JANUARY 2019

AJNR

VOLUME 40 • PP 1–204

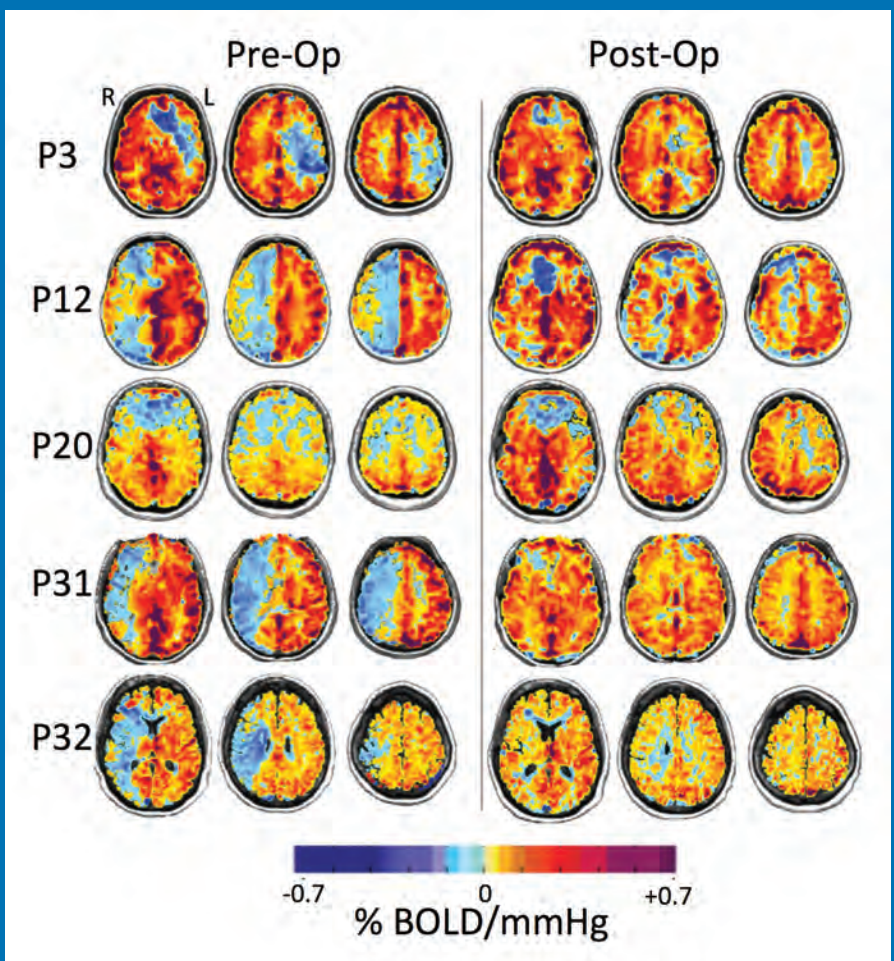
AJNR
AMERICAN JOURNAL OF NEURORADIOLOGY

JANUARY 2019
VOLUME 40
NUMBER 1
WWW.AJNR.ORG

THE JOURNAL OF DIAGNOSTIC AND INTERVENTIONAL NEURORADIOLOGY

Deep learning–based detection of intracranial aneurysms
Type II cortical dysplasia abnormalities
Solitary parathyroid adenoma localization

Official Journal ASNR • ASFNR • ASHNR • ASPNR • ASSR

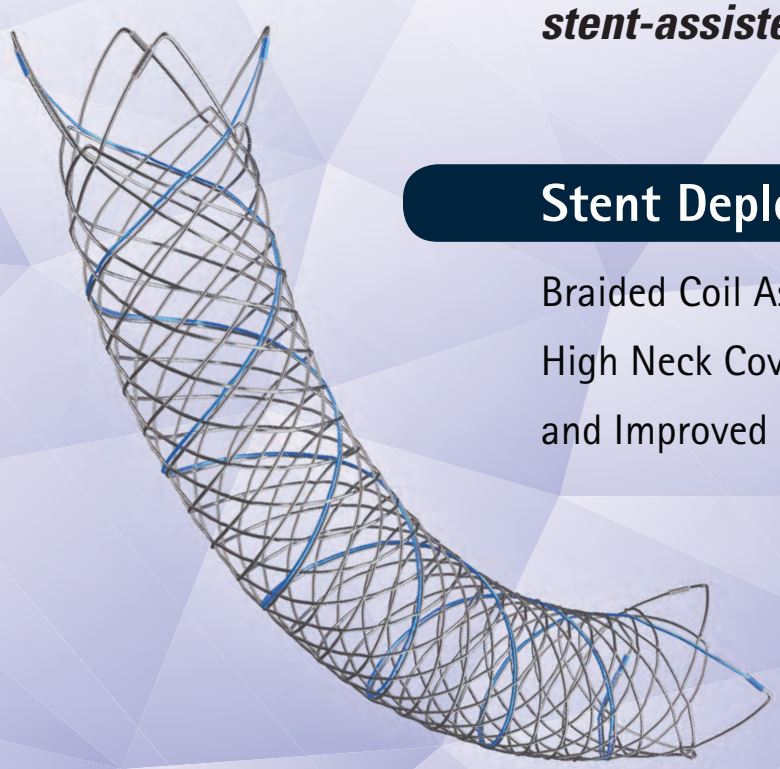


Aneurysm
Therapy
Solutions

LVIS®

Intraluminal Support Device

The first and only stent with Premarket Approval (PMA) for use in stent-assisted coil embolization



Stent Deployment. Refined.

