior injury (hemorrhage and/or encephalomalacia) was also reviewed, and comparisons of all variables were performed to search for potential differences and to define subgroups in our cohort.

All cases were independently reviewed by senior neuroradiologists at each participating institution, and clinical and radiologic features were recorded. Qualitative and quantitative measures and the reporting of additional brainstem malformations were reviewed via independent analysis by the first author (C.A.P.F.A.). Subsequently, all discrepancies and disagreement were solved by a third reviewer (K.J.M.), a pediatric neuroradiologist with >15 years of experience, and a final consensus with the first author (C.A.P.F.A.) was determined for those discrepant cases. The minimum requirements for the inclusion of postnatal MR imaging were having at least 2 series of T1-weighted imaging and T2-weighted imaging with section thicknesses of ≤ 5 mm. Additional sequences were reviewed in most cases (when available). All MR imaging was performed at 1.5T or 3T.

Continuous variables are presented as the mean (SD), while categorical variables are presented as the percentage frequency. The Kolmogorov-Smirnov test was used to examine the normality of continuous variables, and the Mann-Whitney *U* test was used to assess differences among continuous variables. The Pearson χ^2 test and Fisher exact test were used to evaluate the associative significance between 2 categorical variables. Statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS, Version 26; IBM).

The study was conducted after site-specific institutional review board approval with informed consent waived due to its retrospective nature.

RESULTS

Seventy-three patients (44 female, 60.3%; 29 male, 39.7%) were eligible for inclusion (age range, 1 day to 24 years of age). Fifty-seven MR images (78.1%) were obtained using 1.5T scanners, and 16 (21.9%) were obtained using 3T scanners. Genetic abnormalities were present as the presumed cause of DW in 22 (30.1%) patients. Fifteen patients presented with known risk factors for DW: Eight (11%) had vascular injuries; 5 (6.8%) were associated with maternal diseases, including maternal diabetes and congenital heart disease; and 2 (2.7%) were associated with prenatal infections, including prenatal cytomegalovirus. In the remaining 36 (49.3%) patients, there were no known risk factors, and no genetic analysis was performed. The complete descriptive analysis of study subjects is available in the Table 1.

Table 1: Descriptive analysis of patients with DW with brainstem involvement

Variable	
Sex (No.) (%)	
Male	29 (39.7%)
Female	44 (60.3%)
Age	
Age (mo)	25 mo (IQR = 10 days to 15.6 mo)
Birth status (No.) (%)	
Premature	20 (27.4%)
Full-term	35 (47.9%)
Unknown	18 (24.7%)
Imaging features (No.) (%)	
Severe macrocrania	24 (32.9%)
Extreme vermian hypoplasia	12 (16.4%)
Cerebellar dysplasia	27 (37%)
Anteroposterior disproportion	67 (91.8%)
Hydrocephalus	43 (58.9%)
Corpus callosum agenesis	37 (50.7%)
Interhemispheric cyst	11 (15.1%)
Gray matter migration	18 (24.7%)
Calcification or hemorrhage	12 (16.4%)
Fastigial recess angle (mean)	167.1° (SD, 46.6°)
Posterior fossa perimeter (mean) (mm)	172.4 (SD, 110.6)
Clinical features (No.) (%)	
Seizure	31 (42.5%)
Artificial airway dependency	32 (43.8%)
Enteral feeding tube dependency	40 (54.8%)
Mortality	15 (20.5%)

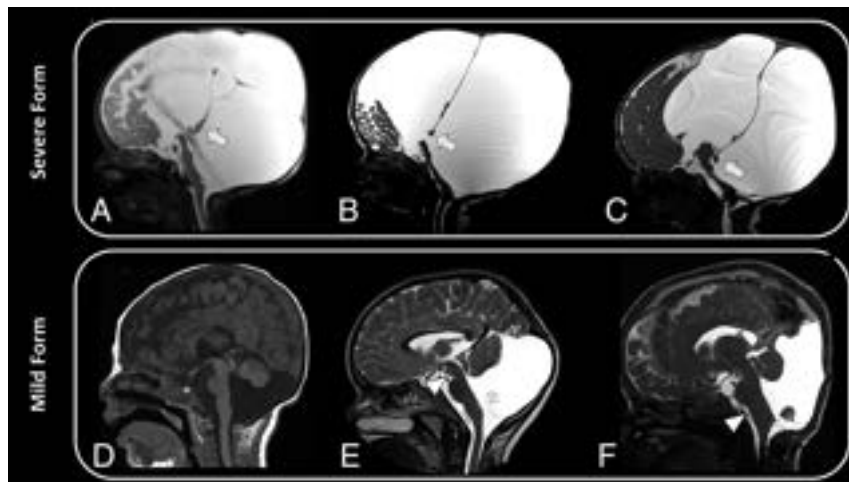


FIG 1. A severe form of brainstem malformation in 3 different patients (A–C): Sagittal T2WI demonstrates a massive enlargement of the posterior fossa with severe hypoplasia and deformity of the vermis (arrows, A–C) and elevation of the torcula. There is associated agenesis of the corpus callosum, severe hydrocephalus, and interhemispheric cysts. A mild form of brainstem malformation in 3 different patients (D–F): Sagittal T1WI (D) shows hypoplasia and rotational displacement of the inferior cerebellar vermis associated with hypoplasia of the pons. Sagittal T2WI (E and F) shows a disproportionately increased craniocaudal size of the midbrain (arrowhead, E) and diffusely increased size of the medulla (arrowhead, F).

Two subgroups of brainstem malformations were observed in our DW cohort with forms defined as the following: 1) mild, 57 (78.1%) patients with changes limited to the size and proportion of the brainstem structures, and 2) severe, 16 (21.9%) patients in whom malformation of the brainstem included posterior tegmental dysplasia with an abnormal posterior fold, bumps, and clefts in the tegmentum (Figs 1 and 2). The comparative

imaging and clinical analyses of these 2 groups are presented in the Table 2. The Kolmogorov-Smirnov test was used to evaluate the normality of the posterior fossa perimeter and fastigial recess angle, with *P* values of .001 and .019, respectively. Therefore, due to the significant difference between these 2 variable distributions from a normal distribution, the Mann-Whitney *U* test was used to compare these 2 variables between the groups.

The mild form of brainstem malformation in the DW cohort was characterized mainly by the presence of pontine hypoplasia (44/57; 77%) and, more rarely, hypoplasia or hyperplasia of other segments of the brainstem, including the medulla and midbrain/tectum (13/57; 23%) (Fig 1).

Patients in the severe brainstem DW cohort had more complex malformations: All presented with tegmental dysgenesis characterized by architectural disorganization of ≥ 1 segment of the brainstem, including evidence of a posterior folding, bumps, or clefts, with flattening of the ventral portions of the brainstem, imaging features with parallels to those of axonal guidance disturbance noted in patients with tegmental and bulbar cap dysplasia (Fig 3 and Online Supplemental Data).^{22,23} Among the significant imaging features observed in this group, the most important findings to highlight include macrocrania in 10 (62.5%), mostly secondary to severe ventriculomegaly; massive enlargement of the posterior fossa and increased rates of the extreme vermian hypoplasia in 9 (56.3%); increased rates of additional malformative features involving the corpus callosum in 12 (75.0%) patients, cerebral interhemispheric cysts in 9 (56.3%) subjects (Fig 4), and gray matter abnormalities in 10 (62.5%) patients compared with the mild group. Clinical differences were also observed between groups, including increased rates of

bulbar dysfunction in 12 (75%) patients and seizures in 11 (68.8%) patients from the severe group, findings indicative of a more severe clinical presentation, revealed by increased mortality in 7 (46.7%) patients, all with bulbar dysfunction.

Among the 73 patients, genetic causes were noted in 22 (30.1%), 5 of 16 (31.2%) had severe brainstem involvement, while 17 of 57 (29.8%) had mild brainstem involvement. No statistical

differences were noted between subgroups. The most common results included 3 patients with trisomy 13, 2 patients with tetrasomy 9p, 2 patients with chromosome 6 deletion involving the *FOXC1* gene, 2 patients with *KMT2D* mutation (Kabuki syndrome), and 2 patients unbalanced X:3 translocations: arr[hg19] Xp22.33p21.2(168,547–29,318,254) x3.

DISCUSSION

Neurologic development is variable in patients with DW,¹ with the clinical course depending not only on the degree of cerebellar hypoplasia but also on the extent and severity of associated CNS malformations. In this study, patients with DW with associated brainstem involvement were evaluated to understand the potential imaging variability of malformations in this structure, their potential associations, and its clinical impact on the DW population. Although all patients with DW and brainstem

involvement shared characteristic posterior fossa cystic malformation with the required features for the current diagnostic criteria,¹⁵ there were 2 different imaging groups of patients with brainstem involvement: 1) those with the mild form: considered a less severe phenotype in which changes were related only to the size and proportion of the brainstem structures (anterior-posterior embryologic defect);²⁴ and 2) those with the severe form in which an additional dysgenetic appearance of the brainstem was noted, with a peculiar posterior tegmentum fold, bump, and cleft. In line with our concept of the critical role of the brainstem in neurologic functions, these patients with extensive morphologic abnormalities in the brainstem were also more severely affected clinically, including a significantly increased frequency of bulbar and other autonomic nervous system dysfunctions, seizures, and higher mortality rates, compared with the mild form. From the genetic perspective, we

noted that genetic abnormalities and associated syndromes in our DW cohort with brainstem malformation (30%) are almost 2 times higher than expected in the overall DW population (16%).³ This observed difference suggests that patients with DW with brainstem malformations may be better candidates for an extensive genetic work-up.

The mild brainstem phenotype was the most frequent form in our cohort (78%), characterized mainly by the presence of pontine hypoplasia and, more rarely, hypoplasia or hyperplasia of other segments of the brainstem, including the medulla and midbrain. Apart from the brainstem dysmorphology resulting from a presumed abnormal embryologic anteroposterior

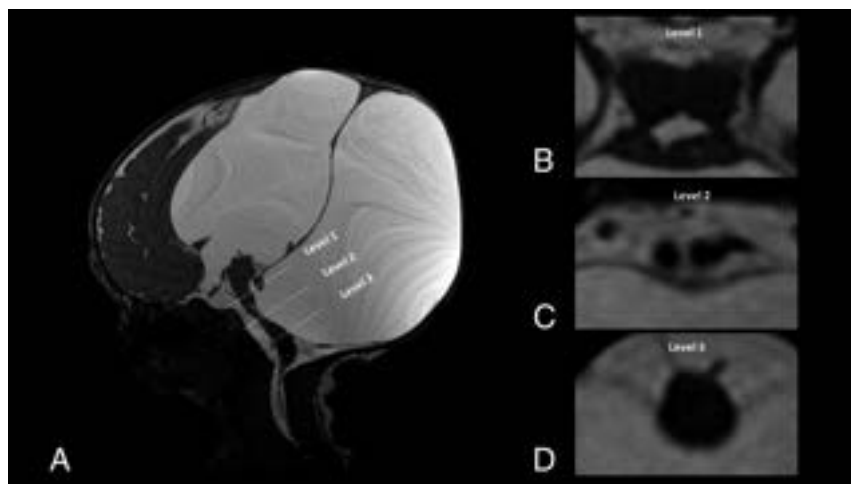


FIG 2. A severe form of brainstem malformation with tegmental dysplasia: Sagittal and axial T2WI demonstrates 3 different craniocaudal brainstem levels, including a complete cleft of the pons (C, level 2). There is associated agenesis of the corpus callosum, severe hydrocephalus, and interhemispheric cyst.

Table 2: Comparison of imaging and clinical characteristics between patient groups by the presence of a tegmental brainstem dysplasia (severe brainstem malformation)

Variable	Tegmental Brainstem Dysplasia (No.) (%)		OR (95% CI)	P Value
	Present (Severe Group)	Absent (Mild Group)		
Mortality	7 (46.7%)	8 (15.1%)	4.922 (1.392–17.399)	.01 ^a
Severe macrocrania	10 (62.5%)	14 (24.6%)	5.119 (1.576–16.629)	.004 ^b
Prematurity	6 (42.87%)	14 (34.4%)	1.446 (0.419–4.997)	.559 ^a
Genetic abnormalities	5 (31.2%)	17 (29.8%)	1.912 (0.319–11.471)	.677 ^b
Bulbar dysfunction	12 (75%)	20 (36.4%)	5.250 (1.492–18.470)	.006 ^a
Enteral feeding tube dependency	12 (75%)	28 (50.9%)	2.893 (0.830–10.087)	.087 ^a
Seizure	11 (68.%)	20 (35.7%)	3.960 (1.205–13.018)	.019 ^a
Extreme vermian hypoplasia	9 (56.25%)	3 (5.3%)	23.143 (5.034–106.401)	<.001 ^b
Cerebellar dysplasia	11 (68.8%)	16 (28.1%)	5.638 (1.690–18.805)	.003 ^a
Hydrocephalus	13 (81.3%)	30 (52.6%)	3.900 (1.002–15.177)	.04 ^a
Corpus callosum agenesis	12 (75.0%)	25 (43.9%)	3.840 (1.104–13.358)	.028 ^a
Interhemispheric cyst	9 (56.3%)	2 (3.5%)	35.357 (6.317–197.9)	<.001 ^b
Gray matter migration	10 (62.5%)	8 (14%)	10.208 (2.901–35.923)	<.001 ^b
Calcification or hemorrhage	4 (25%)	8 (14%)	2.042 (0.526–7.924)	.444 ^b
Posterior fossa perimeter (mean) (mm)	252.4 (SD, 110.3)	149.94 (SD, 100.6)	—	.002 ^c

Note:—En dash indicates not applicable.

^a Pearson χ^2 test.

^b Fisher exact test.

^c Mann-Whitney *U* test.

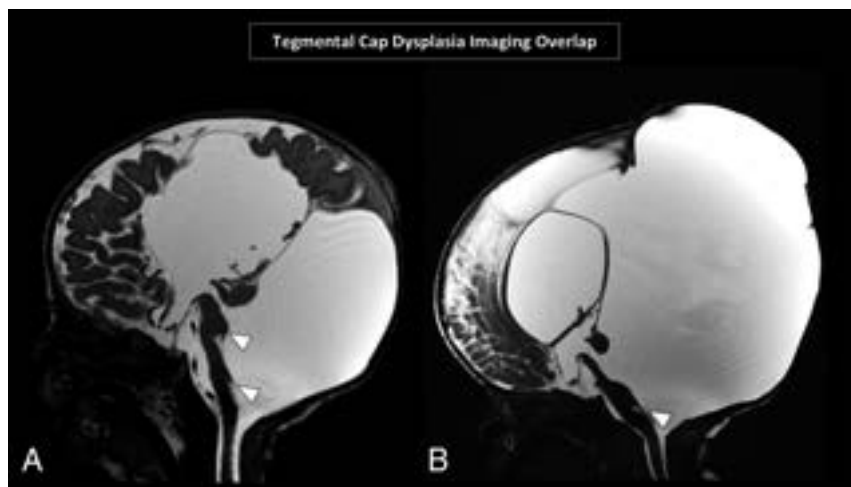


FIG 3. Two patients with severe brainstem malformations and DW. Sagittal T2WI shows severe deformity of the brainstem and flattening of the pons associated with abnormalities of the tegmentum, including a posterior cleft and white matter bump (arrowheads), overlapping features with tegmental cap dysplasia.

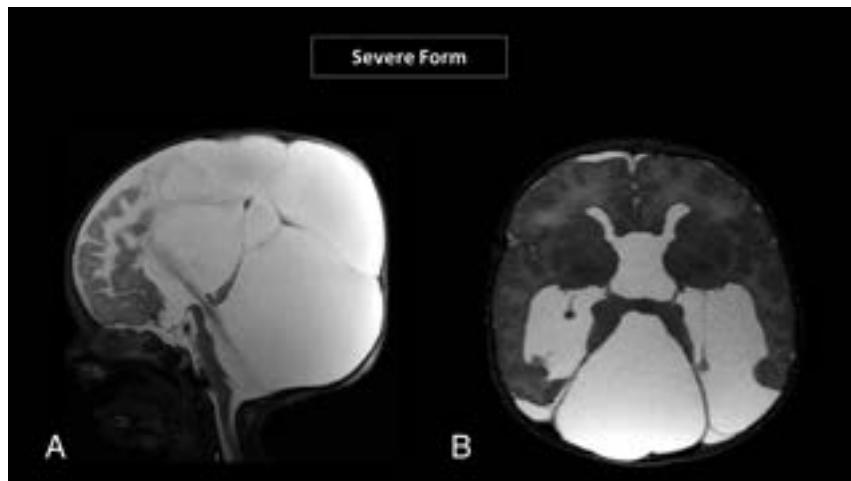


FIG 4. Sagittal (A) and axial T2WI (B) highlighting imaging features of a severe form of DW with brainstem malformation, including massive hydrocephalus and massive enlargement of the posterior fossa, along with complete agenesis of the corpus callosum and interhemispheric cysts.

defect, the measurements and appearance of the posterior fossa, including posterior fossa perimeter, fastigial angle blunting, and choroid plexus/taenia tela choroidea complex morphology, were not significantly different from those in the overall cohort already described in 2022.¹⁵ There were also no differences between the additional abnormalities of the brain or clinical features commonly noted among patients with DW.^{25,26}

Patients with severe brainstem malformation, on the other hand, had additional findings that clearly distinguish them from the mild brainstem DW cohort. In particular, we identified the presence of a posterior tegmental dysplasia as a hallmark of this group. The patients with tegmental brainstem dysplasia characterized by folding, bumps, and clefts in the tegmentum were highly associated with having more severe/extreme vermian hypoplasia, massive hydrocephalus, and larger posterior fossa perimeters compared with the mild brainstem subgroup. Additionally, dysgenesis and agenesis of the corpus callosum

and supratentorial interhemispheric cysts were significantly associated with this phenotype. Most interesting, during the retrospective imaging review and analysis, we noted that some of the patients with brainstem tegmentum dysplasia shared similar features already described in tegmental and bulbar cap dysplasias, including focal posterior protuberance of the brainstem with a small and flat pons, features related to an underlying abnormal axonal guidance pathophysiology.^{22,23} These phenotypic similarities shed light on the likely embryologic developmental intersection between both disorders.

There is substantial evidence that disruption of the rhombic lip underlies the cerebellar DW.^{27,28} The rhombic lip is a progenitor zone located along the dorsal edge of the hindbrain in the rhombomere, one of the anterior hindbrains adjacent to the roof of the fourth ventricle. It gives rise to multiple cell lines that are relevant to the phenotypes we have described in this report. Progenitors from rhombomere 1 rhombic lip develop into cerebellar granule cell progenitors, which migrate over the cerebellar anlage and then undergo massive proliferation, driving most cerebellar growth.²⁸ Earlier in development, these progenitors also give rise to the glutamatergic pontine nuclei neurons, which earlier migrate over the developing cerebellar anlage and continue ventrally to form the pontine nuclei at the brainstem.^{18–20} Failure of generation and/or migration of these early-born brainstem neurons in mice

results in phenotypes similar to those seen in our patient population.²⁰ Additionally, rhombic lip progenitors from rhombomere 1 and other more posterior hindbrain regions also give rise to choroid plexus epithelial cells, which play a role in the expansion of the fourth ventricle plexus from its lateral edges. Mouse models have shown that very early misspecification of cerebellar rhombic lip progenitors toward the choroid plexus fate, at the expense of cerebellar and brainstem fates, contributes to the expansion of the fourth ventricle roof and both cerebellar and brainstem hypoplasia.^{14,29,30}

Therefore, the simultaneous association of brainstem and cerebellar malformation and even fourth ventricle expansion is anticipated in patients with DW. The severity of these abnormalities is likely related to the timing and extent of the insult during early embryologic stages. Early insults during the critical period of rhombic lip development can lead to more severe and extensive abnormalities, whereas insults occurring later in development

may result in milder phenotypes. Thus, the precise timing and extent of the disruption to the rhombic lip progenitors determine the severity of the disorder, emphasizing the importance of early embryologic events in shaping the neurologic outcomes in individuals with DW.

The variability of neuroimaging findings from mild to severe brainstem groups associated with varied severity of clinical presentations and outcomes as well as different rates of underlying genetic causes reinforce the importance of detailed imaging descriptions and neuroimaging stratification to improve the understanding of these conditions and to better support neurologists and geneticists in the clinical care and genetic counseling of these patients and their families. In patients with DW, the brainstem should be carefully scrutinized for differences in size, shape, and/or proportions. Direct comparison with aged-matched normal brain MR imaging examinations may be useful in questionable cases.

Although our study yielded notable findings, it is important to acknowledge its limitations. First, the retrospective study design and cross-sectional analysis focused primarily on postnatal studies. This approach overlooked the consideration of neurodevelopmental outcomes and prenatal imaging findings, which could have provided valuable insights. As a result, there is the possibility of selection and misclassification bias within our groups, which hinders the establishment of a stronger causal relationship. To enhance the reliability of our conclusions, future research should address these limitations by incorporating prospective designs, comprehensive assessments of prenatal factors, and animal model studies. Another limitation was the small number of severe cases in which DW and posterior dysplasia of the brainstem were noted. However, given that these cases were collected from numerous tertiary referral centers with expertise in pediatric neuroimaging, this small number may reflect the true rarity of severe brainstem dysplasia in postnatal life.

CONCLUSIONS

It is of paramount importance to recognize additional brainstem malformations in patients with DW. These malformations can be divided into 2 major subgroups, mild and severe. The severe form, though less prevalent, has characteristic imaging features, including a tegmental dysplasia, and is associated with a more severe clinical presentation that is directly related to brainstem dysfunction and increased mortality.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

- Hirsch JF, Pierre-Kahn A, Renier D, et al. **The Dandy-Walker malformation: a review of 40 cases.** *J Neurosurg* 1984;61:515–22 CrossRef Medline
- Grinberg I, Northrup H, Ardinger H, et al. **Heterozygous deletion of the linked genes *ZIC1* and *ZIC4* is involved in Dandy-Walker malformation.** *Nat Genet* 2004;36:1053–55 CrossRef Medline
- Aldinger KA, Timms AE, Thomson Z, et al. **Redefining the etiologic landscape of cerebellar malformations.** *Am J Hum Genet* 2019;105:606–15 CrossRef Medline
- Aldinger KA, Lehmann OJ, Hudgins L, et al. ***FOXC1* is required for normal cerebellar development and is a major contributor to chromosome 6p25.3 Dandy-Walker malformation.** *Nat Genet* 2009;41:1037–42 CrossRef Medline
- Jalali A, Aldinger KA, Chary A, et al. **Linkage to chromosome 2q36.1 in autosomal dominant Dandy-Walker malformation with occipital cephalocele and evidence for genetic heterogeneity.** *Hum Genet* 2008;123:237–45 CrossRef Medline
- Yahyaoui R, Espinosa MG, Gómez C, et al. **Neonatal carnitine palmitoyltransferase II deficiency associated with Dandy-Walker syndrome and sudden death.** *Mol Genet Metab* 2011;104:414–16 CrossRef Medline
- Liao C, Fu F, Li R, et al. **Dandy-Walker syndrome and microdeletions on chromosome 7 [in Chinese].** *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 2012;29:48–51 CrossRef Medline
- Murray JC, Johnson JA, Bird TD. **Dandy-Walker malformation: etiologic heterogeneity and empiric recurrence risks.** *Clin Genet* 1985;28:272–83 CrossRef Medline
- Soares de Oliveira-Szejnfeld P, Levine D, de Oliveira Melo AS, et al. **Congenital brain abnormalities and Zika virus: what the radiologist can expect to see prenatally and postnatally.** *Radiology* 2016;281:203–18 CrossRef Medline
- Limperopoulos C, Folkerth R, Barnewolt CE, et al. **Posthemorrhagic cerebellar disruption mimicking Dandy-Walker malformation: fetal imaging and neuropathology findings.** *Semin Pediatr Neurol* 2010;17:75–81 CrossRef Medline
- Pichiecchio A, Decio A, Di Perri C, et al. **Acquired Dandy-Walker malformation and cerebellar hemorrhage: usefulness of serial MRI.** *Eur J Paediatr Neurol* 2016;20:188–91 CrossRef Medline
- Cho IY, Hwang SK, Kim SH. **Dandy-Walker malformation associated with neurocutaneous melanosis.** *J Korean Neurosurg Soc* 2011;50:475–77 CrossRef Medline
- Di Stasi M, Mankad K, Carney O, et al. **Congenital melanocytic naevus syndrome and Dandy-Walker malformation: a mistaken association: case report and literature review.** *Neuroradiology* 2023;65:1077–86 CrossRef Medline
- Haldipur P, Gillies GS, Janson OK, et al. ***FOXC1* dependent mesenchymal signalling drives embryonic cerebellar growth.** *Elife* 2014;3:e03962 CrossRef Medline
- Whitehead MT, Barkovich MJ, Sidpra J, et al. **Refining the neuroimaging definition of the Dandy-Walker phenotype.** *AJNR Am J Neuroradiol* 2022;43:1488–93 CrossRef Medline
- Stroustrup Smith A, Levine D. **Appearance of an interhemispheric cyst associated with agenesis of the corpus callosum.** *AJNR Am J Neuroradiol* 2004;25:1037–40 Medline
- Spennato P, Mirone G, Nastro A, et al. **Hydrocephalus in Dandy-Walker malformation.** *Childs Nerv Syst* 2011;27:1665–81 CrossRef Medline
- Watson C, Bartholomaeus C, Puelles L. **Time for radical changes in brainstem nomenclature-applying the lessons from developmental gene patterns.** *Front Neuroanat* 2019;13:10 CrossRef Medline
- Leto K, Arancillo M, Becker EB, et al. **Consensus paper: cerebellar development.** *Cerebellum* 2016;15:789–828 CrossRef Medline
- Kratochwil CF, Maheshwari U, Rijli FM. **The long journey of pontine nuclei neurons: from rhombic lip to cortico-ponto-cerebellar circuitry.** *Front Neural Circuits* 2017;11:33 CrossRef Medline
- D'Angelo E. **Physiology of the cerebellum.** *Handb Clin Neurol* 2018;154:85–108 CrossRef
- Jissendi-Tchofo P, Doherty D, McGillivray G, et al. **Pontine tegmental cap dysplasia: MR imaging and diffusion tensor imaging features of impaired axonal navigation.** *AJNR Am J Neuroradiol* 2009;30:113–19 CrossRef Medline
- Gafner M, Garel C, Leibovitz Z, et al. **Medullary tegmental cap dysplasia: fetal and postnatal presentations of a unique brainstem malformation.** *AJNR Am J Neuroradiol* 2023;44:334–40 CrossRef Medline
- Barkovich AJ, Millen KJ, Dobyns WB. **A developmental classification of malformations of the brainstem.** *Ann Neurol* 2007;62:625–39 CrossRef Medline

25. Fischer EG. **Dandy-Walker syndrome: an evaluation of surgical treatment.** *J Neurosurg* 1973;39:615–21 CrossRef Medline
26. Bindal AK, Storrs BB, McLone DG. **Management of the Dandy-Walker syndrome.** *Pediatr Neurosurg* 1990;16:163–69 CrossRef Medline
27. Haldipur P, Aldinger KA, Bernardo S, et al. **Spatiotemporal expansion of primary progenitor zones in the developing human cerebellum.** *Science* 2019;366:454–60 CrossRef Medline
28. Haldipur P, Bernardo S, Aldinger KA, et al. **Evidence of disrupted rhombic lip development in the pathogenesis of Dandy-Walker malformation.** *Acta Neuropathol* 2021;142:761–76 CrossRef Medline
29. Aldinger KA, Elsen GE, Prince VE, et al. **Model organisms inform the search for the genes and developmental pathology underlying malformations of the human hindbrain.** *Semin Pediatr Neurol* 2009;16:155–63 CrossRef Medline
30. Haldipur P, Dang D, Aldinger KA, et al. **Phenotypic outcomes in mouse and human *FOXC1* dependent Dandy-Walker cerebellar malformation suggest shared mechanisms.** *Elife* 2017;6:e20898 CrossRef Medline

Utility of Gadolinium-Based Contrast in Initial Evaluation of Seizures in Children Presenting Emergently

¹Denas Andrijauskis, ²Graham Woolf, Alexander Kuehne, ³Khalid Al-Dasuqi, ⁴Cicero T. Silva, ⁵Syedmehtdi Payabvash, and ⁶Ajay Malhotra



ABSTRACT

BACKGROUND AND PURPOSE: The frequency and utility of gadolinium in evaluation of acute pediatric seizure presentation is not well known. The purpose of this study was to assess the utility of gadolinium-based contrast agents in MR imaging performed for the evaluation of acute pediatric seizure presentation.

MATERIALS AND METHODS: We identified consecutive pediatric patients with new-onset seizures from October 1, 2016, to September 30, 2021, who presented to the emergency department and/or were admitted to the inpatient unit and had an MR imaging of the brain for the evaluation of seizures. The clinical and imaging data were recorded, including the patient's age and sex, the use of IV gadolinium, and the underlying cause of epilepsy when available.

RESULTS: A total of 1884 patients were identified for inclusion. Five hundred twenty-four (28%) patients had potential epileptogenic findings on brain MR imaging, while 1153 (61%) patients had studies with normal findings and 207 (11%) patients had nonspecific signal changes. Epileptogenic findings were subclassified as the following: neurodevelopmental lesions, 142 (27%); intracranial hemorrhage (traumatic or germinal matrix), 89 (17%); ischemic/hypoxic, 62 (12%); hippocampal sclerosis, 44 (8%); neoplastic, 38 (7%); immune/infectious, 20 (4%); phakomatoses, 19 (4%); vascular anomalies, 17 (3%); metabolic, 3 (<1%); and other, 90 (17%). Eight hundred seventy-four (46%) patients received IV gadolinium. Of those, only 48 (5%) cases were retrospectively deemed to have necessitated the use of IV gadolinium: Fifteen of 48 (31%) cases were subclassified as immune/infectious, while 33 (69%) were neoplastic. Of the 1010 patients with an initial noncontrast study, 15 (1.5%) required repeat MR imaging with IV contrast to further evaluate the findings.

CONCLUSIONS: Gadolinium contrast is of limited additive benefit in the imaging of patients with an acute onset of pediatric seizures in most instances.

ABBREVIATIONS: ACR = American College of Radiology; GBC = gadolinium-based contrast; SPR = Society for Pediatric Radiologists

It is estimated that as many as 1% of children will experience at least 1 afebrile seizure by adolescence, and population-based studies estimate the incidence of epilepsy in childhood at around 0.5–8 per 1000 person-years.^{1–5} The causes of epilepsy in the pediatric population are numerous and can be divided into 6 groups: genetic, structural, metabolic, immune, infectious, and unknown.^{6,7} MR imaging is the criterion standard for the detection of structural epileptogenic abnormalities, making it critical in helping to define the electroclinical syndrome and for identifying surgically amenable lesions.^{6,8} MR imaging has also been shown to be an important prognostic tool for predicting medical refractoriness.^{9–13}

While the utility of MR imaging in the diagnosis and management of new-onset seizures is well-established, the extra yield of adding gadolinium-based contrast (GBC) to the initial evaluation is not well-defined. Current American College of Radiology (ACR) guidelines state that MR imaging of the brain without and with contrast may be appropriate.¹³ In a recent Society for Pediatric Radiologists (SPR) survey, nearly 24% of respondents stated that they used GBC “always or usually” for the initial evaluation of seizures, and another 32% responded that they used it “sometimes.”¹⁴

Multiple studies have shown evidence of trace amounts of GBC being retained chronically in the patient's body following administration.^{15–17} Unlike in the acute setting, the chronic effects and safety profile of GBC are not yet known.¹⁸

The purpose of this study was to assess the use and utility of GBC in MR imaging performed for the evaluation of acute-seizure presentation in the pediatric population.

Received June 16, 2023; accepted after revision August 2.

From the Department of Radiology and Biomedical Imaging, Yale School of Medicine, New Haven, Connecticut.

Please address correspondence to Ajay Malhotra, MD, MMM, Department of Radiology and Biomedical Imaging, Yale School of Medicine, Box 208042, Tompkins East 2, New Haven, CT 06520-8042; e-mail: ajay.malhotra@yale.edu; @AjayMalhotraRad
<http://dx.doi.org/10.3174/ajnr.A7976>

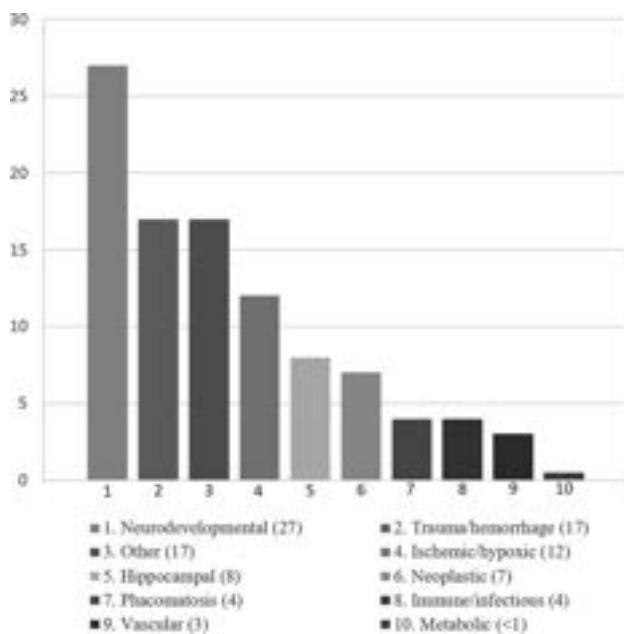


FIGURE. Frequencies of epileptogenic lesions stratified by etiologic categories in percentages (brackets).

MATERIALS AND METHODS

This Health Insurance Portability and Accountability Act–compliant retrospective study was approved by the institutional review board. Informed consent was waived for this retrospective analysis.

Retrospective analysis of the clinical and imaging data was performed of consecutive patients with new-onset seizures, younger than 18 years of age, from October 1, 2016, to September 30, 2021, who presented to the emergency department and/or were admitted to the inpatient unit of an academic, tertiary care center and had an MR imaging of the brain for the evaluation of seizures. Patients with a previously established epilepsy diagnosis were excluded from the study. Patients with a previously known neurologic diagnosis and new-onset seizures were further subclassified and accounted for separately. MR imaging at the institution is performed as ordered by the providers and is not monitored in real-time for the administration of contrast.

All MR imaging cases with positive findings were reviewed, and the indication for GBC was determined on the basis of whether the abnormality was visualized on the noncontrast portion of examination, the additive value of GBC such as enhancement or better delineation of the lesion in the postcontrast portion of the examination, or the recommendation of the interpreting radiologist to order a follow-up MR imaging examination with contrast for better delineation of the lesion found on the initial noncontrast study. The review of cases with positive findings was performed by 2 residents supervised by an attending physician with >12 years of experience.

RESULTS

A total of 1884 patients were identified for inclusion. The median age was 7 years (range, 3 days to 17.9 years) and 55% were male. Two hundred seventy of 1884 (14.3%) patients were younger than

1 year of age, including 102/1884 (5.4%) patients younger than 1 month of age. Of the 1884 patients, 524 (28%) had potential epileptogenic findings on brain MR imaging, while 1153 (61%) had normal findings and 207 (11%) had nonspecific signal changes of unknown clinical significance. One hundred forty-six of 270 (54%) infants and 41/102 (40%) neonates had normal findings on MR imaging.

The most common type of epileptogenic finding was neurodevelopmental lesions detected in 142 (27%) children. The most frequent abnormalities in this group consisted of focal cortical dysplasia in 47 (33%), followed by polymicrogyria in 24 (17%). Neurodevelopmental lesions were seen in 18/270 (6.7%) infants and in 6/102 (5.9%) neonates. The second most common category of epileptogenic findings was intracranial hemorrhage, accounting for 89 (17%) cases. Hemorrhage was seen in a higher proportion of infants (52/270; 19.2%) and neonates (25/102; 24.5%). Eighteen of 52 hemorrhages in infants were due to germinal matrix bleeds, and all these patients had bleeds detected on head ultrasound before the MR imaging obtained for seizure work-up. Birth trauma or nonaccidental trauma was the cause of hemorrhage in 19 infants, 7 of them neonates. Hemorrhagic infarct was seen in 2 patients, and the remaining 13 had nonspecific bleeds. Evidence of a hypoxic-ischemic injury was identified in 62 (12%) children with etiologically relevant lesions on brain MR imaging. Of note, this was the most common seizure etiology among the neonate population (27; 43%). Epileptogenic lesions associated with a hippocampal etiology were noted in 44 (8%) children. The next most common types of epileptogenic lesions were neoplastic (38; 7%) and immune/infectious (20; 4%). Phacomatosis as an etiology of seizures was detected in 19 (4%) children with 14/19 (74%) diagnosed with tuberous sclerosis; 4/19 (21%), with neurofibromatosis; and 1 patient, with Sturge-Weber syndrome. Vascular anomalies were identified in 17 (3%) patients. The least common epileptogenic category was metabolic (eg, metachromatic leukodystrophy, Gaucher disease), which was detected in 3 patients (<1%). The rest of the identified epileptogenic lesions were subclassified as “other” (17%). Sample findings in this category include structural asymmetry, periventricular leukomalacia, or large arachnoid cysts with mass effect.

The frequency of epileptogenic lesions stratified by etiologic categories is shown in the Figure. The frequencies of epileptogenic lesions were not statistically different between the male and female pediatric population ($P > .05$).

A total of 874 (46%) patients received IV gadolinium for the brain MR imaging. Of those, 48 (5%) cases were retrospectively deemed to have necessitated the use of IV gadolinium to establish the initial diagnosis, in which enhancement or better delineation of the lesion in the postcontrast portion of the examination was noted. There was no case in which the abnormality was seen only on the postcontrast images. Fifteen of 48 (31%) cases were subclassified as immune/infectious, while 33 (69%) were neoplastic. Five of these 15 patients with infections were infants (3 neonates). The immune/infectious cause was clinically suspected for 9 of 15 patients (eg, meningitis, encephalitis), while 6 patients did not have an established clinical differential for the seizure etiology, and it turned out to be MS, Lyme disease, or a nonspecific

demyelination process. Fifteen of 33 patients with a neoplastic seizure cause had an already known and previously confirmed CNS malignancy; therefore, contrast-enhanced MR imaging was ordered on the basis of prior medical history. Eighteen cases of neoplastic process were newly diagnosed and not suspected on clinical examination: pilocytic astrocytoma ($n = 3$), ganglioglioma ($n = 3$), choroid plexus papilloma ($n = 2$), pineal glioma ($n = 2$), anaplastic astrocytoma ($n = 1$), pleomorphic xanthoastrocytoma ($n = 1$), schwannoma ($n = 1$), tectal glioma ($n = 1$), and low-grade gliomas, not biopsied ($n = 4$).

The remainder of the patients did not show the additional value of GBC for initial diagnostic purposes.

Of 1010 patients with an initial noncontrast MR imaging study, 15 (1.5%) required repeat MR imaging with IV contrast for further evaluation as per the recommendation of interpreting radiologist. These cases were the following: indeterminate lesions with broad differential ($n = 4$), low-grade glioma, not biopsied ($n = 3$), encephalitis ($n = 1$), MS ($n = 1$), anaplastic astrocytoma ($n = 1$), choroid plexus papilloma ($n = 1$), Lyme disease ($n = 1$), pilocytic astrocytoma ($n = 1$), pineal glioma ($n = 1$), and ganglioglioma ($n = 1$).

DISCUSSION

Our results showed that 28% of patients with new-onset pediatric seizures had a potential epileptogenic focus on brain MR imaging. This finding is consistent with those in previous studies that have validated MR imaging as an important diagnostic and prognostic tool for the evaluation of new-onset seizures in the pediatric population.^{9,19,20}

Forty-six percent of our study cohort received IV gadolinium for the initial brain MR imaging study. This is consistent with the results of the 2017 SPR survey, in which 24.7% of respondents indicated that they “always or usually” administered GBCA agents for patients with seizures, while 32% indicated that they used it “sometimes.”¹⁴ The current ACR guidelines for patients with seizures state that MR imaging with GBC may be appropriate for the initial evaluation of new seizures in a patient without known trauma.¹³ While the ACR guidelines do report that contrast should be used when images without IV contrast are found to be insufficient or when there is clinical suspicion of a neoplastic or inflammatory condition, this recommendation is based on and cites a review article.⁸

Specifically, for children 1 month to 18 years of age with focal seizures, MR imaging with GBC was deemed “maybe appropriate, with disagreement” due to insufficient medical literature evidence.²¹ There is a lack of data regarding the diagnostic advantage of GBC-enhanced MR imaging. Our study specifically investigated the utility of MR imaging contrast in the evaluation of new-onset pediatric seizures and found limited benefit from GBC in most cases. In a recent retrospective study, Nelson et al¹⁹ found that in adults presenting with acute seizure, contrast was not necessary in all cases. Our results were concordant, and very few patients had potentially epileptogenic findings that required contrast for characterization. Forty-eight of 874 (5.5%) patients who received GBC had infection/inflammation or neoplastic etiology, and in 24 of these patients, there was previous clinical concern for infection or known CNS malignancy. Fifteen of 1010

(1.5%) patients with an initial noncontrast MR imaging underwent a repeat study with IV contrast to further evaluate as per the recommendation of the interpreting radiologist. Most patients with previously unsuspected or incompletely characterized lesions on noncontrast MR imaging had low-grade nonenhancing tumors.

The value of GBC for the detection of brain neoplasms and CNS infections is well-established in the literature.^{13,22,23} Thus, when there is clinical suspicion of tumor or an immune/infectious process, such as fever, suspected encephalitis, or a history of malignancy or primary CNS tumor, an initial evaluation with MR imaging of the brain with and without contrast is warranted. However, infectious processes and neoplasms account for only a subset of seizure etiologies and epileptogenic findings on MR imaging. In our study, these accounted for only 11% of all potentially epileptogenic findings and were identified in only 3% of patients overall.

The administration of GBC is not without risks. Along with known potential acute allergic and physiologic reactions, recent literature has increasingly demonstrated trace, chronic gadolinium deposition within tissue.¹⁵⁻¹⁷ The potential adverse effects of chronic gadolinium deposition within tissues are not yet well-understood.²⁴

With the potential long-term consequences of gadolinium deposition still unknown and our study showing that in the overwhelming majority of cases, GBC administration has little diagnostic benefit, it stands to reason that contrast administration should be restricted in the pediatric population. Restricting the use of gadolinium would also lead to a potential reduction in scan times and the need for sedation. Contrast may be reserved for specific clinical scenarios where there is clinical suspicion of neoplasm or infection or subsequent to an abnormality on noncontrast MR imaging warranting postcontrast imaging. Implementing this stepwise approach may be met with concerns over potential delays in diagnosis and the inconvenience of having to recall patients for subsequent imaging. However, at most, in our study, <1/40 patients had findings requiring contrast for further evaluation that would have required a callback, and most of these were nonurgent, low-grade lesions. Real-time monitoring of these patients for the use of contrast during scanning might also be beneficial.

Limitations of our study primarily revolve around the retrospective nature of the chart and imaging review used to collect our data for analysis. Contrast may have been useful in excluding meningitis as a potential cause in patients with suspected infection. However, there is little literature evidence that gadolinium use can exclude underlying infection. We also used data from a single tertiary care center, which does not account for possible regional or institutional variation in the use of GBC for first-time pediatric seizures.

CONCLUSIONS

Our study shows limited additive benefit of GBC in the imaging of pediatric patients with acute-onset seizures. These findings support reserving contrast use for specific clinical circumstances or when indicated by a finding on a noncontrast examination. More restrictive use of GBC in pediatric seizure imaging may be of

particular importance in the pediatric population because the potential chronic adverse effects of GBC are not yet well-understood.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

1. Fiest KM, Sauro KM, Wiebe S, et al. **Prevalence and incidence of epilepsy: a systematic review and meta-analysis of international studies.** *Neurology* 2017;88:296–303 CrossRef Medline
2. Hauser WA, Annegers JF, Kurland LT. **Prevalence of epilepsy in Rochester, Minnesota: 1940–1980.** *Epilepsia* 1991;32:429–45 CrossRef Medline
3. Oka E, Ohtsuka Y, Yoshinaga H, et al. **Prevalence of childhood epilepsy and distribution of epileptic syndromes: a population-based survey in Okayama, Japan.** *Epilepsia* 2006;47:626–30 CrossRef Medline
4. Aaberg KM, Gunnes N, Bakken IJ, et al. **Incidence and prevalence of childhood epilepsy: a nationwide cohort study.** *Pediatrics* 2017;139:e20163908 CrossRef Medline
5. Russ SA, Larson K, Halfon N. **A national profile of childhood epilepsy and seizure disorder.** *Pediatrics* 2012;129:256–64 CrossRef Medline
6. Scheffer IE, Berkovic S, Capovilla G, et al. **ILAE classification of the epilepsies: position paper of the ILAE Commission for Classification and Terminology.** *Epilepsia* 2017;58:512–21 CrossRef Medline
7. Beghi E. **The epidemiology of epilepsy.** *Neuroepidemiology* 2020;54:185–91 CrossRef Medline
8. Cendes F, Theodore WH, Brinkmann BH, et al. **Neuroimaging of epilepsy.** *Handb Clin Neurol* 2016;136:985–1014 CrossRef Medline
9. Berg AT, Mathern GW, Bronen RA, et al. **Frequency, prognosis and surgical treatment of structural abnormalities seen with magnetic resonance imaging in childhood epilepsy.** *Brain* 2009;132:2785–97 CrossRef Medline
10. Bernasconi A, Cendes F, Theodore WH, et al. **Recommendations for the use of structural magnetic resonance imaging in the care of patients with epilepsy: a consensus report from the International League Against Epilepsy Neuroimaging Task Force.** *Epilepsia* 2019;60:2143–44 CrossRef Medline
11. Tellez-Zenteno JF, Hernandez Ronquillo L, Moien-Afshari F, et al. **Surgical outcomes in lesional and non-lesional epilepsy: a systematic review and meta-analysis.** *Epilepsy Res* 2010;89:310–18 CrossRef Medline
12. Semah F, Picot MC, Adam C, et al. **Is the underlying cause of epilepsy a major prognostic factor for recurrence?** *Neurology* 1998;51:1256–62 CrossRef Medline
13. Lee RK, Burns J, Ajam AA, et al; Expert Panel on Neurological Imaging. **ACR Appropriateness Criteria(R) Seizures and Epilepsy.** *J Am Coll Radiol* 2020;17:S293–304 CrossRef Medline
14. Blumfield E, Moore MM, Drake MK, et al. **Survey of gadolinium-based contrast agent utilization among the members of the Society for Pediatric Radiology: a Quality and Safety Committee report.** *Pediatr Radiol* 2017;47:665–73 CrossRef Medline
15. Lersy F, Boulouis G, Clément O, et al. **Consensus Guidelines of the French Society of Neuroradiology (SFNR) on the use of gadolinium-based contrast agents (GBCAs) and related MRI protocols in neuroradiology.** *J Neuroradiol* 2020;47:441–49 CrossRef Medline
16. McDonald RJ, McDonald JS, Kallmes DF, et al. **Intracranial gadolinium deposition after contrast-enhanced MR imaging.** *Radiology* 2015;275:772–82 CrossRef Medline
17. McDonald JS, McDonald RJ, Jentoft ME, et al. **Intracranial gadolinium deposition following gadodiamide-enhanced magnetic resonance imaging in pediatric patients: a case-control study.** *JAMA Pediatr* 2017;171:705–07 CrossRef Medline
18. McDonald RJ, Levine D, Weinreb J, et al. **Gadolinium retention: a research roadmap from the 2018 NIH/ACR/RSNA Workshop on Gadolinium Chelates.** *Radiology* 2018;289:517–34 CrossRef Medline
19. Nelson KA, Thaker AA, Callen AL, et al. **New-onset seizures in adults: low diagnostic yield of gadolinium contrast in initial brain MRI evaluation.** *J Neuroimaging* 2021;31:874–78 CrossRef Medline
20. Hourani R, Nasreddine W, Dirani M, et al. **When should a brain MRI be performed in children with new-onset seizures? Results of a large prospective trial.** *AJNR Am J Neuroradiol* 2021;42:1695–701 CrossRef Medline
21. Trofimova A, Milla SS, Ryan ME, et al; Expert Panel on Pediatric Imaging. **ACR Appropriateness Criteria® Seizures-Child.** *J Am Coll Radiol* 2021;18:S199–211 CrossRef Medline
22. Villanueva-Meyer JE, Mabray MC, Cha S. **Current clinical brain tumor imaging.** *Neurosurgery* 2017;81:397–415 CrossRef Medline
23. Kastrup O, Wanke I, Maschke M. **Neuroimaging of infections.** *NeuroRx* 2005;2:324–32 CrossRef Medline
24. Noda SM, Oztek MA, Stanescu AL, et al. **Gadolinium retention: should pediatric radiologists be concerned, and how to frame conversations with families.** *Pediatr Radiol* 2022;52:345–53 CrossRef Medline

Dual-Layer Detector Head CT to Maintain Image Quality While Reducing the Radiation Dose in Pediatric Patients

Zhengwu Tan, Lan Zhang, Xiaojie Sun, Ming Yang, Joyman Makamure, Hongying Wu, and Jing Wang



ABSTRACT

BACKGROUND AND PURPOSE: Radiation exposure in the CT diagnostic imaging process is a conspicuous concern in pediatric patients. This study aimed to evaluate whether 60-keV virtual monoenergetic images of the pediatric cranium in dual-layer CT can reduce the radiation dose while maintaining image quality compared with conventional images.

MATERIALS AND METHODS: One hundred six unenhanced pediatric head scans acquired by dual-layer CT were retrospectively assessed. The patients were assigned to 2 groups of 53 and scanned with 250 and 180 mAs, respectively. Dose-length product values were retrieved, and noise, SNR, and contrast-to-noise ratio were calculated for each case. Two radiologists blinded to the reconstruction technique used evaluated image quality on a 5-point Likert scale. Statistical assessment was performed with ANOVA and the Wilcoxon test, adjusted for multiple comparisons.

RESULTS: Mean dose-length product values were 717.47 (SD, 41.52) mGy×cm and 520.74 (SD, 42) mGy×cm for the 250- and 180-mAs groups, respectively. Irrespective of the radiation dose, noise was significantly lower, SNR and contrast-to-noise ratio were significantly higher, and subjective analysis revealed significant superiority of 60-keV virtual monoenergetic images compared with conventional images (all $P < .001$). SNR, contrast-to-noise ratio, and subjective evaluation in 60-keV virtual monoenergetic images were not significantly different between the 2 scan groups ($P > .05$). Radiation dose parameters were significantly lower in the 180-mAs group compared with the 250-mAs group ($P < .001$).

CONCLUSIONS: Dual-layer CT 60-keV virtual monoenergetic images allowed a radiation dose reduction of 28% without image-quality loss in pediatric cranial CT.