Braided Coil Assist Stents with
High Neck Coverage, Excellent Visibility
and Improved Conformability*

Low-profile **V**isualized **I**ntraluminal **S**upport

INDICATIONS FOR USE:

The LVIS® and LVIS® Jr. devices are indicated for use with neurovascular embolization coils in patients ≥ 18 years of age for the treatment of wide-neck (neck width ≥ 4 mm or dome to neck ratio < 2) saccular intracranial aneurysms arising from a parent vessel with a diameter ≥ 2.0 mm and ≤ 4.5 mm.

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

The HydroCoil® Embolic System (HES) and MicroPlex® Coil System (MCS) are intended for the endovascular embolization of intracranial aneurysms and other neurovascular abnormalities such as arteriovenous malformations and arteriovenous fistulae. The HES and MCS are also intended for vascular occlusion of blood vessels within the neurovascular system to permanently obstruct blood flow to an aneurysm or other vascular malformation and for arterial and venous embolizations in the peripheral vasculature.

The device should only be used by physicians who have undergone pre-clinical training in all aspects of HES/MCS procedures as prescribed by MicroVention.

CE 0297

stryker

Now you have **24 hours** to make a lifetime of difference in stroke patients like Nora

The Trevo Retriever is the only device cleared to **reduce disability in stroke patients up to 24 hours** from time last seen well.

For more information, visit strykerneurovascular.com/trevo24hours

Trevo® XP
PROVUE RETRIEVER

Copyright © 2018 Stryker
AP002078 v1.0

AJNR
AMERICAN JOURNAL OF NEURORADIOLOGY

CALL FOR AJNR EDITORIAL FELLOWSHIP CANDIDATES

2019 Candidate Information and Requirements

ASNR and AJNR are pleased once again to join efforts with other imaging-related journals that have training programs on editorial aspects of publishing for trainees or junior staff (3–5 years after training), including Radiology (Olmsted fellowship), AJR (Figley and Rogers fellowships), JACR (Bruce J. Hillman fellowship), and Radiologia.

GOALS

- Increase interest in editorial and publication-related activities in younger individuals.
- Increase understanding and participation in the AJNR review process.
- Incorporate into AJNR's Editorial Board younger individuals who have previous experience in the review and publication process.
- Fill a specific need in neuroradiology not offered by other similar fellowships.
- Increase the relationship between "new" generation of neuroradiologists and more established individuals.
- Increase visibility of AJNR among younger neuroradiologists.

ACTIVITIES OF THE FELLOWSHIP

- Serve as Editorial Fellow for one year. This individual will be listed on the masthead as such.
- Review at least one manuscript per month for 12 months. Evaluate all review articles submitted to AJNR.
- Learn how electronic manuscript review systems work.
- Be involved in the final decision of selected manuscripts together with the Editor-in-Chief.
- Participate in all monthly Senior Editor telephone conference calls.
- Participate in all meetings of the Editors during the annual meetings of ASNR and RSNA and the Radiology Editors Forum as per candidate's availability. The Foundation of the ASNR will provide $2000 funding for this activity.
- Evaluate progress and adjust program to specific needs in annual meeting or telephone conference with the Editor-in-Chief.
- Embark on an editorial scientific or bibliometric project that will lead to the submission of an article to AJNR or another appropriate journal as determined by the Editor-in-Chief. This project will be presented by the Editorial Fellow at the ASNR annual meeting.
- Serve as liaison between AJNR and ASNR's Young Professionals Network. Participate in meetings and telephone calls with this group. Design one electronic survey/year, polling the group regarding readership attitudes and wishes.
- Recruit trainees as reviewers as determined by the Editor-in-Chief.
- Organize and host a Fellows' Journal Club podcast.
- Serve as Guest Editor for an issue of AJNR's News Digest with a timely topic.

QUALIFICATIONS

- Be a fellow in neuroradiology from North America, including Canada (this may be extended to include other countries).
- Be a junior faculty neuroradiology member (< 3 years) in either an academic or private environment.
- Be an "in-training" or member of ASNR in any other category.

APPLICATION

- Include a short letter of intent with statement of goals and desired research project. CV must be included.
- Include a letter of recommendation from the Division Chief or fellowship program director. A statement of protected time to perform the functions outlined is desirable.
- Applications will be evaluated by AJNR's Senior Editors prior to the ASNR meeting. The name of the selected individual will be announced at the meeting.
- Applications should be received by March 1, 2019 and sent to Ms. Karen Halm, AJNR Managing Editor, electronically at khalm@asnr.org.

AXS Infinity LS™ Plus Long Sheath

See package insert for complete indications, contraindications, warnings and instructions for use.

INDICATIONS FOR USE

The AXS Infinity LS Plus Long Sheath is indicated for the introduction of interventional devices into the peripheral, coronary, and neuro vasculature.

RX ONLY

CONTRAINDICATIONS

There are no known contraindications.

POTENTIAL ADVERSE EVENTS

- Acute vessel occlusion
- Air embolism
- Death
- Distal embolization
- Emboli
- False aneurysm formation
- Hematoma or hemorrhage at the puncture site
- Infection
- Intracranial hemorrhage
- Ischemia
- Neurological deficit including stroke
- Vessel spasm, thrombosis, dissection or perforation

WARNINGS

Contents supplied STERILE using an ethylene oxide (EO) process. Do not use if sterile barrier is damaged. If damage is found, call your Stryker Neurovascular representative.