ABBREVIATIONS: CNR = contrast-to-noise ratio; CTDI_{vol} = volume CT dose index; DECT = dual-energy CT; DLCT = dual-layer CT; DLP = dose-length product; GWMA = assessment of GM-WM differentiation; PFAA = assessment of artifacts in posterior fossa; SSA = assessment of the subcalvarial space; SAI = subcalvarial artifact index; VMI = virtual monoenergetic image

Unenhanced CT of the head is the standard technique for detecting intracranial pathologies, including trauma or intracranial hemorrhage, in children in the emergency department.¹ MR imaging can accurately detect traumatic complications but often requires sedation in children because of the examination

length and motion sensitivity, which limits rapid assessment; consequently, CT is considered the first-line imaging technique for suspected intracranial injury because of the short examination duration and high sensitivity for acute hemorrhage.² Because CT use in children has risen dramatically, radiation exposure from CT scanning is of an increasing concern; indeed, pediatric patients may receive high radiation doses and are more susceptible to radiation-related malignancies than adults because of organ masses, volumes, and morphology that are very different in children compared with adults.³⁻⁷ Therefore, it is important to optimize the imaging parameters to be consistent with the patient's size. From the patient's perspective, the benefits of a medically required CT scan far exceed the small increase in radiation-induced cancer risk,⁸ but it is always beneficial to reduce radiation exposure from CT scanning in children. Several strategies, eg, reducing the tube current (milliampere) and voltage (kilovolt) as well as using adaptive statistical iterative reconstruction to maintain image quality

Received April 7, 2023; accepted after revision August 2.

From the Department of Radiology (Z.T., L.Z., X.S., M.Y., J.M., H.W., J.W.), Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China; and Hubei Province Key Laboratory of Molecular Imaging (Z.T., L.Z., X.S., M.Y., J.M., H.W., J.W.), Wuhan, Hubei, China.

This research was supported by the National Natural Science Foundation of Hubei Province, China (No. 2021CFB447).

Please address correspondence to Jing Wang, MD, Department of Radiology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, No. 1277, Jiefang Ave, Wuhan, China; e-mail: jjwinflower@126.com

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

Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A7999>

Image-acquisition parameters, workstation, and viewing modes used for the evaluation

Scanner Model	250 mAs	180 mAs
Scanner sequence	Unenhanced image	Unenhanced image
KVp/tube current	120/250	120/180
Pitch/rotation	0.390/0.5	0.390/0.5
Detector configuration	64 × 0.625	64 × 0.625
Kernel	B	B
Section thickness	1 mm	1 mm
Reconstructed thickness	5 mm	5 mm
Iterative reconstruction algorithm/level	Spectral/level 3	Spectral/level 3
Workstation	IntelliSpace 9.0 ^a	IntelliSpace 9.0 ^a
Viewing mode/VMI	MonoE	MonoE
CT dose		
DLP (mGy × cm)	717.47 (SD, 41.52)	520.74 (SD, 42.00)
DLP, 0–6 yr	711.45 (SD, 47.83)	503.76 (SD, 54.00)
DLP, 7–12 yr	724.75 (SD, 31.80)	531.88 (SD, 27.47)
CTDI _{vol} (mGy)	36	26

Note:—MonoE indicates monoenergetic.

^a Philips Healthcare.

while reducing radiation, are available.^{5,9–11} However, image quality can be reduced by low signal intensity and low contrast as well as artifacts caused mainly by movement, image superimposition of different structures, and other factors.

Since the clinical introduction of dual-energy CT (DECT) with rapid kilovolt switching and dual x-ray sources that acquire 2 attenuation data sets by separating energies at the tube level, virtual monoenergetic images (VMIs) have improved soft-tissue contrast and reduced beam-hardening artifacts.^{12,13} From a neuroimaging perspective, DECT has improved image quality and lesion characterization while reducing radiation exposure in pediatric patients.^{14,15} A previous study of pediatric patients reported that 60-keV virtual monoenergetic imaging maximized image quality for the brain parenchyma.¹⁶ A type of DECT using 2 separate orthogonally oriented x-ray spectrum source detectors generated VMIs on the basis of high- and low-energy data sets (typically at 80 kV[peak] and 140 kVp) from the same anatomic region.¹⁷ However, in most patients, these dual-energy systems are operated as conventional scanners because dual-energy scanning affects scanner performance, either by neutralizing or increasing the radiation dose.^{18–20}

Recently, a detector-based approach, referred to as spectral detector CT, was introduced with a shorter scanning time, faster postprocessing, and a lower radiation dose.^{18,21} Dual-layer CT (DLCT) uses a single x-ray source and a detector comprising 2 scintillation layers to perform spectral separation at the detector level using the different absorption properties of the detector layers for high-energy and low-energy photons. The upper layer of the detector is yttrium-based and records lower-energy photons, while the lower layer is gadolinium oxysulfide-based and records higher-energy photons.¹³ Another benefit of this scanning method is that the dual-energy data set is retrieved from the conventional scan and may be used for retrospective analysis. Previous data showed that the quality of GM-WM matter contrast images can be optimized,²² and VMIs from spectral detector CT enabled a 19% reduction in the radiation dose in adult cranial CT while maintaining superior image quality over conventional images obtained with full-dose acquisition.²³ Despite

these encouraging results for adult head imaging, this technique has not been validated in the pediatric population so far.

Therefore, this study aimed to compare 60-keV VMIs with conventional images obtained with different acquisition protocols to evaluate whether a reduced radiation dose can be achieved in 60-keV VMIs without compromising image quality.

MATERIALS AND METHODS

Patient Population

This retrospective, single-center study was approved by the institutional ethics review board (Union Hospital, Tongji Medical College, Huazhong University of Science and Technology), who

waived the requirement for written informed consent.

A total of 148 CT brain scans were consecutively performed in pediatric patients aged 12 years or younger between December 2020 and August 2022 with a standard DLCT protocol in our clinical center. Fourteen patients were excluded from the analysis for severe encephalomalacia, which makes distinguishing GM and WM difficult ($n = 8$) and artifacts that reduce image quality (eg, metal implants) ($n = 6$). Then, 28 patients were excluded because of incomplete data ($n = 28$). Finally, 106 patients were analyzed in this study and assigned to 2 groups according to tube current (250 and 180 mAs); they were further divided into 2 subgroups according to age (6 years or younger and >older than 6 years) (Online Supplemental Data).

Spectral CT Examination

CT was performed with a dual-layer detector CT scanner (IQon; Philips Healthcare). In all selected patients, 53 scans were obtained with a tube current–time product of 250 mAs, and the other 53 scans, with 180 mAs. All other scan parameters were identical (Table). Conventional images were reconstructed with a hybrid algorithm (iDose4, Filter UB; Philips Healthcare iterative reconstruction algorithm). Meanwhile, 60-keV VMIs were reconstructed with a dedicated spectral image-reconstruction algorithm (Spectral, Filter UB; Philips Healthcare). Denoising was set at a minimum level (level 3 of 7) in both reconstruction approaches. All images were reconstructed using a section thickness of 1 mm and a section increment of 1 mm. The dose-length product (DLP) and volume CT dose index (CTDI_{vol}) values were recorded on the radiation dose report.

Objective Image Analysis

The 60-keV VMIs and conventional images were reviewed by an independent radiologist (reader 1) with 2 years of experience in head CT, and objective image analysis was performed with a proprietary vendor console (Spectral Diagnostics Suite; Philips Healthcare). The reviewers had prior training on ROI placement and data acquisition using an independent data set by a senior neuroradiologist. Axial images were analyzed on the brain window (width/center: 70/35 HU). For each VMI and conventional

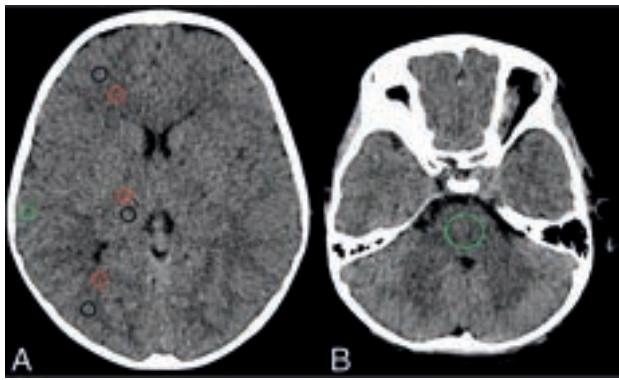


FIG 1. Conventional images. Placement of ROIs in cortical GM and the thalamic parenchyma (ROIs in the *black circle*), in juxtacortical WM and the posterior limb of the internal capsule (ROIs in *red circles*), as well as close to the calvaria and in the medulla oblongata (ROIs in *green circles*) on the axial plane, showing the basal ganglia (A) and the posterior fossa (B).

image data set, reader 1 placed 8 ROIs on a mid-basal ganglia section in the following anatomic locations: 1) the GM of the frontal and parietal lobes, 2) the juxtacortical WM of the frontal and parietal lobes, 3) the thalamic parenchyma, 4) the posterior limb of the internal capsule, and 5) the area near the calvaria (Fig 1A). An additional ROI was placed in the medulla oblongata (6) on the axial plane showing the petrous parts of the temporal bones, as well as the posterior fossa (Fig 1B). The locations and sizes of ROIs were identical for conventional images and VMIs. As previously described, an ROI of 25 mm² was used for all regions of the supratentorial brain versus 200 mm² for the medulla oblongata region of the posterior fossa. Patient-to-patient size-variation adjustment was performed to avoid volume averaging with the adjacent tissues. The respective CT attenuation, with an SD, was obtained from each ROI. Hounsfield unit values and image noise (SD) were averaged for comparing VMI and conventional images.

SNRs for GM and WM as well as the contrast-to-noise ratio (CNR) for GM-WM differentiation were calculated. The SNR of GM and WM for each ROI was determined as the mean CT attenuation divided by the SD. The difference was assessed as the CNR between the ROI measurements of adjacent GM and WM, as follows: $CNR = \frac{\text{Difference of Mean CT Number between GM and WM}}{\text{Square Root of the Sum of their Variances}}$.¹⁴ The posterior fossa artifact index (the SD of the ROI in the posterior fossa) reflected the disturbance of attenuation values from beam-hardening and streak artifacts. The subcalvarial artifact index (SAI, the SD of the ROI close to the calvaria) served as a reference for the beam-hardening artifacts of the skull.

Subjective Image Analysis

The subjective analysis of VMI and conventional images was performed independently by 2 radiologists, readers 2 and 3, with 3 and 5 years of experience in interpreting head CT scans, respectively, who were blinded to the applied reconstruction and scan techniques. Monoenergetic spectral images at 40, 50, 60, 70, 80, 90, and 100 keV were obtained during the reviewing process for comparing them with conventional images. The subjective

overall image quality was rated with a 5-point Likert-type scale (1 = nondiagnostic, 2 = limited, 3 = moderate, 4 = good, and 5 = excellent). Furthermore, the assessment of GM-WM matter differentiation (GWMA) was defined as the ability to distinguish GM from WM (ranging from 1 [no definite differentiation possible] to 5 [excellent differentiation]), subcalvarial space (SSA), and posterior fossa artifacts (PFAA) (ranging from 1 [no distinction of the subcalvarial space/absence of artifacts] to 5 [distinguished subcalvarial space as severe and diagnostically unacceptable]).

Statistical Analysis

SPSS Version 19.0 (IBM), MedCalc Version 19.0.4 (MedCalc Software), and GraphPad Prism Version 7.0 (GraphPad Software) were used for data analysis. Continuous variables were expressed as mean (SD), and categorical variables, as median and range. According to the distribution pattern of continuous variables, the *t* test or Mann-Whitney *U* test was performed to compare objective evaluation indices (mean CT value, SD, SNR, CNR, posterior fossa artifact index, and SAI) between 60-keV VMIs and conventional images. The Wilcoxon test was used to compare Likert scores. Interrater agreement for subjective analysis was assessed using κ coefficients: ≤ 0.20 , poor; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; 0.81–1.00, near-perfect. $P < .05$ was considered statistically significant.

RESULTS

The mean patient age was 6.80 years (SD, 3.20 years), and female and male children were 42 (39.62%) and 64 (60.38%), respectively, for all age groups. The patients were 3.92 years (SD, 1.78) years and 9.38 years (SD, 1.57) years, of whom 19 (38.00%) and 23 (41.07%) were female children aged 6 years or younger and older than 6 years, respectively, and 31 (62.00%) and 33 (58.93%) were male children aged 6 years or younger and older than 6 years, respectively.

CTDI_{vol} values were 36 and 26 mGy, and DLPs were 717.47 (SD, 41.52) mGy×cm and 520.74 (SD, 42.00) mGy×cm in the 250- and 180-mAs groups for all age groups, respectively (Table). The CTDI and DLP were significantly reduced by 28% and 28%, respectively, in the 180- and 250-mAs groups. Considering the age group (6 years or younger versus older than 6 years of age), DLP values did not differ significantly within the 250-mAs group (711.45 [SD, 47.83] mGy×cm versus 724.75 [SD, 31.8] mGy×cm; $P = .324$). However, DLP values within the 180-mAs group significantly increased with age (503.76 [SD, 54.00] mGy×cm versus 531.88 [SD, 27.47] mGy×cm; $P = .001$). Additionally, no significant differences in ROIs were found between the groups (ROIs for supratentorial brain in the 250- and 180-mAs groups were 25.09 [SD, 0.60] mm² and 24.90 [SD, 0.43] mm², respectively ($P = .064$). ROIs for the infratentorial brain were 199.88 (SD, 2.31) mm² and 200.4 (SD, 2.82) mm², respectively ($P = .30$).

Objective Analysis

In the intragroup comparison of the 180- or 250-mAs groups, attenuation in the GM was significantly higher and attenuation in the WM was significantly lower in the 60-keV VMI compared with conventional images ($P < .001$) for all age groups, as well as

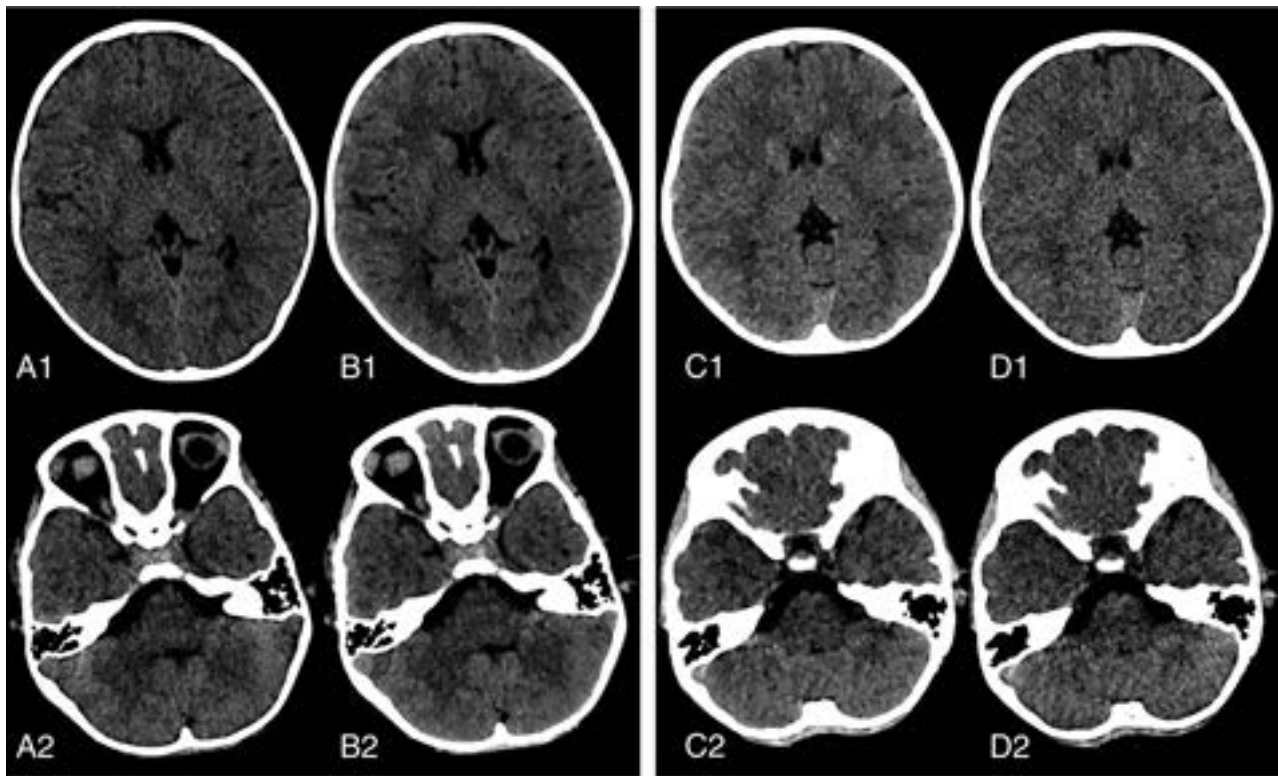


FIG 2. Representative images showing improved image quality in the supratentorial and infratentorial brain parenchyma for conventional images and 60-keV VMIs acquired using 250 and 180 mAs. Axial reconstruction conventional image (A) and 60-keV VMI (B) for nonenhanced brain CT scans acquired with 250 mAs in a 6-year-old girl with neuroblastoma in the retroperitoneal area. Axial reconstruction conventional image (D) and 60-keV VMI (C) for a nonenhanced brain CT scan acquired with 180 mAs in a 6-year-old girl with exotropia in the eye. GWMA (B1, C1), SSA, and PFAA (B2, C2) and overall image quality were better for 60-keV VMIs compared with conventional images (A1, D1, A2, and D2), respectively. Subjective image-quality indices were similar in 60-keV VMIs between the 250-mAs (B1, B2) and 180-mAs (C1, C2) groups. Window settings were kept identical for better comparability (level, 35; width, 70).

for the 6 years and younger and older than 6-year groups (Online Supplemental Data). Compared with conventional images, GM noise, WM noise, and PFAA were significantly lower, and GM SNR, WM SNR, and CNR were significantly higher in 60-keV VMIs ($P < .002$), but SAI values were not significantly different for all age groups, as well as for the 6 years and younger or older than 6-year groups ($P > .05$).

No significant differences were found in GW SNR, SAI, and CNR for 60-keV VMIs in the 250-mAs group compared with the 180-mAs group for all age groups ($P = .064$, $P = .308$, and $P = .150$, respectively). GW noise, WM noise, WM SNR, SAI, and CNR were also not significantly different between the 180- and 250-mAs groups for both the 6 year and younger and older than 6-year groups ($P > .05$). GM noise, WM noise, and PFAA were significantly higher and WM SNR was significantly lower in the 180-mAs group compared with the 250-mAs group for all age groups ($P = .019$, $P < .001$, $P = .001$, and $P < .001$, respectively). PFAA values were also significantly higher for both the 6 year and younger and older than 6-year groups ($P = .024$ and $P = .034$, respectively).

Subjective Analysis

In the intragroup comparison of both the 180- and 250-mAs groups, the assessment of GM-WM differentiation, SSA, overall

image quality, and PFAA had superior Likert scores for 60-keV VMIs compared with conventional images for all age groups, as well as for the 6 year and younger and older than 6-year groups ($P = .013$ to $< .001$; Online Supplemental Data and Fig 2). Specific to 60-keV VMIs, besides PFAA that was significantly different for the older than 6 year group ($P = .001$), all indexes were not significantly different in the 250-mAs group compared with the 180-mAs group for all age groups, as well as for the 6 year and younger and older than 6-year groups (GM-WM differentiation, SSA, the overall image quality, and PFAA; $P = .068$ to $.686$, $P = .063$ to $.283$, $P = .082$ to $.465$, and $P = .055$ to $.912$, respectively) (Fig 2).

Interobserver Agreement

Regarding interobserver agreement, κ values were 0.707 for conventional images and 0.816 for 60-keV VMIs with GWMA, 0.662 for conventional images and 0.718 for 60-keV VMIs with SSA, 0.724 for conventional images and 0.769 for 60-keV VMIs with PFAA assessment, and 0.817 for conventional images and 0.768 for 60-keV VMIs for assessing overall diagnostic quality.

DISCUSSION

In the analysis of unenhanced brain CT images in children, by comparing image quality between 60-keV VMIs and conventional images from spectral detector CT data sets acquired with

different radiation doses, image quality for 60-keV VMIs by DLCT in the 180-mAs group was improved and the radiation dose was reduced by 28% compared with the 250-mAs group based on DLP in cranial CT.

In 60-keV VMIs, significantly higher attenuation in the GM and significantly lower attenuation in the WM were found for the same radiation dose. On the other hand, image noise was significantly lower in 60-keV VMIs compared with conventional images, irrespective of the radiation dose. This noise reduction resulted in significantly higher CNR for GM-WM differentiation on 60-keV VMIs. Hence, objective image-quality parameters were significantly better in 60-keV VMIs than in conventional images, irrespective of the DLP. Accordingly, the subjective image analysis showed a superiority for 60-keV VMIs over conventional images regarding diagnostic assessment.

Because unenhanced head CT is the imaging method of choice in patients with neurologic deficits and neurocranial traumatic lesions, there is a need for improving image quality. Previous studies found better image quality for VMIs from unenhanced head CTs due to increased GM-WM differentiation and reduced beam-hardening artifacts of the skull.²²⁻²⁴

A previous study using rapid-switching DECT in children reported that monoenergetic reconstruction using 60-keV resulted in better image quality because of better GM-WM differentiation and reduced artifacts in the posterior fossa.¹⁶ Because this study was conducted in pediatric patients, the 60-keV VMI was selected for comparative assessment with conventional images. The outcomes supported previous studies,^{16,17} with a superior quality obtained from 60-keV VMIs compared with conventional images. Maximal SNR and CNR and minimal noise were observed on VMIs, irrespective of radiation dose.

To the best of our knowledge, this was the first study to analyze dose reduction in children examined with 60-keV virtual monoenergetic imaging by DLCT, in which a radiation dose reduction of 28% was achieved for the CT scanning process with a tube current–time product of 180 mAs while maintaining comparable image quality versus the 250-mAs group. If one compared GW SNR, SAI, and CNR with 60-keV VMIs, the subjective analysis found no significant differences between the 250- and 180-mAs groups for all age groups; in addition, GW noise, WM noise, WM SNR, SAI, CNR, and subjective analysis were not significantly different for both the 6-year and younger and older than 6-year groups. The technique was used in adults by Reimer et al²³ to demonstrate a 19% reduction in radiation dose from nonenhanced head CT using DLCT by reducing the tube current without a loss of image quality. The CT scanning protocol using 180 mAs was initially selected in this study because it is considered more suitable for radiation dose reduction than CT scanning protocols using 100- and 280-mAs, respectively, as previously reported,^{25,26} and also because approximately 20%–30% of the reduction in the radiation dose would not result in image quality not applicable for diagnosis.¹⁴

A linear relationship was demonstrated between tube current–time product and radiation dose. A decrease in tube current–time product increases image noise.²⁷ In a recent study of unenhanced head CT by Reimer et al,²³ image noise slightly increased from 55 to 49.8 and 44.7 mGy, respectively, showing a

significant difference between the 55- and 44.7-mGy protocols with the same reconstruction technique. Similarly, this study also revealed statistically significant increases in GW noise, WM noise, and PFAA in the 180-mAs group compared with the 250-mAs group for identical 60-keV VMIs for 0–12 years of age. However, no significant differences were observed for GW noise and WM noise in subgroup analysis based on patient age and for subjectively-rated image graininess (noise) for all age groups and age subgroups. These findings indicate that the image quality is acceptable for diagnosis, and the tube current may be further reduced without affecting image noise across age groups. Furthermore, SAI, WM SNR, and CNR were not significantly different, and there were no significant differences in overall image quality scores between the 250- and 180-mAs groups for all age groups and age subgroups.

Therefore, in this study, the degradation of image quality by such artifacts was unlikely to be significant, and overall image quality scores were similar between the 2 scan types and age subgroups. However, there was no statistically significant reduction in SAI, which indicated that our reduction of tube current to 180 mAs did not affect the objective image quality in SAI. The radiation dose for a CT scan should be carefully selected to provide optimal diagnostic image quality with the lowest possible radiation dose. Therefore, changes in the tube current–time product should be considered after consulting a radiologist, a CT technologist, and a medical physicist. In this study, radiation dose reduction was achieved by reducing the tube current–time product to 180 mAs without significantly altering the intrinsic quality of the image. Different denoising levels may impact image quality, and a medium denoising level was selected.²³ The effects of different denoising levels on image quality and whether the radiation dose can be reduced will be further explored.

As DLCT scanning parameters in this study, a tube voltage of 120 kV, a current of 250 mAs, a DLP of 717.47 (SD, 41.52) mGy×cm, and a CTDI_{vol} of 36 mGy were used for conventional CT to view the pediatric brain, which were similar to or slightly lower than diagnostic reference levels for the CT radiation dose in China (804 mGy for the DLP and 39 mGy×cm for CTDI_{vol})²⁸ and other countries.²⁹⁻³⁴ The CTDI_{vol} is the dose of the standard American College of Radiology head phantom according to the reference phantom selection in pediatric CT;^{19,35} the DLP is the actual radiation dose received by the patient. There was no significant difference in DLP between the 6-year and younger and older than 6-year subgroups in the 250-mAs group, while there was a significant difference in DLP between the 6-year and younger and older than 6-year subgroups in the 180-mAs group. There was a trend of DLP increase with increasing age, which may be related to head circumferences at different ages in children and head scan lengths determined by the technologist; this trend had no statistical significance between the 2 groups, with an overlap in mean DLP as observed by the SD of DLP in the 180-mAs group.³⁶ After one reduces the tube current for head CT examination, protocols can be specifically developed for various age groups because skull ossification is age-dependent.⁵ Moreover, it was reported that CT using DLCT with spectral data does not increase the radiation dose compared with CT using DLCT without spectral data.¹⁸

There were several limitations in this study. First, this was a retrospective study performed in a single institution. The sample size was relatively small because radiation dose reduction is not performed routinely. For ethical reasons, we could not use 2 distinct CT protocols for each patient for within-subject analysis. Second, a greater reduction of the radiation dose was not evaluated though the current data suggested that this was achievable. We compared radiation doses on the basis of the DLP alone because there is no established method to normalize the radiation dose to the size of the pediatric head (unlike the size-specific dose estimates for body CT), but grouping by age was performed in this study. Third, the diagnostic value of radiation dose reduction for different brain pathologies was not evaluated because this study aimed to evaluate the possibility of radiation dose reduction while preserving overall image quality on nonenhanced brain CT comparing the 250- and 180-mAs groups. However, the evaluation of diagnostic certainty and the accuracy in pathologies is beyond the scope of this study. Finally, although the current qualitative analysis was performed in a blinded manner, an experienced reader would be likely to detect differences between conventional images and 60-keV VMIs due to differences in image texture.

CONCLUSIONS

In this study, GM noise and WM noise, WM CNR, SAI, and CNR for supratentorial image quality were not different in objective and subjective evaluations between the 180- and 250-mAs groups, and noise in the posterior fossa differed in the objective evaluation by reduced tube current, but image quality is acceptable in the subjective evaluation, both for the 6 years and younger and the older than 6-year groups. On the basis of noise SNR, SAI, CNR, and PFAA in image analysis and observer agreement assessment, we found a dose reduction of 28% while maintaining superior image quality with a scan protocol of 180 mAs compared with 250 mAs on 60-keV VMIs from spectral detector CT in the pediatric head. These data further suggest that an even greater dose reduction is potentially achievable.






Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

- Hemphill JC, Greenberg SM, Anderson CS, et al; Council on Clinical Cardiology. **Guidelines for the management of spontaneous intracerebral hemorrhage: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association.** *Stroke* 2015;46:2032–60 CrossRef Medline
- Ryan ME, Pruthi S, Desai NK, et al; Expert Panel on Pediatric Imaging. **ACR Appropriateness Criteria® Head Trauma-Child.** *J Am Coll Radiol* 2020;17:S125–37 CrossRef Medline
- Meulepas JM, Ronckers CM, Smets AM, et al. **Radiation exposure from pediatric CT scans and subsequent cancer risk in the Netherlands.** *J Natl Cancer Inst* 2019;111:256–63 CrossRef Medline
- Nelson TR. **Practical strategies to reduce pediatric CT radiation dose.** *J Am Coll Radiol* 2014;11:292–99 CrossRef Medline
- Ngo AV, Winant AJ, Lee EY, et al. **Strategies for reducing radiation dose in CT for pediatric patients: how we do it.** *Semin Roentgenol* 2018;53:124–31 CrossRef Medline
- Pearce MS, Salotti JA, Little MP, et al. **Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study.** *Lancet* 2012;380:499–505 CrossRef Medline
- Albert GW, Glasier CM. **Strategies for computed tomography radiation dose reduction in pediatric neuroimaging.** *Neurosurgery* 2015;77:228–32; discussion 232 CrossRef Medline
- Miglioretti DL, Johnson E, Williams A, et al. **The use of computed tomography in pediatrics and the associated radiation exposure and estimated cancer risk.** *JAMA Pediatr* 2013;167:700–07 CrossRef Medline
- McKnight CD, Watcharotone K, Ibrahim M, et al. **Adaptive statistical iterative reconstruction: reducing dose while preserving image quality in the pediatric head CT examination.** *Pediatr Radiol* 2014;44:997–1003 CrossRef Medline
- Park JE, Choi YH, Cheon JE, et al. **Image quality and radiation dose of brain computed tomography in children: effects of decreasing tube voltage from 120 kVp to 80 kVp.** *Pediatr Radiol* 2017;47:710–17 CrossRef Medline
- Nagayama Y, Nakaura T, Tsuji A, et al. **Radiation dose reduction using 100-kVp and a sinogram-affirmed iterative reconstruction algorithm in adolescent head CT: impact on grey-white matter contrast and image noise.** *Eur Radiol* 2017;27:2717–25 CrossRef Medline
- Rapp JB, Biko DM, Barrera CA, et al. **Current and future applications of thoracic dual-energy CT in children: pearls and pitfalls of technique and interpretation.** *Semin Ultrasound CT MR* 2020;41:433–41 CrossRef Medline
- Duan X, Ananthakrishnan L, Guild JB, et al. **Radiation doses and image quality of abdominal CT scans at different patient sizes using spectral detector CT scanner: a phantom and clinical study.** *Abdom Radiol (NY)* 2020;45:3361–68 CrossRef Medline
- Weinman JP, Mirsky DM, Jensen AM, et al. **Dual energy head CT to maintain image quality while reducing dose in pediatric patients.** *Clin Imaging* 2019;55:83–88 CrossRef Medline
- Ginat DT, Mayich M, Daftari-Besheli L, et al. **Clinical applications of dual-energy CT in head and neck imaging.** *Eur Arch Otorhinolaryngol* 2016;273:547–53 CrossRef Medline
- Park J, Choi YH, Cheon JE, et al. **Advanced virtual monochromatic reconstruction of dual-energy unenhanced brain computed tomography in children: comparison of image quality against standard mono-energetic images and conventional polychromatic computed tomography.** *Pediatr Radiol* 2017;47:1648–58 CrossRef Medline
- Gottmukkala RV, Kalra MK, Tabari A, et al. **Advanced CT techniques for decreasing radiation dose, reducing sedation requirements, and optimizing image quality in children.** *Radiographics* 2019;39:709–26 CrossRef Medline
- Ommen F, Jong H, Dankbaar JW, et al. **Dose of CT protocols acquired in clinical routine using a dual-layer detector CT scanner: a preliminary report.** *Eur J Radiol* 2019;112:65–71 CrossRef Medline
- Zhu X, McCullough WP, Mecca P, et al. **Dual-energy compared to single-energy CT in pediatric imaging: a phantom study for DECT clinical guidance.** *Pediatr Radiol* 2016;46:1671–79 CrossRef Medline
- Schick D, Prapat J. **Radiation dose efficiency of dual-energy CT benchmarked against single-source, kilovoltage-optimized scans.** *Br J Radiol* 2016;89:20150486 CrossRef Medline
- Ommen F, Bennink E, Vlassenbroek A, et al. **Image quality of conventional images of dual-layer SPECTRAL CT: a phantom study.** *Med Phys* 2018;45:3031–42 CrossRef Medline
- Neuhaus V, Abdullayev N, Hokamp NG, et al. **Improvement of image quality in unenhanced dual-layer CT of the head using virtual monoenergetic images compared with polyenergetic single-energy CT.** *Invest Radiol* 2017;52:470–76 CrossRef Medline
- Reimer RP, Flatten D, Lichtenstein T, et al. **Virtual monoenergetic images from spectral detector CT enable radiation dose reduction in unenhanced cranial CT.** *AJNR Am J Neuroradiol* 2019;40:1617–23 CrossRef Medline
- Lennartz S, Laukamp KR, Neuhaus V, et al. **Dual-layer detector CT of the head: Initial experience in visualization of intracranial**

- hemorrhage and hypodense brain lesions using virtual monoenergetic images. *Eur J Radiol* 2018;108:177–83 CrossRef Medline
25. Kamdem FE, Ngano SO, Alla Takam C, et al. **Optimization of pediatric CT scans in a developing country.** *BMC Pediatr* 2021;21:44 CrossRef Medline
 26. Tan XM, Shah MT, Chong SL, et al. **Differences in radiation dose for computed tomography of the brain among pediatric patients at the emergency departments: an observational study.** *BMC Emerg Med* 2021;21:106 CrossRef Medline
 27. Brenner DJ, Hall EJ. **Computed tomography—an increasing source of radiation exposure.** *N Engl J Med* 2007;357:2277–84 CrossRef Medline
 28. Xu H, Sun QF, Yue BR, et al. **Results and analysis of examination doses for paediatric CT procedures based on a nationwide survey in China.** *Eur Radiol* 2023;6:10005–7 CrossRef Medline
 29. Almén A, Guðjónsdóttir J, Heimland N, et al. **Paediatric diagnostic reference levels for common radiological examinations using the European guidelines.** *Br J Radiol* 2022;95:20210700 CrossRef Medline
 30. Kanal KM, Butler PF, Chatfield MB, et al. **U.S. diagnostic reference levels and achievable doses for 10 pediatric CT examinations.** *Radiology* 2022;302:164–74 CrossRef Medline
 31. Matsunaga Y, Chida K, Kondo Y, et al. **Diagnostic reference levels and achievable doses for common computed tomography examinations: results from the Japanese nationwide dose survey.** *Br J Radiol* 2019;92:20180290 CrossRef Medline
 32. Hayton A, Wallace A. **Derivation of Australian diagnostic reference levels for paediatric multi detector computed tomography.** *Australas Phys Eng Sci Med* 2016;39:615–26 CrossRef Medline
 33. Bondt TD, Mulkens T, Zanca F, et al. **Benchmarking pediatric cranial CT protocols using a dose tracking software system: a multicenter study.** *Eur Radiol* 2017;27:841–50 CrossRef Medline
 34. Ploussi A, Syrgiamiotis V, Makri T, et al. **Local diagnostic reference levels in pediatric CT examinations: a survey at the largest children's hospital in Greece.** *Br J Radiol* 2020;93:20190358 CrossRef Medline
 35. Chu PW, Yu S, Wang Y, et al. **Reference phantom selection in pediatric computed tomography using data from a large, multicenter registry.** *Pediatr Radiol* 2022;52:445–52 CrossRef Medline
 36. da Silva EH, Baffa O, Elias J jr, et al. **Conversion factor for size specific dose estimation of head CT scans based on age, for individuals from 0 up to 18 years old.** *Phys Med Biol* 2021;66 CrossRef Medline

Arterial Spin-Labeling in the Assessment of Pediatric Nontraumatic Orbital Lesions

 S. Neumane, A. Lesage,  V. Dangouloff-Ros, R. Levy,  C.-J. Roux, M.P. Robert,  D. Bremond-Gignac, and  N. Boddaert



ABSTRACT

SUMMARY: Benign and malignant pediatric orbital lesions can sometimes have overlapping features on conventional MR imaging sequences. MR imaging of 27 children was retrospectively reviewed to describe the signal of some common pediatric extraocular orbital lesions on arterial spin-labeling and to evaluate whether this sequence helps to discriminate malignant from benign masses, with or without ADC value measurements. Qualitative and quantitative assessments of arterial spin-labeling CBF and ADC were performed. All lesions were classified into 3 arterial spin-labeling perfusion patterns: homogeneous hypoperfusion (pattern 1, $n = 15$; benign lesions), heterogeneous hyperperfusion (pattern 2, $n = 9$; cellulitis, histiocytosis, malignant tumors), and homogeneous intense hyperperfusion (pattern 3, $n = 3$; infantile hemangiomas). Arterial spin-labeling can be a valuable tool to improve the diagnostic confidence of some orbital lesions, including infantile hemangioma. An algorithm is proposed.

ABBREVIATIONS: ASL = arterial spin-labeling; av = averaged; IH = infantile hemangioma; LBF = lesion blood flow; OL = orbital lesion; ONG = optic nerve glioma; rLBF = relative lesion blood flow; RMS = rhabdomyosarcoma

A wide variety of benign and malignant lesions can present in the orbital region, ranging from developmental anomalies to infectious processes and malignancies. Imaging is essential for their evaluation and to guide intervention, and MR imaging is the main technique used in children. Many pediatric orbital lesions (OLs) have a distinct appearance on conventional sequences, but benign and malignant processes can have a similar appearance. DWI and ADC values have been increasingly used to characterize OLs. Several studies found relevant differences between benign and malignant lesions, but with a different range of reported sensitivities, specificities, and optimal threshold ADC values for distinguishing between lesions.¹⁻⁴

MR imaging perfusion is increasingly used in the evaluation of head and neck anomalies and cerebral tumors,⁵⁻⁷ but few studies describe its potential use in the evaluation of OLs. Some studies showed that the combination of DWI and dynamic contrast-enhanced MR imaging considerably improved the specificity and sensitivity for differentiating radiologically indeterminate malignant from benign orbital masses.^{8,9} Arterial spin-labeling (ASL) is a noninvasive MR imaging perfusion technique that provides measurements reflecting the degree of lesion angiogenesis and vascular density.⁵ Its role in the characterization of OLs in children has not been studied yet. ASL tumor blood flow and ADC values (alone and in combination) have been shown to be useful in differentiating orbital lymphoma from idiopathic inflammatory pseudotumor in an adult cohort.⁹

The purpose of this study was to describe the ASL signal of some common pediatric extraocular OLs and to evaluate its relevance in discriminating malignant from benign masses, with or without ADC value measurements.

MATERIALS AND METHODS


Study Design and Subjects

This retrospective descriptive study included 27 children (mean age, 4.2 [SD, 4.3] years; range, 0.2–15.0 years) referred to Necker-Enfants Malades University Hospital between January 2015 and September 2022, who underwent MR imaging for evaluation of a nontraumatic extraocular OL with a routine protocol including ASL perfusion. Patients without a confirmed diagnosis were

Received March 8, 2023; accepted after revision August 2.

From the Departments of Pediatric Radiology (S.N., A.L., V.D.-R., R.L., C.-J.R., N.B.) and Ophthalmology (M.P.R., D.B.-G.), Hôpital Universitaire Necker Enfants Malades, Assistance Publique–Hôpitaux de Paris, Université Paris Cité, Paris, France; NeuroSpin, UNIACT (S.N.), Commissariat à l’Énergie Atomique et aux Énergies Alternatives (CEA), Université Paris-Saclay, Gif-sur-Yvette, France; Department of Pediatric Radiology (A.L.), Centre Hospitalier Universitaire Saint-Justine, Montréal, Quebec, Canada; Institut Imagine (V.D.-R., R.L., C.-J.R., N.B.), Institut National de la Santé et de la Recherche Médicale (INSERM) U1163 and U1299, Université Paris Cité, Paris, France; Centre Borelli UMR 9010 (M.P.R.), Université Paris Cité, Université Paris-Saclay, ENS Paris-Saclay, CNRS, SSA, Institut National de la Santé et de la Recherche Médicale (INSERM), Paris, France; and Team T7 (D.B.-G.), Institut National de la Santé et de la Recherche Médicale, UMR51138, Université Paris Cité, Paris, France.

Please address correspondence to Nathalie Boddaert, MD, PhD, Pediatric Radiology Department, Hôpital Universitaire Necker Enfants Malades, AP-HP, Université Paris Cité, F-75015, Paris, France; e-mail: nathalie.boddaert@aphp.fr

 Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A7977>

excluded. Initial symptoms and final diagnoses were obtained from the electronic medical record. Approval was obtained from the institutional ethics board (CRM-2301–324). According to local regulations, consent was waived for this retrospective analysis of anonymized data. This study was recorded in the general register of the Assistance Publique-Hôpitaux de Paris in January 2023 (No. 2023 0126171431).

MR Imaging Acquisition

An ASL perfusion sequence was acquired on a 1.5T or 3T MR imaging scanner (Signa HDxt; GE Healthcare) and consisted of a concurrent spiral 3D pseudocontinuous ASL (12-channel head coil; 3D volume with a 4-mm section thickness using a fast spin-echo acquisition [TR/TE = 4453/10.96 ms; number of axial slices = 40; resolution in plane = 1.875×1.875 mm; postlabeling delay = 1025 ms; flip angle = 155°]). All patients underwent a routine MR imaging examination, including T1 and T2 sequences with additional sequences performed at the discretion of the attending radiologist (gadolinium injection, DWI), depending on the indications.

MR Imaging: Qualitative and Quantitative Analysis

All the cases were reviewed in radio-ophthalmology staff meetings. MR imaging data were retrospectively reviewed in consensus by S.N. and a senior pediatric neuroradiologist, R.L.

ASL. Qualitative assessment of CBF maps was made by coregistration with conventional sequences to describe the signal intensity of individual lesions (hyperintense, warm colors; isointense and hypointense, cold colors) and homogeneity. Lesion blood flow (LBF) measurements (in milliliters/minute/100 g of tissue) were obtained using dedicated postprocessing software on Advantage Windows workstations (GE Healthcare). Freehand ROIs were placed on a minimum of 3 axial plans over the most representative area for small or homogeneous lesions and over the maximal and minimal signal areas of large or heterogeneous masses. Measurements were averaged (av) (LBFav) for each lesion. To take into account general signal variations between subjects and facilitate comparison with other studies, we calculated the relative LBF (rLBF) or lesion-to-cortex signal ratio by comparing the LBFav of each lesion with the average CBF obtained by freehand ROIs drawn over the ipsilateral temporal cortex (CBFav_{cortex}) (rLBF = LBFav/CBFav_{cortex}).

DWI-ADC. The ADC maps were processed in a manner comparable with ASL maps, using similar manual ROIs and similar averaging (ADCav_{lesion} and ADCav_{cortex}), and lesion-to-cortex ADC ratios (rADC = ADCav_{lesion}/ADCav_{cortex}) were calculated.

Data Analysis. Statistical analyses were performed using R 3.3.3 statistical and computing software (<http://www.r-project.org/>). Results are presented as medians and ranges in the text. Quantitative rLBF data were compared among the 3 patterns using 1-way ANOVA.

RESULTS

We analyzed 27 OLs, including 23 benign lesions (vascular [$n=9$] and nonvascular [$n=9$] masses, infectious processes [$n=3$], optic nerve glioma [ONG] [$n=2$]) and 4 malignant

tumors (rhabdomyosarcomas [RMS, $n=2$], metastatic retinoblastoma [$n=1$], and metastatic neuroblastoma [$n=1$]), with histopathologic evidence available for most cases requiring confirmation other than clinical or radiologic (Online Supplemental Data).

ASL

Qualitative and quantitative assessment of ASL maps showed 3 significantly distinct profiles ($F_{(2, 24)} = 86.44$; $P < .001$): pattern 1, homogeneous hypoperfusion ($n=15$); pattern 2, heterogeneous hyperperfusion ($n=9$); and pattern 3, homogeneous intense hyperperfusion ($n=3$) (Online Supplemental Data and Fig 1). Pattern 1 comprised benign OLs, including lymphatic malformation dermoid cysts, dacryops, fibrolipoma, and ONG, with median rLBF = 0.44 (range, 0.21–0.63). On the opposite side of the spectrum, pattern 3 was characteristic of infantile hemangiomas (IHs), with a median rLBF = 10.43 (range, 9.56–18.73). Pattern 2 (median rLBF = 1.23; range, 0.91–2.91) included orbital cellulitis, histiocytosis, and malignant neoplasms with overlapping rLBF (median = 2.75; range, 1.0–2.91; median = 1.07; range, 0.91–1.23; and median = 1.54; range, 1.14–1.88, respectively). There was no overlap between the rLBF range of pattern 3 and that of other groups. Most (78%) benign OLs presented as homogeneous ASL perfusion (patterns 1 and 3), while all malignant tumors were heterogeneous (pattern 2).

DWI-ADC

DWI was available in 78% of cases ($n=21$) (Online Supplemental Data and Fig 2), among which ADC values were decreased in all malignant tumors (median ADC = $0.85 \times 10^{-3} \text{ mm}^2/\text{s}$), abscessed portions of infectious processes (median ADC = $0.77 \times 10^{-3} \text{ mm}^2/\text{s}$), dermoid cysts (median ADC = $0.65 \times 10^{-3} \text{ mm}^2/\text{s}$), and the non-Langerhans cell histiocytosis. ADC values were increased for all IHs (median ADC = $1.69 \times 10^{-3} \text{ mm}^2/\text{s}$), orbital cellulitis (median ADC = $2.20 \times 10^{-3} \text{ mm}^2/\text{s}$), lymphatic malformations (median ADC = $2.60 \times 10^{-3} \text{ mm}^2/\text{s}$), and cases of dacryops, fibrolipoma, and Langerhans cell histiocytosis.

DISCUSSION

Pediatric orbital masses are a heterogeneous group that can present with various and nonspecific manifestations. Clinical presentation and conventional imaging features do not always allow an accurate diagnosis. Given the substantial risk associated with orbital biopsies, additional predictors of the histologic nature of OLs are useful. We have found, in this case series, 3 distinct patterns on ASL maps that allow classifying all lesions: homogeneous hypoperfusion (pattern 1, including only benign lesions), heterogeneous hyperperfusion (pattern 2, including orbital infectious processes, histiocytosis, and malignant neoplasms), and homogeneous very intense hyperperfusion (pattern 3, characteristic of IHs, without rLBF values overlapping with those of other groups) (Figs 3 and 4).

Within pattern 2, there was an rLBF overlap among the OL subgroups, with ADC values refining the characterization of these lesions because ADC values were increased for orbital cellulitis, variable for histiocytosis, and decreased (diffusion restriction) for malignant tumors. Across the entire cohort, ADC values were decreased in malignant tumors, while increased in most benign

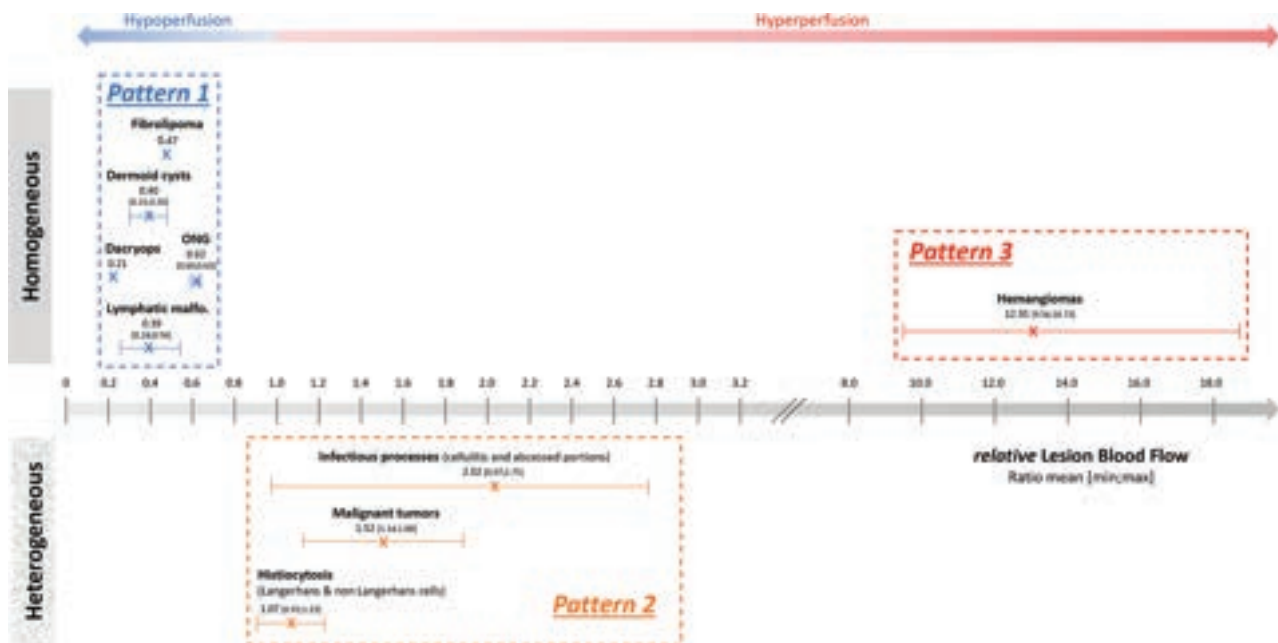


FIG 1. Representation of ASL perfusion profiles of the 27 OMs, according to their homogeneity and signal intensity. Mean rLBF and ranges by diagnostic subgroups show 3 significantly distinct profiles ($P = 1.08 \times 10^{-11}$): homogeneous hypoperfusion (pattern 1, blue), heterogeneous iso- or hyperperfusion (pattern 2, orange), and homogeneous hyperperfusion (pattern 3, red). rLBF corresponds to the ratio between LBF and CBF from the ipsilateral temporal cortex.

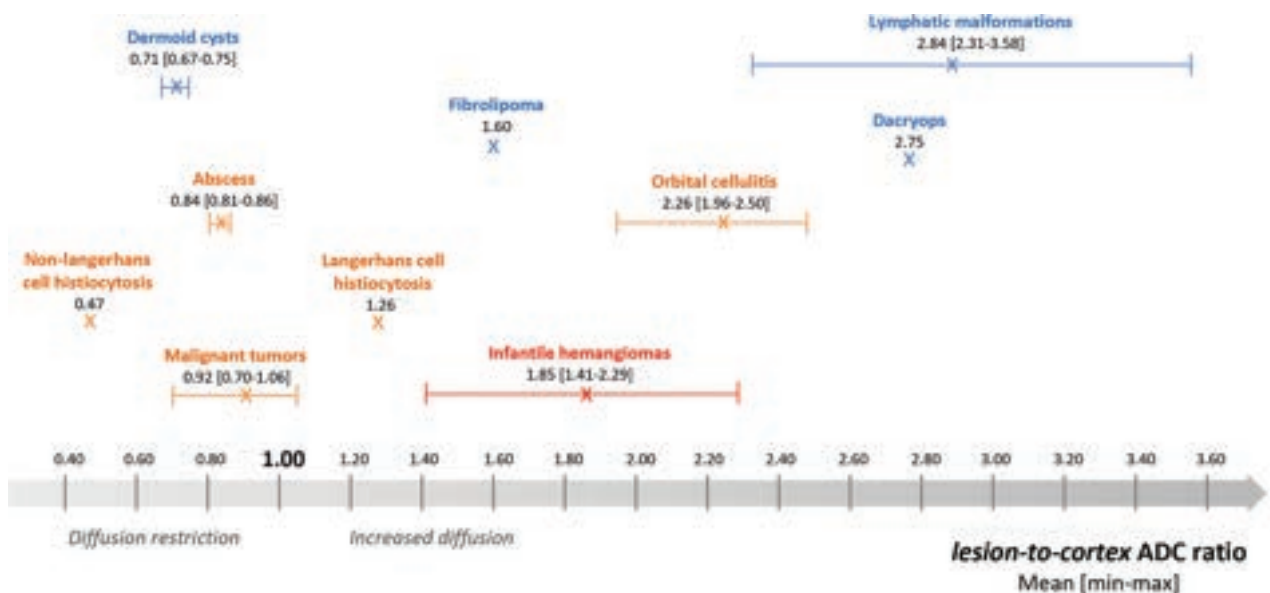


FIG 2. Distribution of ADC ratios ($n = 21$) by diagnostic subgroups. Lesion-to-cortex ADC ratios were calculated by dividing the lesion mean ADC by the ipsilateral temporal cortex mean ADC (both directly obtained for each patient by drawing an ellipsoid ROI over an area similar to that used for ASL measurements, taking care to avoid volume averaging with adjacent structures, particularly fat and bone). The distinct subgroups of infectious processes (eg, abscessed areas with characteristic diffusion restriction and denser tissular inflammatory parts, cellulitis) were assessed separately. Colors correspond to the ASL perfusion profile groups presented in Fig 1 (pattern 1: blue; pattern 2: orange; pattern 3: red).

masses, in line with previous studies reporting lower ADC values in malignant orbital tumors than in benign orbital masses in adult^{1,10} and pediatric populations.^{3,4}

In agreement with previous studies,^{8,9,11} this work confirms the interest in using complementary diffusion and perfusion MR imaging techniques for characterizing orbital masses, while

suggesting that ASL might be more informative than ADC when it comes to differentiating benign vascular lesions from malignant tumors.