For single use only. Do not reuse, reprocess or resterilize. Reuse, reprocessing or resterilization may compromise the structural integrity of the device and/or lead to device failure which, in turn, may result in patient injury, illness or death. Reuse, reprocessing or resterilization may also create a risk of contamination of the device and/or cause patient infection or cross-infection, including, but not limited to, the transmission of infectious disease(s) from one patient to another. Contamination of the device may lead to injury, illness or death of the patient.

After use, dispose of product and packaging in accordance with hospital, administrative and/or local government policy.

1. Do not re-sterilize or reuse, intended for single use only. Re-sterilization and/or reuse may result in cross contamination and/or reduced performance.
2. When the long sheath is exposed to the vascular system, it should be manipulated while under high-quality fluoroscopic observation. Do not advance or retract the long sheath if resistance is met during manipulation; determine the cause of the resistance before proceeding.

PRECAUTIONS

1. Store in a cool, dry, dark place.
2. Do not use kinked, damaged, or opened devices.
3. Use the device prior to the "Use By" date specified on the package.
4. Exposure to temperatures above 54°C (130°F) may damage device. Do not autoclave.
5. Torquing or moving the device against resistance may result in damage to the vessel or device.
6. Maintain a constant infusion of appropriate flush solution.
7. If flow through the device becomes restricted, do not attempt to clear the lumen by infusion. Remove and replace the device.
8. Examine the device to verify functionality and to ensure that its size and shape are suitable for the specific procedure for which it is to be used.
9. The AXS Infinity LS Plus Long Sheath should be used only by physicians trained in percutaneous procedures and/or interventional techniques.
10. Do not use if labeling is incomplete or illegible.

AXS Vecta™ 71 Aspiration Catheter

See package insert for complete indications, contraindications, warnings and instructions for use.

INDICATIONS FOR USE

The AXS Vecta™ Aspiration System, including the AXS Vecta 71 Aspiration Catheter, Aspiration Tubing Set, and VC-701 Cliq Aspirator Pump, is indicated in the revascularization of patients with acute ischemic stroke secondary to intracranial large vessel occlusive disease (within the internal carotid, middle cerebral – M1 and M2 segments, basilar, and vertebral arteries) within 8 hours of symptom onset. Patients who are ineligible for intravenous tissue plasminogen activator (IV t-PA) or who failed IV t-PA therapy are candidates for treatment.

RX ONLY

DEVICE DESCRIPTION

The AXS Vecta Aspiration System consists of the **AXS Vecta 71** Aspiration Catheter, the Aspiration Tubing Set, and the VC 701 Cliq Aspirator Pump.

The AXS Vecta 71 Aspiration Catheter is a single lumen, flexible, variable stiffness catheter. It has a radiopaque marker band on the distal end and a Luer hub at the proximal end. The AXS Vecta 71 Aspiration Catheter shaft has a lubricious coating at the distal end to reduce friction during use. The Scout Introducer may be used in conjunction with the AXS Vecta 71 Aspiration Catheter to facilitate in the introduction of the AXS Vecta 71 Aspiration Catheter into distal vasculature and aid in navigation to distal anatomy. The Scout Introducer has a lubricious coating at the distal end to reduce friction during use. The inner lumen of the AXS Vecta 71 Aspiration Catheters is compatible with the Scout Introducer, guide wires and micro catheters. The inner lumen of the Scout Introducer is compatible with guide wires and micro catheters of an outer diameter of less than 0.044in.

Each package includes one AXS Vecta 71 Aspiration Catheter, one Scout Introducer, one hemostasis valve, and two peel-away introducers. Dimensions of the AXS Vecta 71 Aspiration Catheter and Scout Introducer are included on the individual device label. The AXS Vecta 71 Aspiration Catheters are available in 3 different lengths, the device configurations including the length of the Scout packaged with each catheter and the recommended microcatheter length is presented in the table below.

Catheter part number	INC-11129-115	INC-11129-125	INC-11129-132
Catheter inner diameter (in)	0.071	0.071	0.071
Distal catheter outer diameter (in)	0.082	0.082	0.082
Catheter working length (cm)	115	125	132
Scout Introducer length (cm)	133	143	150
Recommended compatible microcatheter length (cm)	150	160	160
Recommended compatible microcatheter outer diameter (in)	0.044 max	0.044 max	0.044 max
Recommended compatible guidewire outer diameter (in)	0.038 max	0.038 max	0.038 max

The AXS Vecta Aspiration System is recommended for use in the following vessel size ranges based on non-clinical testing:

AXS Vecta 71 Aspiration Catheter	Vessel size (mm)
INC-11129-115	2-4
INC-11129-125	2-4
INC-11129-132	2-4

CONTRAINDICATIONS

The AXS Vecta 71 Aspiration Catheter has not been evaluated for use in the coronary vasculature. Do not use automated high-pressure contrast injection equipment with the AXS Vecta 71 Aspiration Catheter because it may damage the device.

POTENTIAL ADVERSE EVENTS

- Acute vessel occlusion
- Air embolism
- Allergic reaction and anaphylaxis from contrast media
- Arteriovenous fistula
- Death
- Device malfunction
- Distal embolization
- Emboli
- False aneurysm formation
- Hematoma or hemorrhage at the puncture site
- Inability to completely remove thrombus
- Infection
- Intracranial hemorrhage
- Ischemia
- Kidney damage from contrast media
- Neurological deficit including stroke
- Risks associated with angiographic and fluoroscopic radiation including but not limited to: alopecia, burns ranging in severity from skin reddening to ulcers, cataracts, and delayed neoplasia
- Vessel spasm, thrombosis, dissection or perforation

WARNINGS

Contents supplied STERILE using an ethylene oxide (EO) process. Do not use if sterile barrier is damaged. If damage is found, call your Stryker representative.