The specificity of pattern 3 is of particular interest. There can be an overlap between the imaging appearance of IHs and malignant tumors, mainly RMS,^{12,13} because they may have similar signal on

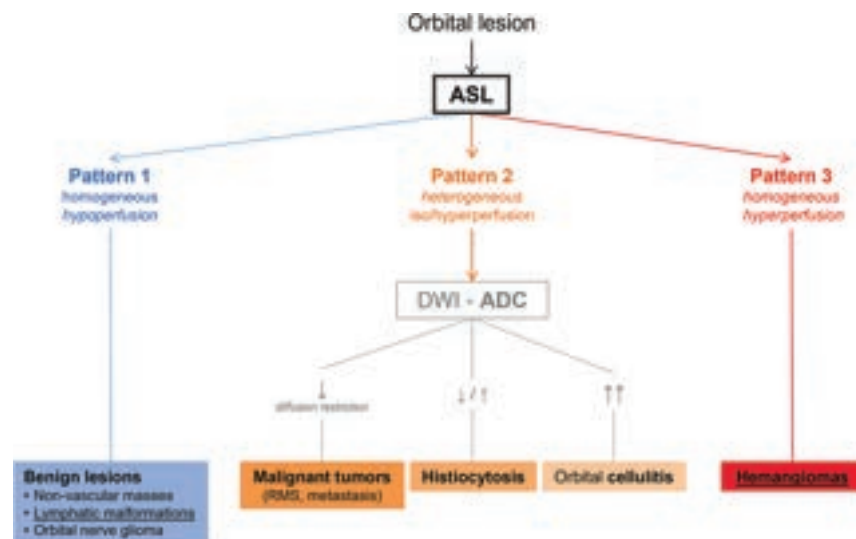


FIG 3. Summary algorithm for radiologic interpretation. There are 3 distinct patterns derived from ASL perfusion analysis: Pattern 1 (blue) corresponds to benign masses presenting homogeneous low lesion blood flow; pattern 2 (orange) describes heterogeneous iso- or hyperperfused lesions, including malignancies; pattern 3 (red) is characteristic of hemangiomas (benign vascular lesions), presenting specific homogeneous and very intense hyperperfusion. ADC assessment within pattern 2 contributed to differentiating malignant tumors (presenting diffusion restriction) from benign lesions like orbital cellulitis (with increased diffusion). Lymphatic malformations (pattern 1) and IHs (pattern 3), two frequent vascular orbital anomalies (underlined) in pediatric population, presented with very distinct ASL perfusion homogeneous profiles.

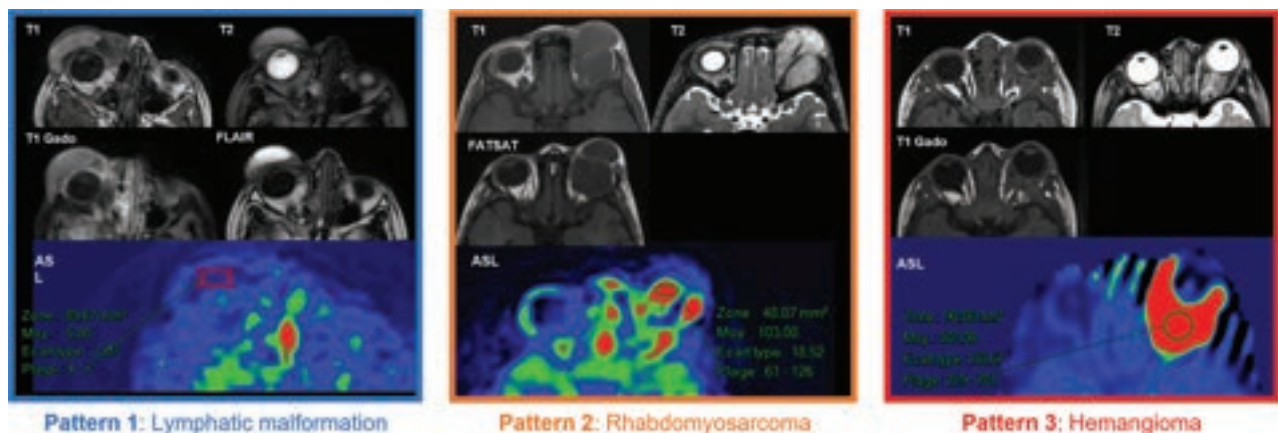


FIG 4. Conventional MR imaging sequences and ASL maps of representative patients for the 3 perfusion patterns. Pattern 1 (homogeneous hypoperfusion profile, characterized by cold dark colors on the CBF map over the lesion area): Lymphatic malformation, right vascular anomaly (patient 12, mean LBF = 10 mL/min/100g tissue, rLBF = 0.24). Pattern 2 (heterogeneous hyperperfusion profile, consisting of multiple bright color patches within the lesion area): Rhabdomyosarcoma, left malignant tumor (patient 24, mean LBF = 74 mL/min/100g tissue, rLBF = 1.84). Pattern 3 (homogeneous very intense hyperperfusion profile, corresponding to a uniform bright warm color area in the CBF maps): IH, left vascular anomaly (patient 16, mean LBF = 515 mL/min/100g tissue, rLBF = 18.73). Conventional sequences: T1 and T2 WI, T1 Gado (T1WI with gadolinium-based contrast medium), FLAIR, FATSAT (fat-suppression T1WI sequence).

T2WI and T1WI+C.^{12,14} IHs present earlier in life, are more often homogeneous on T1WI+C, are usually more hyperintense on T2WI than RMS, and commonly present with flow voids.^{13,14} They also have higher ADC values than RMS.¹² Adding ASL data may help with the differential diagnosis: in our cohort, all IHs (in the proliferating phase at the time of MR imaging) had a very distinct appearance, with homogeneous, very high perfusion signal (pattern 3), in accordance with what has been described in other cervicofacial locations.⁶ No other lesion presented an ASL signal as intense as that in IHs, including malignant tumors.

Moreover, IHs and lymphatic malformations, which are among the most frequent vascular anomalies of the pediatric

orbit, had very distinct ASL perfusion homogeneous profiles in our cohort (patterns 3 and 1, respectively).

ASL is a robust imaging sequence. It has a relatively rapid acquisition time and is noninvasive (does not require a venous access and the injection of exogenous contrast medium), 2 very valuable factors in the pediatric population. It has some limitations, including being sensitive to motion and susceptibility artifacts from metal and having a low SNR. The ASL sequence also has a limited spatial resolution, which makes it less useful in the evaluation of very small lesions.

We have chosen to present average LBF values (instead of maximal values), given that most of the analyzed lesions had a

homogeneous signal. In addition, the generally small size of OLS does not always allow calculating values for different lesion areas.

The main limitation of this study is the number and variety of cases, especially those with malignant tumors. Thus, despite the interesting categorization of pediatric OLS according to their ASL perfusion patterns, the proposed algorithm requires verification in larger cohorts that include a larger spectrum of OLS.

Despite these limitations, our descriptive study offers a novel framework of distinctive hemodynamic properties of some common pediatric OLS, proposing an algorithm that may contribute to the diagnostic orientation of OLS by combining, for the first time, ASL and DWI-ADC.

CONCLUSIONS

ASL perfusion-weighted MR imaging can be a valuable noninvasive tool to improve the discrimination among a full spectrum of benign and malignant orbital processes, by refining the short differential provided by the clinical presentation and conventional sequences. Especially, ASL may improve the diagnostic confidence of benign vascular lesions, in particular IHs requiring medical treatment. Furthermore, information provided by relevant orbital MR imaging is valuable for the oculoplastic surgeon for guiding an intervention. We hope to encourage other teams to integrate ASL into their routine orbital MR imaging protocols to explore its potential.

ACKNOWLEDGMENTS

We thank the children and their parents for their participation in this study.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

1. Sepahdari AR, Politi LS, Aakalu VK, et al. **Diffusion-weighted imaging of orbital masses: multi-institutional data support a 2-ADC threshold model to categorize lesions as benign, malignant, or indeterminate.** *AJNR Am J Neuroradiol* 2014;35:170–75 CrossRef Medline
2. Sepahdari AR, Aakalu VK, Kapur R, et al. **MRI of orbital cellulitis and orbital abscess: the role of diffusion-weighted imaging.** *AJR Am J Roentgenol* 2009;193:W244–50 CrossRef Medline
3. Maldonado FR, Princich JP, Micheletti L, et al. **Quantitative characterization of extraocular orbital lesions in children using diffusion-weighted imaging.** *Pediatr Radiol* 2021;51:119–27 CrossRef Medline
4. Jaju A, Rychlik K, Ryan ME. **MRI of pediatric orbital masses: role of quantitative diffusion-weighted imaging in differentiating benign from malignant lesions.** *Clin Neuroradiol* 2020;30:615–24 CrossRef Medline
5. Dangouloff-Ros V, Deroulers C, Foissac F, et al. **Arterial spin-labeling to predict brain tumor grading in children: correlations between histopathologic vascular density and perfusion MR imaging.** *Radiology* 2016;281:553–66 CrossRef Medline
6. Boulouis G, Dangouloff-Ros V, Boccara O, et al. **Arterial spin-labeling to discriminate pediatric cervicofacial soft-tissue vascular anomalies.** *AJNR Am J Neuroradiol* 2017;38:633–38 CrossRef Medline
7. Mamlouk MD, Hess CP. **Arterial spin-labeled perfusion for vascular anomalies in the pediatric head and neck.** *Clin Imaging* 2016;40:1040–46 CrossRef Medline
8. Xu XQ, Qian W, Ma G, et al. **Combined diffusion-weighted imaging and dynamic contrast-enhanced MRI for differentiating radiologically indeterminate malignant from benign orbital masses.** *Clin Radiol* 2017;72:903.e9–903–15 CrossRef Medline
9. Ro SR, Asbach P, Siebert E, et al. **Characterization of orbital masses by multiparametric MRI.** *Eur J Radiol* 2016;85:324–36 CrossRef Medline
10. Razek AA, Elkhamary S, Mousa A. **Differentiation between benign and malignant orbital tumors at 3-T diffusion MR-imaging.** *Neuroradiology* 2011;53:517–22 CrossRef Medline
11. Eissa L, Abdel Razek AA, Helmy E. **Arterial spin-labeling and diffusion-weighted MR imaging: utility in differentiating idiopathic orbital inflammatory pseudotumor from orbital lymphoma.** *Clin Imaging* 2021;71:63–68 CrossRef Medline
12. Kralik SF, Haider KM, Lobo RR, et al. **Orbital infantile hemangioma and rhabdomyosarcoma in children: differentiation using diffusion-weighted magnetic resonance imaging.** *J AAPOS* 2018;22:27–31 CrossRef Medline
13. Sarioglu FC, Sarioglu O, Guleryuz H, et al. **MRI-based texture analysis for differentiating pediatric craniofacial rhabdomyosarcoma from infantile hemangioma.** *Eur Radiol* 2020;30:5227–36 CrossRef Medline
14. Lope LA, Hutcheson KA, Khademian ZP. **Magnetic resonance imaging in the analysis of pediatric orbital tumors: utility of diffusion-weighted imaging.** *J AAPOS* 2010;14:257–62 CrossRef Medline

Cerebral Hemodynamic and Metabolic Abnormalities in Neonatal Hypocalcemia: Findings from Advanced MRI

Ying Qi, Zixuan Lin, Hanzhang Lu, Jian Mao, Hongyang Zhang, Pengfei Zhao, and Yang Hou



ABSTRACT

BACKGROUND AND PURPOSE: Neonatal hypocalcemia is the most common metabolic disorder, and whether asymptomatic disease should be treated with calcium supplements remains controversial. We aimed to quantify neonatal hypocalcemia's global CBF and cerebral metabolic rate of oxygen (CMRO₂) using physiologic MR imaging and elucidate the pathophysiologic vulnerabilities of neonatal hypocalcemia.

MATERIALS AND METHODS: A total of 37 consecutive patients with neonatal hypocalcemia were enrolled. They were further divided into subgroups with and without structural MR imaging abnormalities, denoted as neonatal hypocalcemia-a ($n = 24$) and neonatal hypocalcemia-n ($n = 13$). Nineteen healthy neonates were enrolled as a control group. Brain physiologic parameters determined using phase-contrast MR imaging, T2-relaxation-under-spin-tagging MR imaging, and brain volume were compared between patients with neonatal hypocalcemia (their subgroups) and controls. Predictors for neonatal hypocalcemia-related brain injuries were identified using multivariate logistic regression analysis and expressed as ORs with 95% CIs.

RESULTS: Patients with neonatal hypocalcemia showed significantly lower CBF and CMRO₂ compared with controls. Furthermore, the neonatal hypocalcemia-a subset (versus controls or neonatal hypocalcemia-n) had significantly lower CBF and CMRO₂. There was no obvious difference in CBF and CMRO₂ between the neonatal hypocalcemia-n subset and controls. CBF and CMRO₂ were independently associated with neonatal hypocalcemia. The ORs were 0.80 (95% CI, 0.65–0.99) and 0.97 (95% CI, 0.89–1.05) for CBF and CMRO₂, respectively.

CONCLUSIONS: Neonatal hypocalcemia with structural damage may exhibit lower hemodynamics and cerebral metabolism. CBF may be useful in assessing the need for calcium supplementation in asymptomatic neonatal hypocalcemia to prevent brain injury.

ABBREVIATIONS: AUC = area under the curve; CMRO₂ = cerebral metabolic rate of oxygen; iCa = ionized calcium; NHC = neonatal hypocalcemia; NRDS = neonatal respiratory distress syndrome; PCMR = phase-contrast MR imaging; PVWM = periventricular WM; TRUST = T2-relaxation-under-spin-tagging; Yv = venous oxygenation

Neonatal hypocalcemia (NHC) is a common abnormality in neonates.¹ Prematurity, perinatal asphyxia, sepsis, maternal

vitamin D deficiency, diabetes, hypoparathyroidism, and hypomagnesemia are the most common causes of NHC.^{2–5} Early- and late-onset NHC occurs before and after the first 3 days of life, respectively. The clinical presentation of NHC is frequently asymptomatic but can be symptomatic, showing neuromuscular irritability, such as myoclonic jerks, jitteriness, exaggerated startles, and seizures. NHC can be life-threatening⁶ and is a risk factor for mortality in critically ill infants.⁷ NHC can be prevented or treated with long-term calcium gluconate therapy specific to its etiology. However, calcium supplementation should be administered carefully in patients with hypoxic-ischemic encephalopathy. Hypoxia causes a rapid influx of calcium into cells, resulting in apoptosis, while calcium supplementation can aggravate brain damage in these patients.⁵ In addition, whether asymptomatic NHC should be treated with calcium supplementation is controversial because it may cause neonatal tetany or seizures.¹

Received June 12, 2023; accepted after revision August 16.

From the Departments of Radiology (Y.Q., H.Z., Y.H.) and Pediatrics (J.M.), Shengjing Hospital of China Medical University, Shenyang, China; Key Laboratory for Biomedical Engineering of Ministry of Education (Z.L.), Department of Biomedical Engineering, College of Biomedical Engineering & Instrument Science, Zhejiang University, Hangzhou, China; Department of Radiology (H.L.), Johns Hopkins University School of Medicine, Baltimore, Maryland; and Department of Pharmacology (P.Z.), School of Pharmaceutical Sciences, China Medical University, Shenyang, China.

This study was supported by the Applied Basic Research Project of Liaoning (2022JH2/101500061), the Medical Education Research Project of Liaoning (No.2022-N006-01), the New Technology Project of Shengjing Hospital, and the 345 Talent Project of Shengjing Hospital of China Medical University.

Pengfei Zhao, and Yang Hou contributed equally to this work.

Please address correspondence to Yang Hou, PhD, 36# of Sanhao St, Heping District, Shenyang, China; e-mail: houyang1973@163.com

Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A7994>

Table 1: Demographic and clinical data of the NHC group and controls

	NHC (n = 37)	NHC-a (n = 24)	NHC-n (n = 13)	Control (n = 19)	<i>P</i> _{NHC & Control} Value
Demographics					
Males (No.) (%)	24 (64.9)	14 (58.3)	7 (53.8)	14 (73.7)	.11
Birth weight (mean) (g)	2024.9 (SD, 614.5)	2080.8 (SD, 532.1)	1676.5 (SD, 695.9)	2434.7 (SD, 1047.5)	.13
Scan age (mean) (wk)	37.0 (SD, 2.9)	37.2 (SD, 3.2)	36.0 (SD, 1.8)	37.0 (SD, 4.0)	.99
Clinical					
Apgar score (mean)					
1 min	7.6 (SD, 3.2)	7.0 (SD, 3.5)	9.4 (SD, 0.5)	9.6 (SD, 0.8)	<.05
5 min	6.2 (SD, 3.3)	5.5 (SD, 3.5)	8.0 (SD, 1.3)	8.6 (SD, 2.1)	<.05
NRDS (No.) (%)	16 (43.2)	13 (54.2)	3 (23.1)	—	—
Twins (No.) (%)	6 (16.2)	5 (20.8)	1 (7.7)	—	—
iCa (mean) (mmol/L)	0.9 (SD, 0.2)	0.9 (SD, 0.2)	0.8 (SD, 0.1)	1.2 (SD, 0.1)	<.05

Note:—The en dash indicates none.

Therefore, gaining a better understanding of the neurobiologic characteristics of NHC and its relationship with serum calcium levels may provide important insights into neonatal pathophysiologic vulnerabilities.

CT perfusion,^{8,9} ¹⁵O-PET,¹⁰ ¹³³Xe-enhanced CT,¹¹ DSC MR imaging,¹² and near-infrared spectroscopy¹³ are common methods of measuring the hemodynamics and metabolism in the neonatal brain but require ionizing radiation and exogenous tracers, which can cause kidney injury and may create motion artifacts or low spatial resolution. Doppler sonography is used to evaluate brain perfusion but does not allow whole-brain perfusion analyses and is not always reproducible, especially when performed by less-experienced neonatal operators.¹⁴ Recent advances in MR imaging technologies allow noninvasive and quantitative measurements of brain hemodynamic parameters without the need for contrast material injection or exposure to ionizing radiation. Global CBF can be measured using arterial spin-labeling¹⁵ and phase-contrast MR imaging (PCMR).¹⁶ CBF measurements using arterial spin-labeling agree with those of ¹⁵O-PET and DSC MR imaging.^{17,18} However, low CBF and neonatal brain volume reduce the accuracy of arterial spin-labeling quantitative measurement, especially in deep white matter.¹⁹ PCMR has been successfully reported in neonates and requires <2 minutes.²⁰ Venous oxygenation (Yv) of the brain can be determined using the T2-relaxation-under-spin-tagging (TRUST) MR imaging technique. Yv and CBF can be combined to assess the global cerebral metabolic rate of oxygen (CMRO₂).^{21,22} All these physiologic measurements require <5 minutes, which is suitable for the neonatal population. Most neonatal studies have focused on low birth weight,²³ brain injury such as hypoxic-ischemic encephalopathy detected with near-infrared spectroscopy,²⁴ and WM lesions detected with PCMR and TRUST MR imaging.²⁰ Few perinatal studies have been conducted on the effects of NHC on cerebral oxygen metabolism and hemodynamics.

The present study aimed to comprehend the relationship among NHC, its hemodynamics, and cerebral oxygen metabolism using PCMR and TRUST MR imaging for a better understanding of its pathophysiologic process and improvement of treatment.

MATERIALS AND METHODS

Ethics Statement

The study protocol was approved by the Ethics Committee of the Shengjing Hospital of China Medical University (IRB2021PS094K).

Written informed consent was waived because the data were collected as part of a clinically indicated MR imaging. None of the authors has any conflict of interest to declare. Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Participants and Study Design

Between June 2021 and April 2022, fifty-six neonates were enrolled. The participants were categorized into the NHC and control groups. The inclusion criteria for the NHC group were asymptomatic neonates with early-onset NHC occurring in the first 3 days of life. Participants were further divided into those with abnormal MR imaging findings, denoted as NHC-a, and those with normal MR imaging findings, denoted as NHC-n. NHC is defined as ionized calcium (iCa) <4 mg/dL (1 mmol/L) in preterm infants and <1.2 mmol/L in term neonates.

The inclusion criteria for the control group were asymptomatic neonates without NHC and abnormal MR imaging findings or laboratory results at birth. They had normal results from umbilical cord blood gas analysis, the liver function test, routine blood examination, and serum and ion analyses at birth. The purpose of MR images was for exclusions of abnormalities of the nervous system. The exclusion criteria were congenital malformations, severe infection, major congenital heart disease, and unusable MR images. Neonates' demographic and clinical data are summarized in Table 1. Detailed information is provided in the Online Supplemental Data (There were 27 cases [sodium levels <135 mmol/L] of hypocalcemia with low serum sodium, including 11 cases with sodium levels of <130 mmol/L [range, NHC-n, 123.4–145.6 mmol/L versus NHC-a, 121.9–144.0 mmol/L]). Two experienced neuro-radiologists (Y.Q. and X.X.Q., each with 12 years of experience and blinded to the group data) diagnosed and graded the injuries on MR imaging as described on a scale of I–IV, according to their size, number, and distribution, as summarized in previous Table 1 reported by Childs et al.²⁵

MR Imaging Protocol

Imaging was performed on a 3T scanner (Ingenia; Philips Healthcare) without sedation. Axial T1WI and T2WI were acquired using the following settings: TR, 200/5000 ms; TE, 2.3/80 ms;

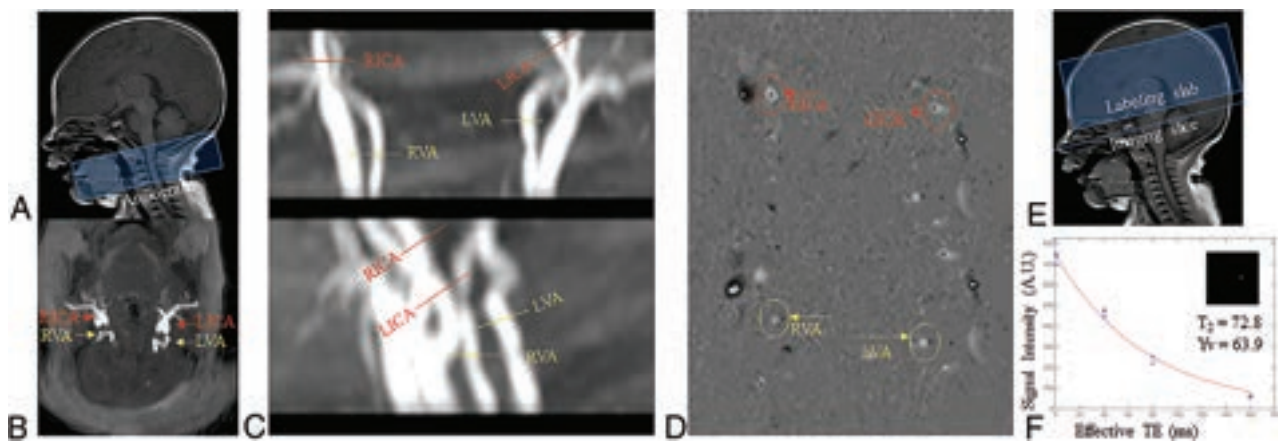


FIG 1. Sample images for CBF and Yv determination. *A*, Positioning of the imaging section. *B* and *C*, PCMR imaging section locations for the LICA, RICA, LVA, and RVA on the MIP image of a TOF angiogram. *D*, The manual delineation of LICA, RICA, LVA, and RVA was conducted to determine flux in these vessels. The sum of fluxes by brain weight yielded the CBF. *E*, The imaging section was positioned parallel to the intercommissural line with a 10-mm distance from the superior sagittal sinus (SSS). TRUST MR imaging used a spin-labeling module to isolate pure venous blood signals in the SSS. *F*, Blood signals from the SSS were fitted to a monoexponential function of the effective TE to yield blood T2. T2 was converted to Yv via a calibration plot. LICA indicates left internal carotid arteries; RICA, right internal carotid arteries; LVA, left vertebral artery; RVA, right vertebral artery; AU, arbitrary unit.

section thickness, 5/5 mm; FOV, $180 \times 150 \times 90 \text{ mm}^3$; matrix size, $224 \times 162/112 \times 112$; scan time, 34.4/40.9 seconds.

CBF Determination

First, TOF MRA was performed to visualize the left and right ICAs and vertebral arteries. The imaging slab was positioned at the center of the epistropheus, with 60-mm saturation slabs placed above to suppress venous signals. Second, PCMR used bipolar gradients to encode flow velocity in the major feeding arteries, thereby providing a measurement of CBF. PCMR settings were the following: single section; voxel size, $0.5 \times 0.5 \times 3 \text{ mm}^3$; FOV, $180 \times 180 \times 3 \text{ mm}^3$; maximum velocity encoding, 20 cm/s; and total scan time, 1.5 minutes. Manual delineation of the arterial ROI was performed on the magnitude image of each PCMR scan and then applied to the velocity maps to obtain the flux in all vessels, the sum of which yielded the total blood flow (milliliters/minute). To convert the value to CBF, we obtained brain volume by manual tracing of the T2WI and assuming a parenchymal density of 1.06 g/mL.²⁶ CBF was calculated by dividing the total flux by the brain weight (Fig 1A–D).

Arterial and Venous Oxygenation Determination

Arterial oxygenation was measured using a pulse oximeter in the right hand. Yv was measured by TRUST MR imaging with the following settings: 4 effective TEs, 0, 40, 80, and 160 ms; TR, 3000 ms; TI, 1022 ms; FOV, $160 \times 160 \times 5 \text{ mm}^3$; matrix size, 64×61 ; sensitivity encoding factor, 3; voxel size, $2.5 \times 2.5 \times 5 \text{ mm}^3$; thermal Carr-Purcell-Meiboom-Gill (τ CPMG), 10 ms; scan time, 72 seconds.²⁰ TRUST MR imaging applies a spin-labeling module to isolate pure venous blood signals, a series of T2-preparation pulses to modulate the MR imaging signal, and uses monoexponential fitting to yield blood T2, which was converted to Yv using a hematocrit-specific T2–Yv calibration curve. Data processing for PCMR and TRUST MR imaging was performed as previously described (Fig 1E, –F).^{16,20,22}

CMRO₂ Calculation

CMRO₂ can be calculated using the Fick principle,²⁶ in which HCT is hematocrit.

$$\text{CMRO}_2 = \text{CBF} \times (\text{Y}_a - \text{Y}_v) \times \frac{\text{HCT} \times 897}{44\%}.$$

Statistical Analysis

To compare baseline demographics, clinical characteristics, and physiologic parameters among the 3 groups, we used a Student *t* test or Wilcoxon 2-sample exact test (for continuous variables) and a χ^2 test (for categorical variables). To evaluate group differences with respect to scan age, birth weight, and twins, multiple linear regressions were examined to predict the measured physiologic parameters (Yv, CBF, and CMRO₂) and brain volume that can be influenced by NHC using 4 operating input parameters: group (NHC-a, NHC-n, or controls), neonatal scan age, birth weight, and twins. Similar analyses were performed to examine the dependence of physiologic parameters and brain volume on iCa levels. Physiologic parameters and brain volume were assigned as dependent variables, whereas the brain iCa level, scan age, birth weight, and twins were assigned as independent variables. Considering the effects of gestational age, birth weight, Apgar score, sex, and blood gas analysis on structural abnormalities, we used univariate and multivariate logistic regression analysis, respectively. The risk factors for predicting NHC with structural MR imaging abnormalities were identified using multivariate logistic regression analysis expressed as ORs with 95% CIs. Receiver operating characteristic curves were used to determine whether physiologic parameters, iCa levels, scan age, birth weight, or the incorporation of above values together (physiological parameters + iCa levels + scan age + birth weight) could facilitate the diagnosis of neonatal brain injuries and to identify the cut-off points for risk factors associated with neonatal brain injuries. All computations were performed using standard software (SPSS; IBM). *P* < .05 indicates a statistically significant difference.

RESULTS

Participant Characteristics

Figure 2 shows a schematic of participant selection. Neonates' MR imaging findings, focal periventricular WM (PVWM) lesion grades, and their physiologic parameters are summarized in Table 2. Figure 3 shows MR imaging findings of representative neonates in the NHC-a subgroup with different severities. The physiologic parameters of the participants in the NHC and control groups are shown in Table 3. The NHC neonates and the NHC-a subset (versus controls) had significantly lower CBF ($P = .03$, $P = .001$) and CMRO₂ ($P = .02$, $P = .002$). CBF and CMRO₂ proved significantly lower in the NHC-a (than in the NHC-n) subset ($P = .001$, $P = .01$). There were no significant differences in CBF and CMRO₂ between the NHC-n subset and controls ($P = .70$, $P = .63$). Yv and brain volume were not significantly different among groups ($P > .05$). Figure 4 shows Yv, CBF, CMRO₂, and brain volume in both groups and subgroups using boxplots. There were 5 neonates with SAH not secondary to parenchymal bleeding primarily from the germinal matrix. Only 1 neonate with neonatal respiratory distress syndrome (NRDS) in grade III diagnosed by x-ray in our study required a surfactant within 2 hours after birth.

Linear Regression Analysis of Measured Physiologic Parameters Associated with Group (NHC-a, NHC-n, or Control Groups), Scan Age, Birth Weight, and Twins

The multilinear regression analysis showed that CBF and CMRO₂ were significantly correlated with group ($\beta = -2.6$, $P = .001$; $\beta = -6.5$, $P = .001$, respectively). Yv and brain volume

had a nonsignificant relationship with category ($\beta = 0.7$, $P = .69$; $\beta = 13.1$, $P = .20$). Yv, CMRO₂, and brain volume had a significant relationship with scan age ($\beta = -1.2$, $P = .03$; $\beta = 1.3$, $P = .03$; $\beta = 14.5$, $P < .001$). CBF was nonsignificantly correlated with scan age ($\beta = 0.4$, $P = .07$). Yv was significantly correlated with birth weight ($\beta = 0.006$, $P = .01$). CBF, CMRO₂, and brain volume were nonsignificantly correlated with birth weight ($\beta = 0.001$, $P = .53$; $\beta = 0.003$, $P = .30$; $\beta = 0.008$, $P = .55$, respectively). Yv, CBF, and CMRO₂ were nonsignificantly correlated with twins ($\beta = -5.0$, $P = .32$; $\beta = 0.6$, $P = .79$; $\beta = 7.4$, $P = .15$, respectively). Brain volume was significantly correlated with twins ($\beta = -77.2$, $P = .007$) (Online Supplemental Data).

Linear Regression Analysis of Measured Physiologic Parameters Associated with iCa Levels, Scan Age, Birth Weight, and Twins

Multilinear regression analysis showed that CBF and CMRO₂ demonstrated significant increases with iCa levels ($\beta = 8.2$, $P = .01$; $\beta = 24.2$, $P = .003$, respectively). Yv and brain volume were not dependent on iCa levels ($\beta = 2.0$, $P = .79$; $\beta = -7.8$, $P = .85$). Brain volume significantly correlated with twins ($\beta = -70.1$, $P = .01$) (Online Supplemental Data).

Logistic Regression Analysis of Measured Physiologic Parameters, Brain Volume, and iCa Levels to Detect NHC with Brain Injuries

The ORs were 0.80 (95% CI, 0.65–0.99) and 0.97 (95% CI, 0.89–1.05) for CBF and CMRO₂, respectively. The combined values together versus single values (CBF, CMRO₂, brain volume, or iCa levels, respectively) showed a superior capacity to detect NHC with brain injuries (area under curve [AUC] = 0.92 [95% CI, 0.83–0.99]) versus CBF (AUC = 0.82 [95% CI, 0.69–0.96]; $P = .001$), CMRO₂ (AUC = 0.73 [95% CI, 0.56–0.90]; $P = .02$), brain volume (AUC = 0.52 [95% CI, 0.32–0.72]; $P = .82$), or iCa levels (AUC = 0.62 [95% CI, 0.44–0.80]; $P = .23$) in Fig 5. The cutoff values for CBF and CMRO₂ were 12.75 mL/100 g/min and 22.63 μ mol/100 g/min, respectively. The sensitivities with 95% CIs of the combined values together, CBF, and CMRO₂ were 85.1% (60.8%–93.3%), 80.5% (53.6%–86.9%), and 79.6% (53.3%–86.9%), respectively. Specificities with 95% CIs were 80.0% (75.0%–99.0%), 76.7% (61.1%–84.5%), and 73.3% (69.2%–81.2%), respectively.

DISCUSSION

In general, patients with NHC with structural abnormalities had impaired hemodynamics and cerebral metabolism. CBF and CMRO₂ were correlated with iCa levels, and CBF may be useful in assessing the need for calcium supplementation in asymptomatic NHC to prevent brain injury.

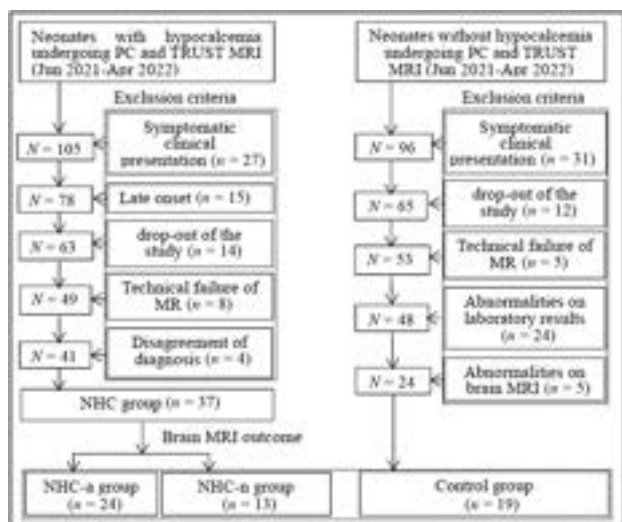


FIG 2. Flow chart of study participants.

Table 2: MR imaging findings between the NHC group and the NHC-a subset

Characteristics of Neonatal MR Imaging	NHC (n = 37)	NHC-a (n = 24)
Periventricular WM lesions I (No.) (%)	6 (16.2)	6 (25)
Periventricular WM lesions II		
+0 (No.) (%)	4 (10.8)	4 (16.7)
+ Intraventricular hemorrhage (No.) (%)	1 (2.7)	1 (4.2)
Periventricular white matter lesions III (No.) (%)	1 (2.7)	1 (4.2)
Periventricular white matter lesions IV (No.) (%)	7 (18.9)	7 (29.2)
SAH without parenchymal hemorrhagic lesions (No.) (%)	5 (13.5)	5 (20.8)

Calcium Levels in the Neonate

Serum calcium levels in neonates are high at birth because calcium readily crosses the placenta to the fetus via active transport during the third trimester. After delivery, the maternal transfer of calcium halts and neonates may experience NHC during the first few days of

life. Although NHC is frequently asymptomatic, this condition can be potentially life-threatening.² Treatment is controversial, and evidence of its benefits is lacking.⁵

Effect of NHC on Physiologic Parameters and Structural Damage

Within the brain microcirculation, parenchymal arterioles are fundamental regulators of CBF.^{27,28} The tone of parenchymal arterioles within the brain is controlled by astrocytes, which synthesize and release prostaglandins and epoxyeicosatrienoic acid via cyclooxygenase-1 and epoxygenase P450, respectively, in a Ca^{2+} -dependent manner. Ca^{2+} is not only indispensable for physiologic activities but is also an important second messenger in cells. The resulting increase in astrocytic Ca^{2+} leads to the synthesis and release of these vasodilators, which cause the dilation of brain arterioles and increase CBF. In addition, Ca^{2+} inhibits the inward current of Na^+ . When

hypocalcemia occurs, the inward current of Na^+ increases, the threshold of the action potential decreases, and depolarization occurs. Subsequently, the voltage-dependent Ca^{2+} channel is activated, and calcium concentration increases, leading to a contraction of vascular smooth-muscle cells and reduction in CBF. This result may explain why the NHC group had poorer hemodynamics.

In addition, Ca^{2+} within the presynaptic terminal is key for chemical synaptic transmission. When intercellular Ca^{2+} is reduced, neurotransmitter release is inhibited. A lower calcium level leads to diminished neural activity, thus lower oxygen need and lower CMRO_2 .

However, this impairment in hemodynamics was more significant in the NHC-a subset but not in the NHC-n subset, compared with controls. This finding might be related to the autoregulation of CBF. CBF regulation in the neonatal brain is distinct from that of the fully matured brain. The preterm rate is relatively high in

this cohort (75.6% in NHC and 63.2% in control group). Infants born prematurely have an anatomically incomplete and underdeveloped cerebral vasculature and cannot fully autoregulate.²⁹ Once their cerebrovascular autoregulation cannot fully compensate for the effect of hypocalcemia, the probability of brain injuries increases (Online Supplemental Data).

The NHC-a group with brain injury mainly included patients with SAH and PVWM lesions. During activation of the clotting mechanism after SAH, coagulation factor IV (Ca^{2+}) is consumed. When SAH or PVWM lesions are accompanied by cerebral edema, Ca^{2+} will flow into the cell, causing a marked overload of intracellular Ca^{2+} , further promoting hypocalcemia.

Although it is possible that other factors may contribute to both impaired hemodynamics and structural lesions, we were able to control for most of the confounding factors by excluding participants with severe diseases and including age, sex, birth weight, Apgar

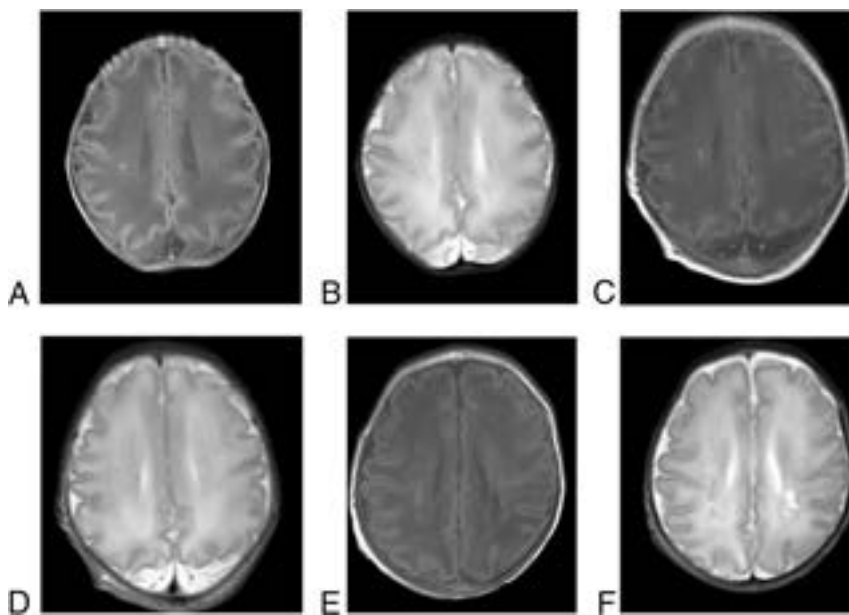


FIG 3. NHC-a subgroup with brain injuries on T1WI and T2WI. A and B, Neonate (scan age, 33.3 weeks; birth weight, 1740 g; male) had focal PVWM lesions (grade I) in the right PVWM area showing increased signal intensity on T1WI and decreased signal intensity on T2WI. C and D, Neonate (scan age, 34.3 weeks; birth weight, 2600 g; male) had PVWM lesions (grade II) in the bilateral PVWM areas, linearly displaying increased signal intensity on T1WI and decreased signal intensity on T2WI. E and F, Neonate (scan age, 37.7 weeks; birth weight, 2200 g; female) had clustered PVWM lesions and cystic lesions (grade IV) in bilateral PVWM areas.

Table 3: Comparison of physiologic parameters between groups and their subgroups (n = 56)^a

Comparators	Yv (%)	CBF (mL/100g/min)	CMRO ₂ (mL/100g/min)	Brain Volume (mL)
NHC group (n = 37)	63.4 (SD, 13.6)	13.3 (SD, 5.3)	31.1 (SD, 12.2)	300.0 (SD, 88.6)
Control group (n = 19)	64.7 (SD, 7.1)	16.5 (SD, 4.7)	40.4 (SD, 14.8)	290.4 (SD, 72.8)
P value	.64	<.05	<.05	.69
NHC-a group (n = 24)	63.2 (SD, 16.1)	11.2 (SD, 4.5)	27.4 (SD, 10.7)	296.9 (SD, 87.3)
NHC-n group (n = 13)	63.8 (SD, 7.5)	17.1 (SD, 4.7)	37.9 (SD, 12.1)	305.6 (SD, 94.2)
P value	.88	<.05	<.05	.78
NHC-a group (n = 24)	63.2 (SD, 16.1)	11.2 (SD, 4.5)	27.4 (SD, 10.7)	296.9 (SD, 87.3)
Control group (n = 19)	64.7 (SD, 7.1)	16.5 (SD, 4.7)	40.4 (SD, 14.8)	290.4 (SD, 72.8)
P value	.70	<.05	<.05	.80
NHC-n group (n = 13)	63.8 (SD, 7.5)	17.1 (SD, 4.7)	37.9 (SD, 12.1)	305.6 (SD, 94.2)
Control group (n = 19)	64.7 (SD, 7.1)	16.5 (SD, 4.7)	40.4 (SD, 14.8)	290.4 (SD, 72.8)
P value	.73	.70	.63	.61

^a Data are expressed as means. P values are based on the Wilcoxon 2-sample exact test.

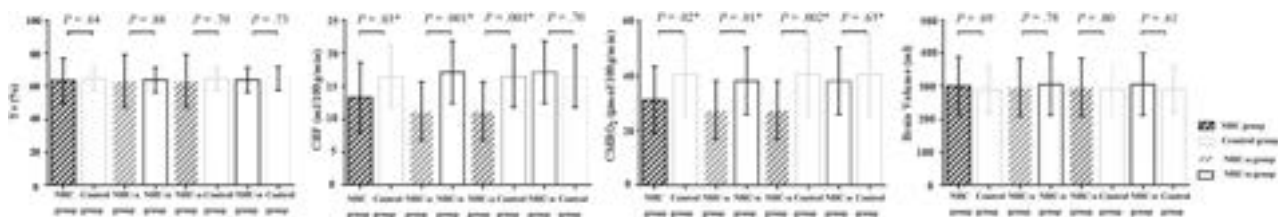


FIG 4. Boxplots of the physiologic parameters in both groups and subgroups. A–D, Boxplots display Yv, CBF, CMRO₂, and brain volume from both groups and subgroups. NHC neonates and the NHC-a subset (versus controls) both had significantly lower CBF ($P = .03$, $P = .001$) and CMRO₂ ($P = .02$, $P = .002$). CBF and CMRO₂ were significantly lower in NHC-a (versus the NHC-n) ($P = .001$, $P = .01$). There was no significant difference in CBF and CMRO₂ between those with NHC-n and controls ($P = .70$, $P = .63$). Yv and brain volume were not significantly different between groups ($P > .05$). The asterisk indicates $P < .05$.

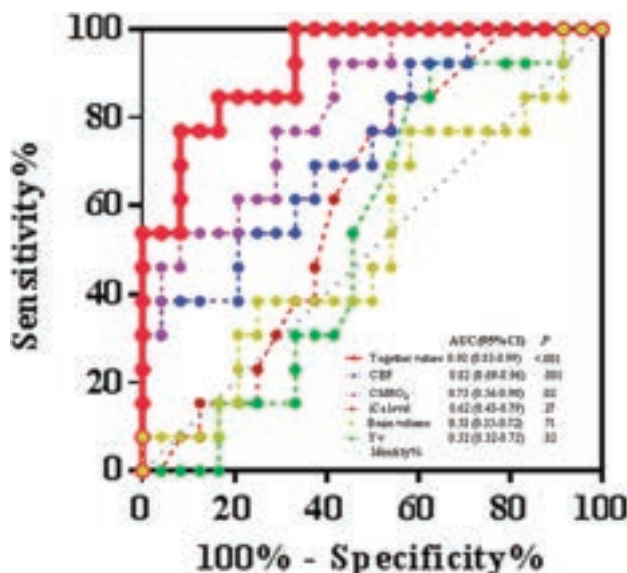


FIG 5. Receiver operating characteristic curve for the combined values together versus a single value (CBF, CMRO₂, brain volume, or iCa levels, respectively) to analyze the association between NHC and brain injuries. The combined values together versus a single value showed superior capacity to detect NHC with brain injuries (AUC = 0.92 [95% CI, 0.83-0.99]) versus CBF (AUC = 0.82 [95% CI, 0.69-0.96]; $P = .001$), CMRO₂ (AUC = 0.73 [95% CI, 0.56-0.90]; $P = .02$), brain volume (AUC = 0.53 [95% CI, 0.35-0.72]; $P = .82$), Yv (AUC = 0.52 [95% CI, 0.32-0.72]; $P = .82$), or iCa levels (AUC = 0.62 [95% CI, 0.44-0.80]; $P = .23$).

score, and blood gas analysis as covariates in multivariate regression analysis. Further studies with larger sample sizes and longitudinal follow-up are needed to better elucidate the causal relationship among calcium level, brain hemodynamics, and structure change.

Comparison with Previous Studies

Our measurements of Yv, CBF, and CMRO₂ in healthy neonates were comparable with those in previous reports,²¹ demonstrating the accuracy and reliability of PCMR and TRUST MR imaging for assessing neonatal cerebral hemodynamics and metabolism.

There were limited reports on the cerebral hemodynamics and oxygen metabolism in NHC. CBF and CMRO₂ showed significant increases with iCa levels. The OR of NHC with brain injuries was 0.80 (95% CI, 0.65–0.99) for CBF. Thus, CBF may be a protective factor for asymptomatic NHC with brain injury.

When a neonate with hypocalcemia has CBF of <12.75 mL/100 g/min (cutoff value), calcium supplementation might be considered to reduce the risk of brain injury.

Limitations

Our study has a moderate sample size. In future research, we will broaden the scope of our study to include follow-up of patients with NHC to clarify the impact of decreased CBF, CMRO₂, and iCa on neurocognitive deficits. In addition, SAH is negligible secondary to natural delivery. However, Wang et al³⁰ and Peng et al³¹ reported that the myelin sheath was shown to be rapidly damaged and slowly recovered after SAH. Future studies for measurement of whether there are hemodynamic and metabolic abnormalities in neonates with SAH without parenchymal hemorrhagic lesions might be performed. There was 1 neonate with NRDS in grade III diagnosed by x-ray in our study who was administered a surfactant within 2 hours after birth. The administration of surfactant is performed by liquid bolus instillation into the endotracheal tube. Transient airway obstruction has been associated with rapid changes in oxygen saturation, heart rate, systemic blood pressure, and CBF. However, Terry et al³² reported that CBF did not change significantly after surfactant administration. CBF that was not affected by fluctuations in arterial blood pressure during surfactant administration may indicate active autoregulation of flow by the cerebral vasculature. CBF remained unchanged until the mean arterial pressures fell below 35 mm Hg. In addition, MR dose not scan until neonate is stable. In our only case with NRDS grade III, an MR imaging examination was performed 7 days after birth, at which time the surfactant had little influence on CBF.