For single use only. Do not reuse, reprocess or resterilize. Reuse, reprocessing or resterilization may compromise the structural integrity of the device and/or lead to device failure which, in turn, may result in patient injury, illness or death. Reuse, reprocessing or resterilization may also create a risk of contamination of the device and/or cause patient infection or cross-infection, including, but not limited to, the transmission of infectious disease(s) from one patient to another. Contamination of the device may lead to injury, illness or death of the patient.

After use, dispose of product and packaging in accordance with hospital, administrative and/or local government policy.

1. The AXS Vecta 71 Aspiration Catheter has not been evaluated for more than one (1) clot retrieval attempt.
2. The AXS Vecta 71 Aspiration Catheter was evaluated for an average duration of direct aspiration of 4 minutes.
3. This product is intended for single use only, do not re-sterilize or reuse. Re-sterilization and/or reuse may result in cross contamination and/or reduced performance.
4. When the catheter is exposed to the vascular system, it should be manipulated while under high-quality fluoroscopic observation. Do not advance or retract the catheter if resistance is met during manipulation; determine the cause of the resistance before proceeding.
5. Operators should take all necessary precautions to limit x-radiation doses to patients and themselves by using sufficient shielding, reducing fluoroscopy times, and modifying x-ray technical factors where possible.

PRECAUTIONS

1. Store in a cool, dry, dark place.
2. Do not use kinked, damaged, or opened devices.
3. Use the device prior to the "Use By" date specified on the package.
4. Exposure to temperatures above 54°C (130°F) may damage device. Do not autoclave.
5. Torqueing or moving the device against resistance may result in damage to the vessel or device.
6. Maintain a constant infusion of appropriate flush solution.
7. If flow through the device becomes restricted, do not attempt to clear the lumen by infusion. Remove and replace the device.
8. Examine the device to verify functionality and to ensure that its size and shape are suitable for the specific procedure for which it is to be used.
9. The AXS Vecta Aspiration System should be used only by physicians trained in percutaneous procedures and/or interventional techniques.
10. The Scout Introducer should be used with a guidewire and microcatheter inserted when in vasculature.
11. If using the AXS Vecta Aspiration System for thrombectomy, monitor the canister fluid level and replace the canister if the fill level reaches 75% of the canister volume.
12. Administration of anticoagulants and antiplatelets should be suspended until 24 hours post-treatment. Medical management and acute post stroke care should follow the American Stroke Association (ASA) guidelines.
13. Any neurological determination should be evaluated by urgent CT scan and other evaluations as indicated according to investigator/hospital best practice.
14. As in all surgical interventions, monitoring of intra-procedural blood loss is recommended so that appropriate management may be instituted.
15. Limit the usage of the AXS Vecta 71 Aspiration Catheter to arteries greater than the catheter's outer diameter.
16. Excessive aspiration with the distal tip of the AXS Vecta 71 Aspiration Catheter covered by the vessel wall may cause vessel injury. Carefully investigate location of the distal tip under fluoroscopy prior to aspiration.
17. There is an inherent risk with the use of angiography and fluoroscopy.
18. When transporting the VC-701 Cliq pump, utilize the pump handle.
19. Do not use if labeling is incomplete or illegible.

Stryker Neurovascular
47900 Bayside Parkway
Fremont, CA 94538

strykerneurovascular.com

Date of Release: SEP/2018

EX_EN_US

Stryker Corporation or its divisions or other corporate affiliated entities own, use or have applied for the following trademarks or service marks: AXS Infinity LS, AXS Vecta, Stryker. All other trademarks are trademarks of their respective owners or holders.

Scout is a trademark of InNeuroCo, Inc.

Copyright © 2018 Stryker

Trevo® XP ProVue Retrievers

See package insert for complete indications, complications, warnings, and instructions for use.

INDICATIONS FOR USE

1. The Trevo Retriever is indicated for use to restore blood flow in the neurovasculature by removing thrombus for the treatment of acute ischemic stroke to reduce disability in patients with a persistent, proximal anterior circulation, large vessel occlusion, and smaller core infarcts who have first received intravenous tissue plasminogen activator (IV t-PA). Endovascular therapy with the device should start within 6 hours of symptom onset.
2. The Trevo Retriever is intended to restore blood flow in the neurovasculature by removing thrombus in patients experiencing ischemic stroke within 8 hours of symptom onset. Patients who are ineligible for intravenous tissue plasminogen activator (IV t-PA) or who fail IV t-PA therapy are candidates for treatment.
3. The Trevo Retriever is indicated for use to restore blood flow in the neurovasculature by removing thrombus for the treatment of acute ischemic stroke to reduce disability in patients with a persistent, proximal anterior circulation, large vessel occlusion of the internal carotid artery (ICA) or middle cerebral artery (MCA)-M1 segments with smaller core infarcts (0-50cc for age < 80 years, 0-20cc for age ≥ 80 years). Endovascular therapy with the device should start within 6-24 hours of time last seen well in patients who are ineligible for intravenous tissue plasminogen activator (IV t-PA) or who fail IV t-PA therapy.