CONCLUSIONS

NHC with structural damage is likely to affect neonates with lower cerebral metabolism and hemodynamics. CBF may be useful in assessing the need for calcium supplementation in patients with asymptomatic NHC to prevent brain injury.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

1. Moss CR. Neonatal hypocalcemia in the infant of a diabetic mother. *Neonatal Netw* 2020;39:200–04 CrossRef Medline

2. Cho WI, Yu HW, Chung HR, et al. **Clinical and laboratory characteristics of neonatal hypocalcemia.** *Ann Pediatr Endocrinol Metab* 2015;20:86–91 CrossRef Medline
3. Yilmaz B, Aygun C, Cetinoglu E. **Vitamin D levels in newborns and association with neonatal hypocalcemia.** *J Matern Fetal Neonatal Med* 2018;31:1889–93 CrossRef Medline
4. Levy-Shraga Y, Dallalzadeh K, Stern K, et al. **The many etiologies of neonatal hypocalcemic seizures.** *Pediatr Emerg Care* 2015;31:197–201 CrossRef Medline
5. Perino JM. **Calcium levels in the neonate.** *Neonatal Netw* 2020;39:35–39 CrossRef Medline
6. Kutilek S, Vracovska M, Pecenkova K, et al. **Calcemia and inflammatory markers in early-onset neonatal infection.** *Acta Medica (Hradec Kralove)* 2019;62:58–61 CrossRef Medline
7. Liu Y, Chai Y, Rong Z, et al. **Prognostic value of ionized calcium levels in neonatal sepsis.** *Ann Nutr Metab* 2020;76:193–200 CrossRef Medline
8. Wintermark M, Lepori D, Cotting J, et al. **Brain perfusion in children: evolution with age assessed by quantitative perfusion computed tomography.** *Pediatrics* 2004;113:1642–52 CrossRef Medline
9. Koziak AM, Winter J, Lee TY, et al. **Validation study of a pulsed arterial spin labeling technique by comparison to perfusion computed tomography.** *Magn Reson Imaging* 2008;26:543–53 CrossRef Medline
10. Wright EA, d'Esterre CD, Morrison LB, et al. **Absolute cerebral blood flow infarction threshold for 3-hour ischemia time determined with CT perfusion and 18F-FFMZ-PET imaging in a porcine model of cerebral ischemia.** *PLoS One* 2016;11:e0158157 CrossRef Medline
11. Pryds O, Greisen G, Skov LL, et al. **Carbon dioxide-related changes in cerebral blood volume and cerebral blood flow in mechanically ventilated preterm neonates: comparison of near infrared spectrophotometry and ¹³³xenon clearance.** *Pediatr Res* 1990;27:445–49 CrossRef Medline
12. Dallery F, Bouzerar R, Michel D, et al. **Perfusion magnetic resonance imaging in pediatric brain tumors.** *Neuroradiology* 2017;59:1143–53 CrossRef Medline
13. Diop M, Kishimoto J, Toronov V, et al. **Development of a combined broadband near-infrared and diffusion correlation system for monitoring cerebral blood flow and oxidative metabolism in preterm infants.** *Biomed Opt Express* 2015;6:3907–18 CrossRef Medline
14. Tortora D, Severino M, Rossi A. **Arterial spin labeling perfusion in neonates.** *Semin Fetal Neonatal Med* 2020;25:101130 CrossRef Medline
15. Ouyang M, Liu P, Jeon T, et al. **Heterogeneous increases of regional cerebral blood flow during preterm brain development: preliminary assessment with pseudo-continuous arterial spin labeled perfusion MRI.** *Neuroimage* 2017;147:233–42 CrossRef Medline
16. Liu P, Qi Y, Lin Z, et al. **Assessment of cerebral blood flow in neonates and infants: a phase-contrast MRI study.** *Neuroimage* 2019;185:926–33 CrossRef Medline
17. Ye FQ, Berman KF, Ellmore T, et al. **H(2)(15)O PET validation of steady-state arterial spin tagging cerebral blood flow measurements in humans.** *Magn Reson Med* 2000;44:450–56 CrossRef Medline
18. Armitage PA, Skipper N, Connolly DJ, et al. **A qualitative comparison of arterial spin labelling and dynamic susceptibility contrast MRI in 52 children with a range of neurological conditions.** *Br J Radiol* 2017;90:20160495 CrossRef Medline
19. van Osch MJ, Teeuwisse WM, van Walderveen MA, et al. **Can arterial spin labeling detect white matter perfusion signal?** *Magn Reson Med* 2009;62:165–73 CrossRef Medline
20. Qi Y, Liu P, Lin Z, et al. **Hemodynamic and metabolic assessment of neonates with punctate white matter lesions using phase-contrast MRI and T2-relaxation-under-spin-tagging (TRUST) MRI.** *Front Physiol* 2018;9:233 CrossRef Medline
21. Liu P, Huang H, Rollins N, et al. **Quantitative assessment of global cerebral metabolic rate of oxygen (CMRO2) in neonates using MRI.** *NMR Biomed* 2014;27:332–40 CrossRef Medline
22. Shetty AN, Lucke AM, Liu P, et al. **Cerebral oxygen metabolism during and after therapeutic hypothermia in neonatal hypoxic-ischemic encephalopathy: a feasibility study using magnetic resonance imaging.** *Pediatr Radiol* 2019;49:224–33 CrossRef Medline
23. Qi Y, Wang X, Mao J. **Quantitative assessment of cerebral metabolism and hemodynamics in small-for-gestational-age (SGA) newborns.** *Quant Imaging Med Surg* 2021;11:2321–32 CrossRef Medline
24. Wintermark P, Hansen A, Warfield SK, et al. **Near-infrared spectroscopy versus magnetic resonance imaging to study brain perfusion in newborns with hypoxic-ischemic encephalopathy treated with hypothermia.** *Neuroimage* 2014;85(Pt 1):287–93 CrossRef Medline
25. Childs AM, Cornette L, Ramenghi LA, et al. **Magnetic resonance and cranial ultrasound characteristics of periventricular white matter abnormalities in newborn infants.** *Clin Radiol* 2001;56:647–55 CrossRef Medline
26. Herscovitch P, Mintun MA, Raichle ME. **Brain oxygen utilization measured with oxygen-15 radiotracers and positron emission tomography: generation of metabolic images.** *J Nucl Med* 1985;26:416–17 Medline
27. Elsary AY, Elgameel AA, Mohammed WS, et al. **Neonatal hypocalcemia and its relation to vitamin D and calcium supplementation.** *Saudi Med J* 2018;39:247–53 CrossRef Medline
28. Dabertrand F, Nelson MT, Brayden JE. **Ryanodine receptors, calcium signaling, and regulation of vascular tone in the cerebral parenchymal microcirculation.** *Microcirculation* 2013;20:307–16 CrossRef Medline
29. Rhee CJ, da Costa CS, Austin T, et al. **Neonatal cerebrovascular autoregulation.** *Pediatr Res* 2018;84:602–10 CrossRef Medline
30. Wang Y, Xu J, You W, et al. **Roles of Rufy3 in experimental subarachnoid hemorrhage-induced early brain injury via accelerating neuronal axon repair and synaptic plasticity.** *Mol Brain* 2022;15:35 CrossRef Medline
31. Peng J, Pang J, Huang L, et al. **LRP1 activation attenuates white matter injury by modulating microglial polarization through Shc1/PI3K/Akt pathway after subarachnoid hemorrhage in rats.** *Redox Biol* 2019;21:101121 CrossRef Medline
32. Terry MH, Merritt TA, Harding B, et al. **Pulmonary distribution of lucinactant and poractant alfa and their peridosing hemodynamic effects in a preterm lamb model of respiratory distress syndrome.** *Pediatr Res* 2010;68:193–98 CrossRef Medline

Neuroimaging Findings in Axenfeld-Rieger Syndrome: A Case Series

Samuel White, Ajay Taranath, Prasad Hanagandi, Deepa A. Taranath, Minh-Son To, Emmanuelle Souzeau, Owen M. Siggs, and Jamie E. Craig



ABSTRACT

SUMMARY: Axenfeld-Rieger syndrome is an autosomal dominant condition associated with multisystemic features including developmental anomalies of the anterior segment of the eye. Single nucleotide and copy number variants in the paired-like homeodomain transcription factor 2 (*PITX2*) and forkhead box C1 (*FOXC1*) genes are associated with Axenfeld-Rieger syndrome as well as other CNS malformations. We determined the association between Axenfeld-Rieger syndrome and specific brain MR imaging neuroradiologic anomalies in cases with or without a genetic diagnosis. This case series included 8 individuals with pathogenic variants in *FOXC1*; 2, in *PITX2*; and 2 without a genetic diagnosis. The most common observation was vertebrobasilar artery dolichoectasia, with 46% prevalence. Other prevalent abnormalities included WM hyperintensities, cerebellar hypoplasia, and ventriculomegaly. Vertebrobasilar artery dolichoectasia and absent/hypoplastic olfactory bulbs were reported in >50% of individuals with *FOXC1* variants compared with 0% of *PITX2* variants. Notwithstanding the small sample size, neuroimaging abnormalities were more prevalent in individuals with *FOXC1* variants compared those with *PITX2* variants.

ABBREVIATION: ARS = Axenfeld-Rieger syndrome

Axtenfeld-Rieger syndrome (ARS) is an autosomal dominant genetic condition associated with multisystemic features. It is characterized primarily by ocular features that result from developmental anomalies of the anterior segment of the eye, including posterior embryotoxon (a thickened and anteriorly displaced Schwalbe ring), iris hypoplasia, corectopia (displaced pupil), pseudopolycoria (additional pupillary opening), and iridocorneal adhesions.^{1,2} The developmental anomalies of the structures allowing drainage of the aqueous humor lead to an increased risk of secondary glaucoma. Commonly reported systemic features include facial dysmorphism; dental, umbilical, cardiovascular,

and endocrinological anomalies; hearing impairment; and developmental delay.^{2,3}

ARS has been associated with variants in the paired-like homeodomain transcription factor 2 (*PITX2*) and forkhead box C1 (*FOXC1*) genes.⁴⁻⁶ *PITX2* belongs to the homeobox gene family and is fundamental to the embryonic development of several tissues, including an essential role in left-right patterning.^{7,8} *FOXC1* encodes a forkhead family transcription factor and is also involved in embryonic development.^{9,10} Pathogenic and likely pathogenic variants in *FOXC1* and *PITX2* have been reported in ~40% of individuals with a clinical diagnosis of ARS,^{11,12} with unique ocular and systemic phenotypes associated with each gene.¹³⁻¹⁶

Variants in both *FOXC1* and *PITX2* have also been associated with a range of CNS malformations, including hydrocephalus,¹⁶⁻¹⁸ classic commissural agenesis (corpus callosum agenesis),¹⁷⁻¹⁹ and cerebellar malformations (Dandy-Walker phenotype,²⁰ mega cisterna magna, and cerebellar vermis hypoplasia).^{18,21,22} More recently, cerebral small-vessel disease has been reported in individuals with *FOXC1* or *PITX2* variants, with the presence of WM hyperintensities, dilated perivascular spaces, and lacunar infarcts on MR imaging.^{23,24}

Here, using a series of ARS cases with accompanying MR imaging of the brain, we systematically determined the association between ARS and specific neuroradiologic anomalies in cases with or without a genomic diagnosis.

Received February 2, 2023; accepted after revision August 16.

From the Robinson Research Institute (S.W.), Faculty of Medicine and Health Sciences, University of Adelaide, Adelaide, South Australia, Australia; Department of Radiology (A.T.), Women's and Children's Hospital, Adelaide, South Australia, Australia; Department of Neuroradiology (P.H.), King Abdulaziz Medical City, Ministry of National Guard Health Affairs, Riyadh, Saudi Arabia; Department of Ophthalmology (D.A.T., M.-S.T., E.S., O.M.S., J.E.C.), Flinders University, Bedford Park, South Australia, Australia; and Garvan Institute of Medical Research (O.M.S.), Darlinghurst, New South Wales, Australia.

This work was supported by the Australian National Health and Medical Research Council Centres of Excellence Research Grant (APPI16360). J.E.C. was supported by an Australian National Health and Medical Research Council Practitioner Fellowship (APPI154824). E.S. was supported by an Early Career Fellowship from the Hospital Research Foundation, and O.M.S. was supported by a Snow Fellowship.

Please address correspondence to Samuel White, MD, Robinson Research Institute, Ground Floor, Norwich Centre, 55 King William Str, North Adelaide, South Australia 5006, Australia; e-mail: s.white@adelaide.edu.au

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

<http://dx.doi.org/10.3174/ajnr.A7995>

Table 1: Cohort demographics^a

	All	FOXC1 Variant	PITX2 Variant	No Genetic Diagnosis
Prevalence in overall cohort (No.) (%)	NA	8 (61.5)	2 (16.7)	2 (16.7)
Sex, female (No.) (%)	9 (75)	6 (75.0)	2 (100)	1 (50)
European ancestry (No.) (%)	100 (100)	8 (100)	2 (100)	2 (100)
Age (yr)				
Mean	37.3 (SD, 21.1)	29.2 (SD, 16.1)	51.8 (SD, 3.04)	10.6 (SD, 14.6)
Range	8–73	8–49	48–54	0–31

Note:—NA indicates not applicable.

^aBoth sex and ancestry were self-reported.

Table 2: Mean globe parameters

Parameter	Mean Length (mm)				General Population ^{27,28}
	All (n = 12)	FOXC1 Variant (n = 8)	PITX2 Variant (n = 2)	No Genetic Diagnosis (n = 2)	
Anterior-posterior diameter of globe	22.1	22.4	24.0	20.1	24.2
Transverse diameter of globe	22.9	23.5	24.0	20.7	24.2
Anterior chamber depth	2.37	2.62	1.89	1.51	2.62

MATERIALS AND METHODS

Subjects, Genetic Testing, and Neuroimaging

The study was conducted in accordance with the revised Declaration of Helsinki. Ethics approval was obtained from the Southern Adelaide Clinical Research Ethics Committee, and all participants or their caregivers provided written informed consent. Individuals with a clinical diagnosis of ARS were drawn from the Australian and New Zealand Registry of Advanced Glaucoma as previously described.²⁵ *FOXC1* and *PITX2* genetic testing was performed in a National Association of Testing Authorities–accredited laboratory by Sanger sequencing or multiplex ligation-dependent probe amplification as previously described.^{16,26} Brain MRIs were assessed by 2 pediatric neuroradiologists (A.T. and P.H.) blinded to the genetic results of each participant.

RESULTS

Twelve individuals with ARS were included. The mean age at evaluation was 37.3 (SD, 21.2) years (range, 2 months–73 years), 77% (10/13) were female, and all were of European ancestry (Table 1). Eight individuals had heterozygous pathogenic or likely pathogenic variants in *FOXC1* (including 7 with sequence variants and 1 with a full gene deletion), 2 had heterozygous pathogenic or likely pathogenic variants in *PITX2*, and 2 had no genetic diagnosis despite testing.

Globe and Optic Chiasm

Mean globe parameters are outlined in Table 2 and are compared with ocular biometry from the general population.^{27,28} A thin optic chiasm was reported in 42% (5/12) of the cohort. Both individuals with *PITX2* variants had optic chiasm thinning, whereas only 38% (3/8) of those with *FOXC1* variants had a thin optic chiasm (Fig 1A and Table 3).

Cortex

Five subjects had nonspecific WM hyperintensities. Other WM changes included reduced WM volume ($n = 1$) and delayed

myelination ($n = 1$). Prominent perivascular spaces were noted in 3 individuals. Four patients had corpus callosal thinning. Three individuals had colpocephaly/ventriculomegaly, and another had a ventriculoperitoneal shunt. Other corpus callosal abnormalities included thickening of the splenium ($n = 1$) and genu ($n = 1$). The prevalence of prominent perivascular spaces in the *FOXC1* variant group was 38% (3/8). Corpus callosum thinning was observed in 38% (3/8) of *FOXC1* variants (Fig 1B). Thirty-eight percent (3/8) of individuals with *FOXC1* variants had ventriculomegaly or a ventriculoperitoneal shunt in situ, whereas none (0/2) of the *PITX2* variant group had a ventricular abnormality.

Cerebellum

Hemispheric or global cerebellar hypoplasia was reported in 42% (5/12) of subjects (Fig 1C). Superior vermis hypoplasia was identified in 1 patient, and inferior vermis hypoplasia, in another. Other cerebellar findings included tonsillar ectopia (defined as inferior tonsillar location 3–5 mm below the plane of foramen magnum) ($n = 1$) and mega cisterna magna (defined as distance from the posterior aspect of the cerebellar vermis to the inside of the occipital bone of > 10 mm) ($n = 1$) (Fig 1C). One of the 2 individuals with *PITX2* variants had superior vermis hypoplasia, and 38% (3/8) of the *FOXC1* variant group had global or hemispheric cerebellar hypoplasia.

Brainstem

An oblong pons was observed in 3 patients. Three subjects had brainstem indentation secondary to tortuous vertebral arteries. Another individual had medullary elongation. Brainstem indentation secondary to tortuous vertebral arteries was observed in 38% (3/8) of individuals with *FOXC1* variants. No brainstem abnormalities were reported in the *PITX2* variant group.

Vessels

Twenty-five percent (3/12) of individuals had circle of Willis abnormalities on MRA. Vertebrobasilar artery dolichoectasia was reported in 6 patients (Fig 1D). Four of these 6 individuals also had anterior circulation dolichoectasia. All 3 patients with circle

of Willis abnormalities on MRA had *FOXC1* variants, accounting for 38% of this group. Most (75%) of the *FOXC1* variant group had vertebrobasilar artery dolichoectasia. Of these patients, two-thirds also had anterior circulation dolichoectasia.

Other Findings

Absent or hypoplastic olfactory bulbs were reported in 50% (6/12) of subjects (Fig 1E). Other findings included cochlear nerve

hypoplasia ($n = 1$), a widened opercula ($n = 1$), hippocampal malrotation ($n = 1$), and bilateral absent posterior semicircular canals ($n = 1$). Absent or hypoplastic olfactory bulbs were reported in 63% (5/8) of subjects with *FOXC1* variants but in none (0/2) of the subjects with *PITX2* variants. Except for 1 subject with right unicoronal craniosynostosis, most of the cohort did not have craniofacial dysmorphism. The pituitary gland and hypothalamus were normal across the cohort. There was no abnormality of the deep gray nuclei. No dental abnormalities were observed.

DISCUSSION

This case series reviewed the neuroradiologic features of 12 individuals diagnosed with ARS, comprising 8 individuals with *FOXC1* variants, 2 with *PITX2* variants, and 2 with unsolved genetic defects. No single anatomic abnormality was observed in most individuals. The most common observation was vertebrobasilar artery dolichoectasia (50% prevalence), which was associated with anterior circulation dolichoectasia in most cases. Other prevalent abnormalities included WM hyperintensities (42%), hemispheric or global cerebellar hypoplasia (42%), corpus callosal thinning (33%), and ventriculomegaly (25%). Optic chiasm thinning was observed in both members of the *PITX2* variant group and in 38% of the *FOXC1* variant group. Vertebrobasilar artery dolichoectasia was reported in 75% of individuals with *FOXC1* variants compared with 0% of individuals with *PITX2* variants. Similarly, while 63% of individuals with *FOXC1* variants had absent or hypoplastic olfactory bulbs, they were not observed in any of the individuals with *PITX2* variants. Circle of Willis abnormalities on MRA and ventricular abnormalities both had a prevalence of 38% in the *FOXC1* group compared with 0% in the *PITX2* group.

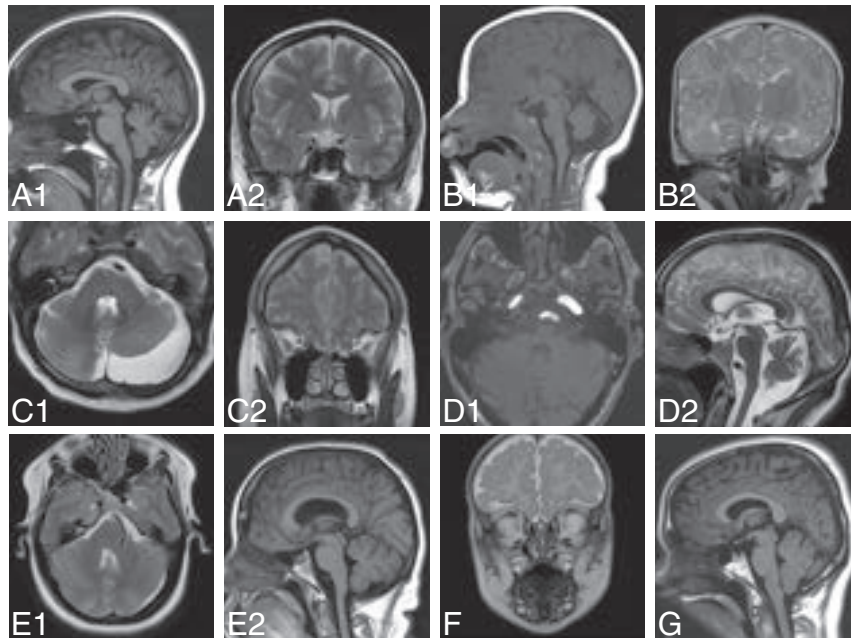


FIGURE. A, A 49-year-old woman with a *PITX2* mutation had superior vermis volume loss on the T1-weighted sagittal image (A1) and optic chiasm thinning on the T2-weighted coronal image (A2). B, A 2-month-old girl with an unsolved mutation had inferior vermis hypoplasia and a widened tectovermian angle on the T1-weighted sagittal image (B1) and malrotated hippocampi on the T2-weighted coronal image (B2). C, A 43-year-old man with a *FOXC1* mutation had mega cisterna magna and an ectatic basilar artery and a hypoplastic left cerebellar hemisphere on the T2-weighted axial image (C1) and hypoplastic olfactory bulbs on the T2-weighted coronal image (C2). D, A 46-year-old man with a *FOXC1* mutation had a tortuous basilar artery and an ectatic cavernous segment of the left ICA on axial TOF angiography (D1); a short mesencephalon with loss of the normal relationship among the mesencephalon, pons, and medulla; loss of volume in the superior vermis and ectatic basilar artery; and flow void seen end-on on the T2-weighted sagittal image (D2). E, A 41-year-old woman with a *FOXC1* mutation had a tortuous basilar artery flow void on the T2-weighted axial image (E1), a short mesencephalon with loss of the normal relationship among the mesencephalon, pons, and medulla, and bowing of the corpus callosum secondary to ventriculomegaly on the T1-weighted sagittal image (E2). F, A 3-month-old boy with an unsolved mutation had absent olfactory bulbs on the T2-weighted coronal image. G, A 31-year-old woman with an unsolved mutation had a thickened splenium and tonsillar ectopia on the T1-weighted sagittal image.

Table 3: Prevalence of neuroradiologic anomalies^a

(N (%))	All	<i>FOXC1</i>	<i>PITX2</i>	No Genetic Diagnosis
Thin optic chiasm	5/12 (41.7)	3/8 (37.5)	2/2 (100)	0/2 (0.0)
Nonspecific WM hyperintensities	5/12 (41.7)	4/8 (50.0)	1/2 (50.0)	0/2 (0.0)
Corpus callosal thinning	4/12 (33.3)	3/8 (37.5)	0/2 (0.0)	1/2 (50.0)
Ventriculomegaly or ventriculoperitoneal shunt	4/12 (33.3)	3/8 (37.5)	0/2 (0.0)	0/2 (0.0)
Hemispheric or global cerebellar hypoplasia	5/12 (41.7)	3/8 (37.5)	1/2 (50.0)	0/2 (0.0)
Oblong pons	3/12 (25.0)	2/8 (25.0)	0/2 (0.0)	1/2 (50.0)
Brainstem indentation secondary to tortuous vessels	3/12 (25.0)	3/8 (37.5)	0/2 (0.0)	0/2 (0.0)
Circle of Willis abnormalities	3/12 (25.0)	3/8 (37.5)	0/2 (0.0)	0/2 (0.0)
Vertebrobasilar artery dolichoectasia	6/12 (50.0)	6/8 (75.0)	0/2 (0.0)	0/2 (0.0)
Absent/hypoplastic olfactory bulb	6/12 (50.0)	5/8 (62.5)	0/2 (0.0)	1/2 (50.0)

^a Data are (No.) (%).

Mean globe parameters were smaller in the cohort without a genetic diagnosis. Although we acknowledge the very small sample size, this group might have an underlying genetic etiology that also impacts globe development, though this remains to be validated. No individuals with *FOXC1* or *PITX2* variants had craniofacial dysmorphism or abnormalities of the hypothalamus, pituitary, deep gray nuclei, or dentition.

The specific endocrinologic manifestations of ARS have not been comprehensively reported in the literature. Santini et al²⁹ have described a patient with growth hormone deficiency associated with ARS. Notably, growth hormone deficiency is a prevalent feature among individuals with septo-optic dysplasia,³⁰ which, like ARS, often involves olfactory bulb–tract hypoplasia.³¹

While there have been several isolated ARS neuroimaging case studies reported,^{32–34} to our knowledge, there is only 1 case series (Reis et al³⁵) that characterized the genetic and phenotypic features of an ARS cohort comprising 128 individuals with *FOXC1* or *PITX2* variants, including 18 with neuroimaging. The authors observed WM hyperintensities in 94% of *FOXC1* variants and 50% of *PITX2* variants³⁵ (compared with 50% for both *FOXC1* and *PITX2* variant groups in our study). Seventy-one percent of individuals with *FOXC1* variants had colpocephaly/ventriculomegaly (compared with 25% in our study). Reis et al also reported a 31% prevalence of arachnoid cysts in the *FOXC1* variant group, which was not assessed in our case series. Most interesting, the same study observed no correlation between the extent of neuroimaging anomalies and the presence or severity of cognitive impairment in patients with *FOXC1* variants.