COMPLICATIONS

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

COMPATIBILITY

3x20mm retrievers are compatible with Trevo® Pro 14 Microcatheters (REF 90231) and Trevo® Pro 18 Microcatheters (REF 90238). 4x20mm retrievers are compatible with Trevo® Pro 18 Microcatheters (REF 90238). 4x30mm retrievers are compatible with Excelsior® XT-27® Microcatheters (150cm x 6cm straight REF 275081) and Trevo® Pro 18 Microcatheters (REF 90238). 6x25mm Retrievers are compatible with Excelsior® XT-27® Microcatheters (150cm x 6cm straight REF 275081). Recommended minimum vessel ID for all Retriever sizes is 2.5mm. Compatibility of the Retriever with other microcatheters has not been established. Performance of the Retriever device may be impacted if a different microcatheter is used.

Balloon Guide Catheters (such as Merci® Balloon Guide Catheter and FlowGate® Balloon Guide Catheter) are recommended for use during thrombus removal procedures.

Copyright © 2018 Stryker

Retrievers are compatible with the Abbott Vascular DOC® Guide Wire Extension (REF 22260).

Retrievers are compatible with Boston Scientific Rotating Hemostatic Valve (Ref 421242).

SPECIFIC WARNINGS FOR INDICATION 1

- The safety and effectiveness of the Trevo Retrievers in reducing disability has not been established in patients with large core infarcts (i.e. ASPECTS ≤ 7). There may be increased risks, such as intracerebral hemorrhage, in these patients.
- The safety and effectiveness of the Trevo Retrievers in reducing disability has not been established or evaluated in patients with occlusions in the posterior circulation (e.g., basilar or vertebral arteries) or for more distal occlusions in the anterior circulation.

SPECIFIC WARNINGS FOR INDICATION 2

- To reduce risk of vessel damage, take care to appropriately size Retriever to vessel diameter at intended site of deployment.

SPECIFIC WARNINGS FOR INDICATION 3

- The safety and effectiveness of the Trevo Retrievers in reducing disability has not been established in patients with large core infarcts (i.e., ASPECTS ≤ 7). There may be increased risks, such as intracerebral hemorrhage, in these patients.
- The safety and effectiveness of the Trevo Retrievers in reducing disabillity has not been established or evaluated in patients with occlusions in the posterior circulation (e.g., basilar or vertebral arteries) or for more distal occlusions in the anterior circulation.
- Users should validate their imaging software analysis techniques to ensure robust and consistent results for assessing core infarct size.

WARNINGS APPLIED TO ALL INDICATIONS

- Administration of IV t-PA should be within the FDA-approved window (within 3 hours of stroke symptom onset).
- To reduce risk of vessel damage, adhere to the following recommendations:
  - Do not perform more than six (6) retrieval attempts in same vessel using Retriever devices.
  - Maintain Retriever position in vessel when removing or exchanging Microcatheter.
- To reduce risk of kinking/fracture, adhere to the following recommendations:
  - Immediately after unsheathing Retriever, position Microcatheter tip marker just proximal to shaped section. Maintain Microcatheter tip marker just proximal to shaped section of Retriever during manipulation and withdrawal.
  - Do not rotate or torque Retriever.
  - Use caution when passing Retriever through stented arteries.
- The Retriever is a delicate instrument and should be handled carefully. Before use and when possible during procedure, inspect device carefully for damage. Do not use a device that shows signs of damage. Damage may prevent device from functioning and may cause complications.
- Do not advance or withdraw Retriever against resistance or significant vasospasm. Moving or torquing device against resistance or significant vasospasm may result in damage to vessel or device. Assess cause of resistance using fluoroscopy and if needed resheath the device to withdraw.
- If Retriever is difficult to withdraw from the vessel, do not torque Retriever. Advance Microcatheter distally, gently pull Retriever back into Microcatheter, and remove Retriever and Microcatheter as a unit. If undue resistance is met when withdrawing the Retriever into the Microcatheter, consider extending the Retriever using the Abbott Vascular DOC guidewire extension (REF 22260) so that the Microcatheter can be exchanged for a larger diameter catheter such as a DAC® Catheter. Gently withdraw the Retriever into the larger diameter catheter.
- Administer anti-coagulation and anti-platelet medications per standard institutional guidelines.
- Users should take all necessary precautions to limit X-radiation doses to patients and themselves by using sufficient shielding, reducing fluoroscopy times, and modifying X-ray technical factors where possible.

PRECAUTIONS

- Prescription only – device restricted to use by or on order of a physician.
- Store in cool, dry, dark place.
- Do not use open or damaged packages.
- Use by "Use By" date.
- Exposure to temperatures above 54°C (130°F) may damage device and accessories. Do not autoclave.
- Do not expose Retriever to solvents.
- Use Retriever in conjunction with fluoroscopic visualization and proper anti-coagulation agents.
- To prevent thrombus formation and contrast media crystal formation, maintain a constant infusion of appropriate flush solution between guide catheter and Microcatheter and between Microcatheter and Retriever or guidewire.
- Do not attach a torque device to the shaped proximal end of DOC® Compatible Retriever. Damage may occur, preventing ability to attach DOC® Guide Wire Extension.

DOC is a trademark of Abbott Laboratories.

Stryker Neurovascular
47900 Bayside Parkway
Fremont, CA 94538

strykerneurovascular.com

Date of Release: APR/2018

EX_EN_US

stryker

Redefine aspiration.

AXS Vecta 71 Aspiration Catheter

Big 0.071in ID aspiration lumen to ingest more clot

Deliver through a 0.088in ID long sheath or **the new 0.091in AXS Infinity LS™ Plus Long Sheath**

Packaged with the Scout Introducer, a 0.044in lumen nitinol cross coil catheter that replaces the need for a 3MAX or other delivery catheter

AXS Vecta™ 71
ASPIRATION CATHETER

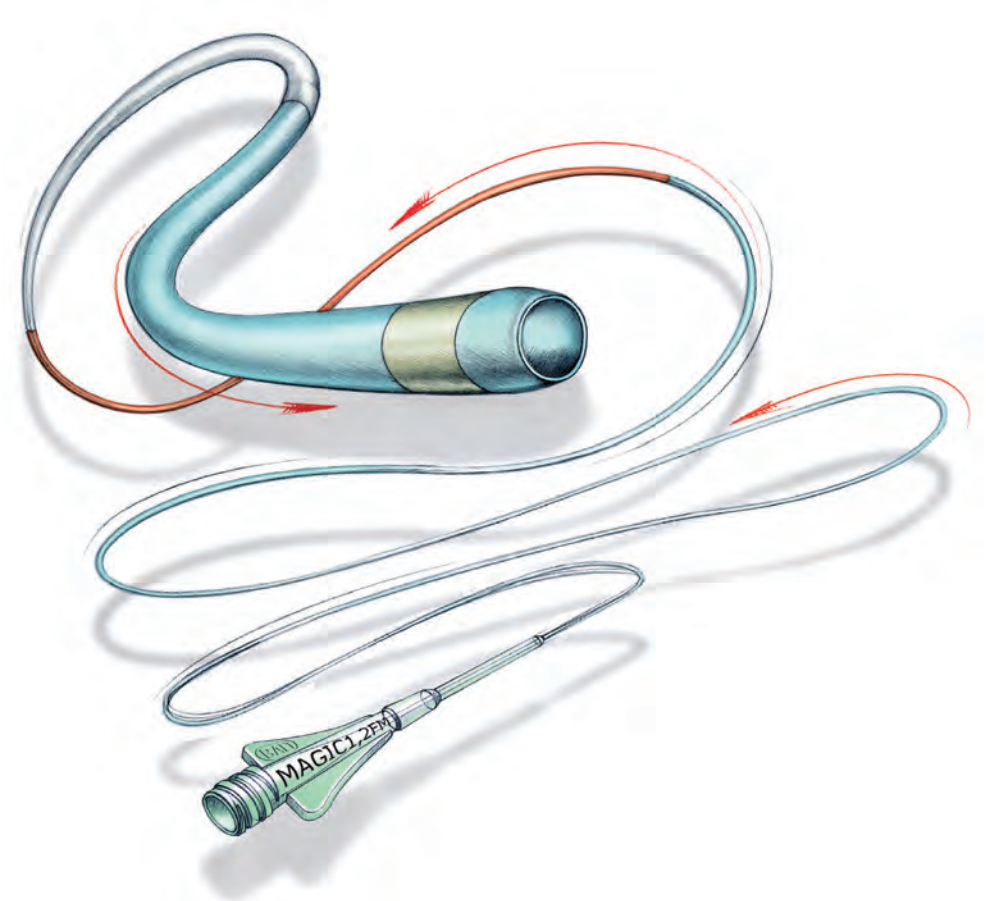
Copyright © 2018 Stryker

Magic

BALT

FLOW-DEPENDENT MICROCATHETER

NOW AVAILABLE THROUGH BLOCKADE™ MEDICAL



ORDERMAGICS@BLOCKADEMEDICAL.COM

MAGIC catheters are designed for general intravascular use. They may be used for the controlled, selective regional infusion of therapeutic agents or embolic materials into vessels.[1]

1. Magic Catheters IFU – Ind 19

MKTG-068 Rev. A

18 Technology Drive #169, Irvine Ca 92618
P 949.788.1443 | **F** 949.788.1444

BLOCKADE MEDICAL
A BALT Company

AJNR

AMERICAN JOURNAL OF NEURORADIOLOGY

JANUARY 2019
VOLUME 40
NUMBER 1
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.

1 **PERSPECTIVES** *T. Peker*

EDITORIAL

2 **Women Rising to the Top: The Tipping Point for the ASNR** *C.C. Meltzer, et al.*

REVIEW ARTICLE

5 **Imaging of Surgical Free Flaps in Head and Neck Reconstruction** *J.L. McCarty, et al.* — HEAD & NECK

PRACTICE PERSPECTIVES

14 **The ASNR-ACR-RSNA Common Data Elements Project: What Will It Do for the House of Neuroradiology?** *A.E. Flanders, et al.*

19 **Bias in Neuroradiology Peer Review: Impact of a "Ding" on "Dinging" Others** *P. Charkhchi, et al.*

GENERAL CONTENTS

25 **Deep Learning–Based Detection of Intracranial Aneurysms in 3D TOF-MRA** *T. Sichtermann, et al.* — ADULT BRAIN FUNCTIONAL

33 **Automated ASPECTS on Noncontrast CT Scans in Patients with Acute Ischemic Stroke Using Machine Learning** *H. Kuang, et al.* — ADULT BRAIN FUNCTIONAL

39 **Radiomics-Based Intracranial Thrombus Features on CT and CTA Predict Recanalization with Intravenous Alteplase in Patients with Acute Ischemic Stroke** *W. Qiu, et al.* — ADULT BRAIN FUNCTIONAL

45 **Improved White Matter Cerebrovascular Reactivity after Revascularization in Patients with Steno-Occlusive Disease** *L. McKetton, et al.* — ADULT BRAIN FUNCTIONAL

51 **Infarct Volume Predicts Hospitalization Costs in Anterior Circulation Large-Vessel Occlusion Stroke** *C.D. Streib, et al.* — ADULT BRAIN