Due to the nature of the imaging technique used in this study (MR imaging of the brain), we were not able to reliably detect extracerebral abnormalities such as craniofacial dysmorphism and dental anomalies, which are more sensitively detected by physical examination and specific dental imaging modalities such as orthopantomogram. Reis et al³⁵ reported classic dental anomalies such as hypodontia/oligodontia and microdontia in 91% of individuals with *PITX2* variants, with similar anomalies reported in 100% (23/23) of an Australian *PITX2* cohort¹⁶ drawn from the same registry as the current study. In contrast, these classic dental anomalies were considerably less common among the *FOXC1* group, who had a tendency to present with more atypical anomalies such as enamel hypoplasia/frequent caries (16%) or dental crowding (16%).³⁵ In a similar vein, while craniofacial dysmorphism was not observed on MR imaging of the brain in any of the individuals included in our case series, features such as thin upper lip and maxillary hypoplasia were reported in 78% of individuals with *FOXC1* variants and 93% of individuals with *PITX2* variants in 1 previous study,³⁵ and in another study (which included all cases described here), hypertelorism/telecanthus and low-set ears were found to be more prevalent in those with *FOXC1* variants compared with those with *PITX2* variants.¹⁵

This case series was limited by its small cohort size ($n = 12$), particularly with respect to the *PITX2* variant group ($n = 2$), which like other ARS case series precluded statistical comparisons.³³

CONCLUSIONS

This study is novel in its description of the relative prevalence of neuroimaging findings among patients with *FOXC1* and *PITX2*

variants, and overall, we observed that the *FOXC1* variant group had a higher prevalence of most (70%) neuroimaging abnormalities assessed (Table 3). The most common observation was verte-brobasilar artery dolichoectasia, which was reported in 75% of individuals with *FOXC1* variants compared with 0% of individuals with *PITX2* variants.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

- Shields MB, Buckley E, Klintworth GK, et al. **Axenfeld-Rieger syndrome: a spectrum of developmental disorders.** *Surv Ophthalmol* 1985;29:387–409 CrossRef Medline
- Alward WL. **Axenfeld-Rieger syndrome in the age of molecular genetics.** *Am J Ophthalmol* 2000;130:107–15 CrossRef Medline
- Fitch N, Kaback M. **The Axenfeld syndrome and the Rieger syndrome.** *J Med Genet* 1978;15:30–34 CrossRef Medline
- Semina EV, Reiter R, Leysens NJ, et al. **Cloning and characterization of a novel bicoid-related homeobox transcription factor gene, RIEG, involved in Rieger syndrome.** *Nat Genet* 1996;14:392–99 CrossRef Medline
- Nishimura DY, Swiderski RE, Alward WL, et al. **The forkhead transcription factor gene FKHL7 is responsible for glaucoma phenotypes which map to 6p25.** *Nat Genet* 1998;19:140–47 CrossRef Medline
- Mears AJ, Jordan T, Mirzayans F, et al. **Mutations of the forkhead/winged-helix gene, FKHL7, in patients with Axenfeld-Rieger anomaly.** *Am J Hum Genet* 1998;63:1316–28 CrossRef Medline
- Kitamura K, Miura H, Miyagawa-Tomita S, et al. **Mouse Pitx2 deficiency leads to anomalies of the ventral body wall, heart, extra- and periocular mesoderm and right pulmonary isomerism.** *Development* 1999;126:5749–58 CrossRef Medline
- Lu MF, Pressman C, Dyer R, et al. **Function of Rieger syndrome gene in left-right asymmetry and craniofacial development.** *Nature* 1999;401:276–78 CrossRef Medline
- Kume T, Deng K, Hogan BL. **Murine forkhead/winged helix genes Foxc1 (Mf1) and Foxc2 (Mfh1) are required for the early organogenesis of the kidney and urinary tract.** *Development* 2000;127:1387–95 CrossRef Medline
- Kume T, Jiang H, Topczewska JM, et al. **The murine winged helix transcription factors, Foxc1 and Foxc2, are both required for cardiovascular development and somitogenesis.** *Genes Dev* 2001;15:2470–82 CrossRef Medline
- D'haene B, Meire F, Claerhout I, et al. **Expanding the spectrum of FOXC1 and PITX2 mutations and copy number changes in patients with anterior segment malformations.** *Invest Ophthalmol Vis Sci* 2011;52:324–33 CrossRef Medline
- Reis LM, Tyler RC, Volkmann Kloss BA, et al. **PITX2 and FOXC1 spectrum of mutations in ocular syndromes.** *Eur J Hum Genet* 2012;20:1224–33 CrossRef Medline
- Prem Senthil M, Knight LSW, Taranath D, et al. **Comparison of anterior segment abnormalities in individuals with FOXC1 and PITX2 variants.** *Cornea* 2022;41:1009–15 CrossRef Medline
- Knight LS, Ruddle JB, Taranath DA, et al. **Childhood and early onset glaucoma classification and genetic profile in a large Australasian disease registry.** *Ophthalmology* 2021;128:1549–60 CrossRef Medline
- Souzeau E, Siggs OM, Pasutto F, et al. **Gene-specific facial dysmorphism in Axenfeld-Rieger syndrome caused by FOXC1 and PITX2 variants.** *Am J Med Genet A* 2021;185:434–39 CrossRef Medline
- Souzeau E, Siggs OM, Zhou T, et al. **Glaucoma spectrum and age-related prevalence of individuals with FOXC1 and PITX2 variants.** *Eur J Hum Genet* 2017;25:839–47 CrossRef Medline
- Maclean K, Smith J, St Heaps L, et al. **Axenfeld-Rieger malformation and distinctive facial features: clues to a recognizable 6p25 microdeletion syndrome.** *Am J Med Genet A* 2005;132A:381–85 CrossRef Medline

18. Delahaye A, Khung-Savatovsky S, Aboura A, et al. **Pre- and postnatal phenotype of 6p25 deletions involving the FOXC1 gene.** *Am J Med Genet A* 2012;158A:2430–38 CrossRef Medline
19. Raybaud C. **The corpus callosum, the other great forebrain commissures, and the septum pellucidum: anatomy, development, and malformation.** *Neuroradiology* 2010;52:447–77 CrossRef Medline
20. Whitehead MT, Barkovich MJ, Sidpra J, et al. **Refining the neuroimaging definition of the Dandy-Walker phenotype.** *AJNR Am J Neuroradiol* 2022;43:1488–93 CrossRef Medline
21. Idrees F, Vaideanu D, Fraser SG, et al. **A review of anterior segment dysgeneses.** *Surv Ophthalmol* 2006;51:213–31 CrossRef Medline
22. Aldinger KA, Lehmann OJ, Hudgins L, et al. **FOXC1 is required for normal cerebellar development and is a major contributor to chromosome 6p25.3 Dandy-Walker malformation.** *Nat Genet* 2009;41:1037–42 CrossRef Medline
23. Cellini E, Disciglio V, Novara F, et al. **Periventricular heterotopia with white matter abnormalities associated with 6p25 deletion.** *Am J Med Genet A* 2012;158A:1793–97 CrossRef Medline
24. French CR, Seshadri S, Destefano AL, et al. **Mutation of FOXC1 and PITX2 induces cerebral small-vessel disease.** *J Clin Invest* 2014;124:4877–81 CrossRef Medline
25. Souzeau E, Goldberg I, Healey PR, et al. **Australian and New Zealand Registry of Advanced Glaucoma: methodology and recruitment.** *Clin Exp Ophthalmol* 2012;40:569–75 CrossRef Medline
26. Siggs OM, Souzeau E, Pasutto F, et al. **Prevalence of FOXC1 variants in individuals with a suspected diagnosis of primary congenital glaucoma.** *JAMA Ophthalmol* 2019;137:348–55 CrossRef Medline
27. Augusteyn RC, Nankivil D, Mohamed A, et al. **Human ocular biometry.** *Exp Eye Res* 2012;102:70–75 CrossRef Medline
28. Hashemi H, Khabazkhoob M, Miraftab M, et al. **The distribution of axial length, anterior chamber depth, lens thickness, and vitreous chamber depth in an adult population of Shahroud, Iran.** *BMC Ophthalmol* 2012;12:50 CrossRef Medline
29. Santini AJ, Canales Ramos NM, Burgos Ortega NI, et al. **MON-257 Axenfeld-Rieger syndrome: an uncommon cause of growth hormone deficiency.** *J Endocr Soc* 2020;4(Suppl 1):MON-257 CrossRef
30. Koizumi M, Ida S, Shoji Y, et al. **Endocrine status of patients with septo-optic dysplasia: fourteen Japanese cases.** *Clin Pediatr Endocrinol* 2017;26:89–98 CrossRef Medline
31. Benson JC, Nascene D, Truwit C, et al. **Septo-optic dysplasia: assessment of associated findings with special attention to the olfactory sulci and tracts.** *Clin Neuroradiol* 2019;29:505–13 CrossRef Medline
32. Whitehead M, Choudhri A, Salim S. **Magnetic resonance imaging findings in Axenfeld-Rieger syndrome.** *Clin Ophthalmol* 2013;7:911–16 CrossRef Medline
33. Kumar M, Chambers C, Dhamija R. **Axenfeld-Rieger syndrome and leukoencephalopathy caused by a mutation in FOXC1.** *Pediatr Neurol* 2017;66:113–14 CrossRef Medline
34. Annakie D, Gasimova U, Shafi M, et al. **Brain matter white changes in Axenfeld-Rieger syndrome: things to keep in mind.** *Turk J Neurol* 2022;28:124–26 CrossRef
35. Reis LM, Maheshwari M, Capasso J, et al. **Axenfeld-Rieger syndrome: more than meets the eye.** *J Med Genet* 2023;60:368–79 CrossRef Medline

Venous Sinus Stenosis with Prominent Emissary Veins: A New Common Cranial MRI Finding of Mucopolysaccharidosis I

Shiwei Huang, Ashish Gupta, Paul Orchard, Troy Lund, and David Nascene



ABSTRACT

SUMMARY: Mucopolysaccharidosis I-Hurler (MPSIH) syndrome is the most severe form of a group of hereditary lysosomal diseases. This study aims to describe previously unreported common cranial findings of sigmoid sinus stenosis with prominent emissary veins in MPSIH. A retrospective review was conducted of 66 patients with MPSIH who were treated at our institution. A total of 12 cranial MR imaging studies from 12 different patients demonstrating the venous sinus anatomy were reviewed. All 12 patients exhibited various degrees of sigmoid or transverse sinus stenosis. Eleven had various forms of emissary veins. Of those 12 patients with imaging of the venous sinuses, 9 had a lumbar puncture within the same months as the acquisition of the venogram without any correlation between elevated opening pressure and the severity of the venous sinus stenosis. Stenotic cerebral venous sinuses with associated emissary veins, common in patients with MPSIH, may be abnormal findings due to posterior fossa horns from glycosaminoglycan depositions rather than signs of elevated intracranial pressure or requirement of CSF diversion.

ABBREVIATIONS: GAG = glycosaminoglycan; HSCT = hematopoietic stem cell transplant; MPSIH = mucopolysaccharidosis I-Hurler; OP = opening pressure; VP = ventriculoperitoneal

Mucopolysaccharidosis type I, Hurler (MPSIH) syndrome is a severe, rare genetic lysosomal storage disease characterized by the accumulation of glycosaminoglycan (GAG) within the lysosomes due to α -L-iduronidase deficiency. Elevated GAG levels can be found in many tissues and body fluids, leading to progressive organ damage, cardiovascular complications, neurocognitive delay, and skeletal abnormalities.¹

The brain and cranial imaging manifestations of MPSIH have been well-studied and include WM signal-intensity abnormality, ventricular dilation, hydrocephalus, enlarged perivascular spaces, a J-shaped sella turcica, an enlarged pituitary gland, and posterior fossa horns (hypertrophy of the occipitomastoid sutures).²⁻⁵ Hydrocephalus is an early manifestation in patients with MPSIH, and prompt recognition and treatment can potentially prevent neurocognitive decline or even mortality.^{4,6} One common hypothesis is that skull abnormalities could lead to venous hypertension, which contributes to hydrocephalus in MPSIH. However, to our knowledge, no studies have evaluated the venous anatomy in

patients with MPSIH or its correlation with intracranial pressure.^{7,8} Studies in idiopathic intracranial hypertension (pseudotumor cerebri) have described the role of posterior fossa emissary veins as collateral venous outlets in transverse sinus stenosis that help to prevent or lessen intracranial pressures.^{9,10} In this study, we describe the high prevalence of posterior fossa horns, venous sinus stenosis, and prominent emissary veins in patients with MPSIH and examine their possible correlation with intracranial pressure.

MATERIALS AND METHODS

A total of 66 patients with MPSIH treated with hematopoietic stem cell transplant (HSCT) at our institution between 2008 and 2020 were retrospectively reviewed. Of those patients, 8 patients had dedicated MRVs and another 4 had high-resolution, contrast-enhanced 3D T1-weighted MPRAGE images with sufficient detail to allow evaluation of the venous sinuses. The venograms were obtained in accordance with our institution's existing stem cell transplant protocol. The contrast-enhanced MR images were obtained to assess infection, ophthalmologic pathology, and hormonal deficiencies. The radiologic assessments were performed by a resident and confirmed by an attending neuroradiologist for posterior fossa horns, venous stenosis, and emissary veins. Charts of the 12 patients were reviewed for concurrent CSF opening pressure (OP) measurements within 30 days of MR imaging. Lumbar punctures were performed with the patient in the lateral

Received June 7, 2023; accepted after revision August 16.

From the Department of Neurosurgery (S.H.), Division of Pediatric Blood and Marrow Transplant (A.G., P.O., T.L.), and Department of Radiology (D.N.), University of Minnesota, Minneapolis, Minnesota.

Please address correspondence to Shiwei Huang, MD, Department of Neurosurgery, University of Minnesota, 420 Delaware St SE MMC 96, Room D429 Mayo Building, Minneapolis, MN 55455; e-mail: huan2256@umn.edu



Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A7997>

decubitus position with maintained end-tidal CO₂ of 25–40 mm Hg.¹¹ Whether ventriculoperitoneal (VP) shunts were present was also assessed. Ten children without MPSIH who underwent MRV in 2020 were reviewed and included as a control population. This study was performed in accordance with the rules and regulations of Committee on the Use of Human Subjects in Research at the University of Minnesota (IRB STUDY00010613).

RESULTS

All 12 patients with sufficient imaging demonstrated some degree of venous sinus stenosis and posterior fossa horns, with an average age of 3.1 years at the time of the MRV (Online Supplemental Data). Eleven of the 12 patients had emissary veins or venous plexuses draining from either the sigmoid sinus, torcula, superior sagittal sinus, or jugular bulb (Fig 1). Nine of the 12 patients had an OP measured with a mean of 24 cm H₂O (range, 7–37 cm H₂O). Three of those 9 patients had an OP of >28 cm H₂O. Two of the 12 patients had a VP shunt placed 4 years after HSCT due to the discovery of a papilledema refractory to medical therapy but with normal OP and ventricular size, and their contrast-enhancing MR images were obtained 5 years after VP shunt placement to assess a hormonal deficiency (9 years after HSCT). On the basis of MR imaging, 1 patient was found to have acute hydrocephalus 22 days after HSCT, for which a VP shunt was eventually placed, and this patient's MRV was obtained before HSCT as part of the pretransplantation work-up. The average age at the time of MRV was 2.1 years for the 10 controls. Three of them underwent MRV for a work-up of intracerebral hemorrhage; 2, for seizure work-up; 2, for visual disturbance work-up; and 3, for surveillance of known vascular abnormalities (Online Supplemental Data). None of the patients in the control group had intracranial hypertension or shunts. Only 2 had some degree of posterior fossa horns, and only 1 had emissary veins (Fig 2).

DISCUSSION

MPSIH is associated with varying degrees of CNS involvement and a resultant impact on neurodevelopment. Previous studies have summarized various types of CNS pathology, including WM

signal-intensity abnormality, ventricular dilation, hydrocephalus, enlarged perivascular spaces, a J-shaped sella turcica, an enlarged pituitary, and recently described posterior fossa horns (hypertrophy of the occipitomastoid sutures).^{2–5,7} Several case reports have reported large occipital emissary veins in healthy cohorts, and the presence of emissary veins potentially improves cerebellar venous outflow and thus prevents the increase in intracranial hypertension.^{7,8} To our knowledge, prominent emissary veins in MPSIH have not been reported in the literature. Recently in the literature, a high prevalence posterior fossa horns due to internal hypertrophy of the occipitomastoid sutures and the subsequent regression of posterior fossa horns have been described in MPSIH.^{3,4,12} We propose a possible interaction between the posterior fossa horns, the impaired venous outflow, and intracranial hypertension. More specifically, the prominent posterior fossa horns, which develop possibly due to GAG deposits, just inferior to the transverse sinuses and posterior to the sigmoid sinuses narrow the posterior fossa. The posterior fossa horns compress the cerebellar hemispheres, which, in turn, compress the sigmoid and transverse sinuses, resulting in venous hypertension and poor CSF resorption.

In adult patients with idiopathic intracranial hypertension, various degrees of venous sinus stenosis are seen, and alleviation of intracranial hypertension via venous sinus stent placement in those patients supports the hypothesis that elevation in venous pressure restricts CSF resorption.¹³ Furthermore in those patients, alternative cerebral venous drainage has been reported, and occipital emissary veins and extrajugular venous drainage have been described as hallmarks of idiopathic intracranial hypertension.^{9,10,14,15} We, therefore, postulated that the compression of the sigmoid and transverse sinuses drives the development or enlargement of emissary veins in patients with MPSIH to provide alternative venous outflow, alleviate the venous hypertension, and ultimately improve CSF resorption and decrease the likelihood of developing hydrocephalus.

Although the exact incidence of hydrocephalus in MPSIH is unclear, 1 large multicenter study reported 30.6% of patients with hydrocephalus before HSCT and 5.9% with persistent hydrocephalus after HSCT.^{11,16,17} In our study of 12 patients, we discovered that 3 of the 12 patients had VP shunt placement. All 12 patients

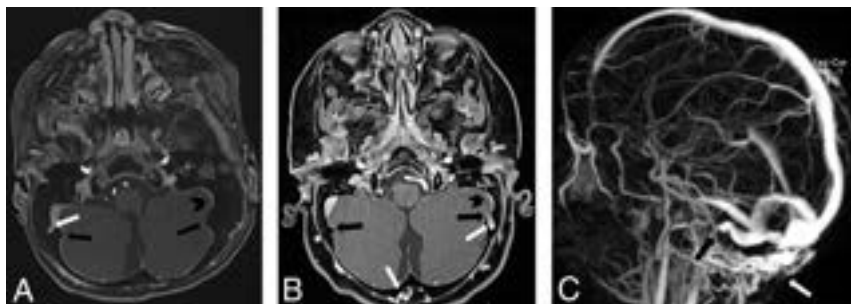


FIG 1. A, Axial T1-weighted image with contrast (patient 3) demonstrates prominent bilateral posterior fossa horns (black arrows) with a draining emissary vein from the right sigmoid sinus (white arrow) and a stenotic left sigmoid sinus (black arrowhead). B, Axial T1-weighted images with contrast (patient 10) demonstrate left-greater-than-right bilateral posterior fossa horns (black arrows) with a draining emissary vein from the left sigmoid sinus (white arrow), an emissary venous plexus (white arrow), and a stenotic left sigmoid sinus (black arrowhead). C, MRV (patient 8) shows a hypoplastic right transverse sinus (black arrow) with a large emissary venous plexus from the distal jugular bulb (white arrow).

had some degree of venous stenosis and posterior fossa horns compared with only 2 patients with posterior fossa horns and 1 patient with bilateral emissary veins in the control group. Almost all of them (11 of the 12) had prominent emissary veins from the sigmoid sinus or torcula. The mean OP among 9 of the 12 patients was 24 cm H₂O, which was the reported average OP in 25 patients with MPSIH in a previous study.¹¹ Only 3 of those 9 patients had an OP of >28 H₂O cm, which is the recently established upper normal limit of OP in the pediatric population.¹⁸ In our study of 12 patients, at the same time that the venogram was obtained, all 3 patients with elevated OP had

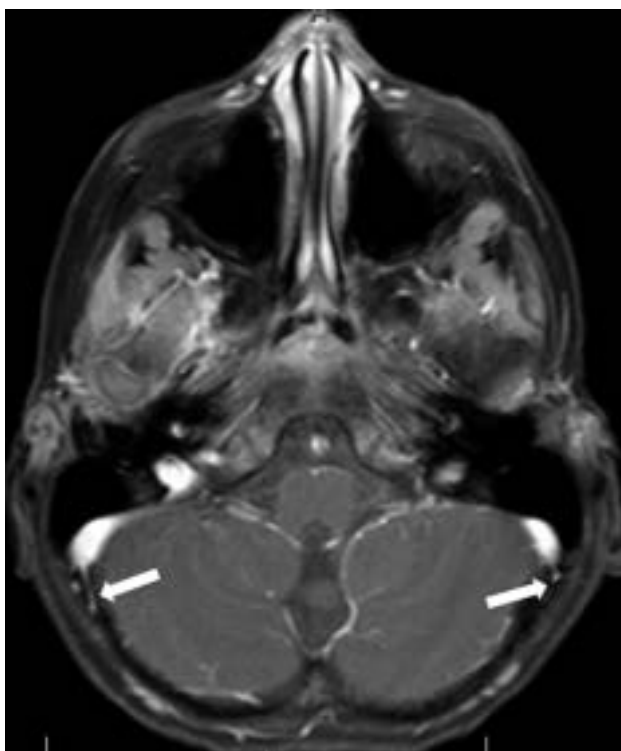


FIG 2. Axial T1-weighted image with contrast of patient 10 in the control population demonstrates small draining emissary veins from the bilateral sigmoid sinuses (white arrows).

venous stenosis and emissary veins, but none of them ultimately required shunt placement. Three of the 12 had VP shunts placed. One of the 3 patients developed acute ventriculomegaly and hydrocephalus on MR imaging 22 days after HSCT and required VP shunt placement, so the patient did not have OP measured and the patient's MRV was obtained as part of the pretransplantation evaluation. The other 2 patients with VP shunt placement had undergone the procedure 4 years after HSCT due medical refractory papilledema with normal OP, and their MR images were obtained 5 years after shunt placement for evaluation of hormonal deficiency. There was no apparent association of elevated OP or the incidence of VP shunting with the severity of posterior fossa horns or the degree of venous sinus stenosis. Emissary veins are equally prevalent, but they are found in patients with MPSIH with both normal and elevated OP, contrary to the patient group of similar ages that did not have MPSIH. Only one of them had bilateral emissary veins without venous sinus stenosis. The authors postulated that the various venous stenoses and subsequent development of emissary veins, in fact, might be a common finding in patients with MPSIH.

The retrospective nature of this study and the small size are 2 main limitations. Not all patients had a dedicated venogram, and not all patients had OP measurements. Several patients had venography and OP measurements at different stages of their treatment, ie, some of the studies were performed before HSCT and some were performed after HSCT. Multiple studies have reviewed the effects of HSCT in patients with MPSIH and found improvement in the incidence of hydrocephalus, possibly due to improvement in GAG deposits and CSF flow.^{17,19} Therefore, the OP may be affected by the intrinsic characteristics of the CSF and is not a surrogate for venous hypertension. However, our study

illustrated the high prevalence of venous stenosis and emissary veins, which was not previously described. Whether posterior fossa horns or venous sinus stenosis contributes to the development of hydrocephalus in MPSIH requires further investigation with a larger cohort.

CONCLUSIONS

Venous hypertension has long been postulated as one of the causes of hydrocephalus in patients with MPSIH. Adult studies in idiopathic intracranial hypertension have postulated that increased venous congestion leads to increased intracranial pressure. This study demonstrates that a similar mechanism may occur in patients with MPSIH. The posterior fossa horns likely cause venous sinus compression which, in turn, drives the development of compensatory collateral venous outflow, ie, emissary veins. However, given the high prevalence of venous emissary veins but the low incidence of a shunt, larger studies are needed to further characterize these features and hopefully help identify the patients who will develop hydrocephalus and guide the timing of CSF diversion.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

1. Muenzer J. **Overview of the mucopolysaccharidoses.** *Rheumatology (Oxford)* 2011;50(Suppl 5):v4–12 CrossRef Medline
2. Zafeiriou DI, Batzios SP. **Brain and spinal MR imaging findings in mucopolysaccharidoses: a review.** *AJNR Am J Neuroradiol* 2013;34:5–13 CrossRef Medline
3. Damar Ç, Derinkuyu BE, Kiliçkaya MA, et al. **Posterior fossa horns; a new calvarial finding of mucopolysaccharidoses with well-known cranial MRI features.** *Turk J Med Sci* 2020;50:1048–61 CrossRef Medline
4. Huang S, Lund T, Orchard P, et al. **Dilated optic nerve sheath in mucopolysaccharidosis I: common and not necessarily high intracranial pressure.** *AJNR Am J Neuroradiol* 2023;44:91–94 CrossRef Medline
5. Huang S, Beatty ZJ, McKinney AM, et al. **Increased pituitary volumes in patients with Sanfilippo syndrome (mucopolysaccharidosis type 3, MPS III).** *Neuroradiology* 2023;65:1381–86 CrossRef Medline
6. Dalla Corte A, de Souza CFM, Anés M, et al. **Hydrocephalus and mucopolysaccharidoses: what do we know and what do we not know?** *Childs Nerv Syst* 2017;33:1073–80 CrossRef Medline
7. Vedolin LM, Schwartz IV, Komlos M, et al. **Correlation of MR imaging and MR spectroscopy findings with cognitive impairment in mucopolysaccharidosis II.** *AJNR Am J Neuroradiol* 2007;28:1029–33 CrossRef Medline
8. Matheus MG, Castillo M, Smith JK, et al. **Brain MRI findings in patients with mucopolysaccharidosis types I and II and mild clinical presentation.** *Neuroradiology* 2004;46:666–72 CrossRef Medline
9. Salem M, Dolati P, Fusco MR, et al. **Abnormal large central occipital emissary vein: a case report and review of literature.** *Cureus* 2016;8:1–5 CrossRef
10. Hedjoudje A, Piveteau A, Gonzalez-Campo C, et al. **The occipital emissary vein: a possible marker for pseudotumor cerebri.** *AJNR Am J Neuroradiol* 2019;40:973–78 CrossRef Medline
11. Raymond GV, Pasquali M, Polgreen LE, et al. **Elevated cerebral spinal fluid biomarkers in children with mucopolysaccharidosis I-H.** *Sci Rep* 2016;6:4–9 CrossRef
12. Huang S, Hall D, Nascene D. **Posterior fossa horns in Hurler syndrome: prevalence and regression.** *AJNR Am J Neuroradiol* 2023;44:983–86 CrossRef Medline
13. Arun A, Amans MR, Higgins N, et al. **A proposed framework for cerebral venous congestion.** *Neuroradiol J* 2022;35:94–111 CrossRef Medline

14. Marsot-Dupuch K, Gayet-Delacroix M, Elmaleh-Bergès M, et al. **The petrosquamosal sinus: CT and MR findings of a rare emissary vein.** *AJNR Am J Neuroradiol* 2001;22:1186–93 Medline
15. Sattur MG, Amans M, Fargen KM, et al. **Angiographic evaluation of cranial venous outflow patterns in patients with and without idiopathic intracranial hypertension.** *Oper Neurosurg (Hagerstown)* 2023;24:e29–35 CrossRef Medline
16. Alden TD, Amartino H, Dalla Corte A, et al. **Surgical management of neurological manifestations of mucopolysaccharidosis disorders.** *Mol Genet Metab* 2017;122S:41–48 CrossRef Medline
17. Aldenhoven M, Wynn RF, Orchard PJ, et al. **Long-term outcome of Hurler syndrome patients after hematopoietic cell transplantation: an international multicenter study.** *Blood* 2015;125:2164–72 CrossRef Medline
18. Avery RA, Shah SS, Licht DJ, et al. **Reference range for cerebrospinal fluid opening pressure in children.** *N Engl J Med* 2010;363:891–93 CrossRef Medline
19. Eisengart JB, Rudser KD, Xue Y, et al. **Long-term outcomes of systemic therapies for Hurler syndrome: an international multicenter comparison.** *Genet Med* 2018;20:1423–29 CrossRef Medline