59 **MRI Quantitative T2* Mapping to Predict Dominant Composition of In Vitro Thrombus** *R. Bourcier, et al.* — ADULT BRAIN

65 **Two-Layered Susceptibility Vessel Sign and High Overestimation Ratio on MRI Are Predictive of Cardioembolic Stroke** *R. Bourcier, et al., on behalf of the THRACE Investigators* — ADULT BRAIN

AJNR (Am J Neuroradiol ISSN 0195–6108) is a journal published monthly, owned and published by the American Society of Neuroradiology (ASNR), 800 Enterprise Drive, Suite 205, Oak Brook, IL 60523. Annual dues for the ASNR include $170.00 for journal subscription. The journal is printed by Cadmus Journal Services, 5457 Twin Knolls Road, Suite 200, Columbia, MD 21045; Periodicals postage paid at Oak Brook, IL and additional mailing offices. Printed in the U.S.A. POSTMASTER: Please send address changes to American Journal of Neuroradiology, P.O. Box 3000, Denville, NJ 07834, U.S.A. Subscription rates: nonmember $410 ($480 foreign) print and online, $320 online only; institutions $470 ($540 foreign) print and basic online, $935 ($1000 foreign) print and extended online, $380 online only (basic), extended online $825; 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, Meta/CZI and ReadCube. Copyright © American Society of Neuroradiology.

68 **Differences in the Calculated Transvenous Pressure Drop between Chronic Hydrocephalus and Idiopathic Intracranial Hypertension** *G.A. Bateman, et al.* — ADULT BRAIN

74 **Absence of Disproportionately Enlarged Subarachnoid Space Hydrocephalus, a Sharp Callosal Angle, or Other Morphologic MRI Markers Should Not Be Used to Exclude Patients with Idiopathic Normal Pressure Hydrocephalus from Shunt Surgery** *S. Agerskov, et al.* — ADULT BRAIN

80 **Brain β-Amyloid and Atrophy in Individuals at Increased Risk of Cognitive Decline** *I.K. Martikainen, et al.* — ADULT BRAIN

86 **Defining the Normal Dorsal Contour of the Corpus Callosum with Time** *K.L. Krause, et al.* — ADULT BRAIN

92 **Compressed Sensing–Sensitivity Encoding (CS-SENSE) Accelerated Brain Imaging: Reduced Scan Time without Reduced Image Quality** *J.E. Vranic, et al.* — ADULT BRAIN

99 **Determinants of Deep Gray Matter Atrophy in Multiple Sclerosis: A Multimodal MRI Study** *G. Pontillo, et al.* — ADULT BRAIN

107 ***Commentary***
What Causes Deep Gray Matter Atrophy in Multiple Sclerosis?
M.M. Schoonheim, et al.

109 **Clinical Feasibility of Zero TE Skull MRI in Patients with Head Trauma in Comparison with CT: A Single-Center Study** *S.B. Cho, et al.* — ADULT BRAIN

116 **Coil Embolization in Patients with Recurrent Cerebral Aneurysms Who Previously Underwent Surgical Clipping** *S.-T. Kim, et al.* — INTERVENTIONAL

122 **Y-Stent-Assisted Coiling of Wide-Neck Bifurcation Intracranial Aneurysms: A Meta-Analysis** *F. Cagnazzo, et al.* — INTERVENTIONAL

129 **Osseous versus Nonosseous Spinal Epidural Arteriovenous Fistulas: Experiences of 13 Patients** *Y. Song, et al.* — INTERVENTIONAL, SPINE

135 **Bone Subtraction Iodine Imaging Using Area Detector CT for Evaluation of Skull Base Invasion by Nasopharyngeal Carcinoma** *T. Hiyama, et al.* — HEAD & NECK

142 **Solitary Parathyroid Adenoma Localization in Technetium Tc99m Sestamibi SPECT and Multiphase Multidetector 4D CT** *T.H. Vu, et al.* — HEAD & NECK

150 **Parapharyngeal Space Venous Malformation: An Imaging Mimic of Pleomorphic Adenoma** *C.M. Tomblinson, et al.* — HEAD & NECK

154 **MR Imaging–Based Radiomic Signatures of Distinct Molecular Subgroups of Medulloblastoma** *M. Iv, et al.* — PEDIATRICS, FUNCTIONAL

162 **Early Diagnosis of Spastic Cerebral Palsy in Infants with Periventricular White Matter Injury Using Diffusion Tensor Imaging** *H. Jiang, et al.* — PEDIATRICS, FUNCTIONAL

169 **Cesarean Delivery Impacts Infant Brain Development** *S.C. Deoni, et al.* — PEDIATRICS

178 **MR Imaging of the Brain in Neurologic Wilson Disease** *X.-E. Yu, et al.* — PEDIATRICS

184 **MRI Abnormalities Predominate in the Bottom Part of the Sulcus with Type II Focal Cortical Dysplasia: A Quantitative Study** *Z. Liu, et al.* — PEDIATRICS

191 **Reliability of MR Imaging–Based Posterior Fossa and Brain Stem Measurements in Open Spinal Dysraphism in the Era of Fetal Surgery** *M. Aertsen, et al.* — PEDIATRICS, SPINE

199 **"Ears of the Lynx" MRI Sign Is Associated with SPG11 and SPG15 Hereditary Spastic Paraplegia** *B. Pascual, et al.* — PEDIATRICS

204 35 YEARS AGO IN AJNR

ONLINE FEATURES

LETTERS

E1 **T1-Hyperintense Plaques on Intracranial-versus-Extracranial Vessel Wall MRI** *W. Yuan, et al.*

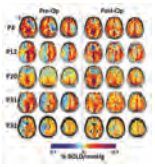
E2 ***Reply*** *C. Zhu, et al.*

E3 **Disproportionate International Contributions to Subspecialties of Neuroradiology in the *American Journal of Neuroradiology*** *S. Emamzadehfard, et al.*

E5 **ERRATUM**

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

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



BOLD MRI cerebrovascular reactivity maps (CVR) of 5 patients who underwent STA-MCA bypass. The first column shows CVR maps before bypass, and the second column shows CVR maps within a 1-year follow-up after revascularization surgery. Areas shown in blue demonstrate decreased and negative CVR. CVR maps visually demonstrate improvement (reduction of steal) postbypass.

Indicates Editor's Choices selection

Indicates Fellows' Journal Club selection

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

Indicates article with supplemental on-line table

Indicates article with supplemental on-line photo

Indicates article with supplemental on-line video

Evidence-Based Medicine Level 1

Evidence-Based Medicine Level 2

AJNR

AMERICAN JOURNAL OF NEURORADIOLOGY

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.

JANUARY 2019 • VOLUME 40 • NUMBER 1 • WWW.AJNR.ORG

Official Journal:
American Society of Neuroradiology
American Society of Functional Neuroradiology
American Society of Head and Neck Radiology
American Society of Pediatric Neuroradiology
American Society of Spine Radiology

EDITOR-IN-CHIEF

Jeffrey S. Ross, MD
Professor of Radiology, Department of Radiology, Mayo Clinic College of Medicine, Phoenix, AZ

SENIOR EDITORS

Harry J. Cloft, MD, PhD
Professor of Radiology and Neurosurgery, Department of Radiology, Mayo Clinic College of Medicine, Rochester, MN

Thierry A.G.M. Huisman, MD
Radiologist-in-Chief, Texas Children's Hospital, Houston, TX

Yvonne W. Lui, MD
Associate Professor of Radiology, Chief of Neuroradiology, New York University School of Medicine, New York, NY

C.D. Phillips, MD, FACR
Professor of Radiology, Weill Cornell Medical College, Director of Head and Neck Imaging, New York-Presbyterian Hospital, New York, NY

Pamela W. Schaefer, MD
Clinical Director of MRI and Associate Director of Neuroradiology, Massachusetts General Hospital, Boston, Massachusetts, Associate Professor, Radiology, Harvard Medical School, Cambridge, MA

Charles M. Strother, MD
Professor of Radiology, Emeritus, University of Wisconsin, Madison, WI

STATISTICAL SENIOR EDITOR

Bryan A. Comstock, MS
Senior Biostatistician, Department of Biostatistics, University of Washington, Seattle, WA

ARTIFICIAL INTELLIGENCE DEPUTY EDITOR

Christopher G. Filippi, MD
Professor and Vice Chair of Biomedical and Translational Science, Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Lenox Hill Hospital and Greenwich Village Healthplex, New York, NY

EDITORIAL BOARD

Ashley H. Aiken, *Atlanta, GA*
Lea M. Alhilali, *Phoenix, AZ*
Kubilay Aydin, *Istanbul, Turkey*
John D. Barr, *Dallas, TX*
Ari Blitz, *Baltimore, MD*
Barton F. Branstetter IV, *Pittsburgh, PA*
Jonathan L. Brisman, *Lake Success, NY*
Keith Cauley, *Danville, PA*
James Y. Chen, *San Diego, CA*
Asim F. Choudhri, *Memphis, TN*
Daniel Chow, *Irvine, CA*
J. Matthew Debnam, *Houston, TX*
Seena Dehkharghani, *New York, NY*
Yonghong Ding, *Rochester, MN*
Clifford J. Eskey, *Hanover, NH*
Saeed Fakhran, *Phoenix, AZ*
Massimo Filippi, *Milan, Italy*
Reza Forghani, *Montreal, Quebec, Canada*
Nils D. Forkert, *Calgary, Alberta, Canada*
Wende N. Gibbs, *Los Angeles, CA*
Christine M. Glastonbury, *San Francisco, CA*
John L. Go, *Los Angeles, CA*
Philipp Gölitz, *Erlangen, Germany*
Allison Grayev, *Madison, WI*
Brent Griffith, *Detroit, MI*
Ajay Gupta, *New York, NY*
Rakesh Kumar Gupta, *Haryana, India*
Lotfi Hacein-Bey, *Sacramento, CA*
Christopher P. Hess, *San Francisco, CA*
Andrei Holodny, *New York, NY*
Benjamin Huang, *Chapel Hill, NC*
Mahesh V. Jayaraman, *Providence, RI*
Valerie Jewells, *Chapel Hill, NC*
Christof Karmonik, *Houston, TX*
Timothy J. Kaufmann, *Rochester, MN*
Hillary R. Kelly, *Boston, MA*
Toshibumi Kinoshita, *Akita, Japan*
Kennith F. Layton, *Dallas, TX*
Alexander Lerner, *Los Angeles, CA*
Michael Lev, *Boston, MA*
Karl-Olof Lovblad, *Geneva, Switzerland*
Franklin A. Marden, *Chicago, IL*
Joseph C. McGowan, *Merion Station, PA*
Stephan Meckel, *Freiburg, Germany*
Christopher J. Moran, *St. Louis, MO*
Takahisa Mori, *Kamakura City, Japan*
Suresh Mukherji, *Ann Arbor, MI*
Alexander J. Nemeth, *Chicago, IL*
Renato Hoffmann Nunes, *Sao Paulo, Brazil*
Sasan Partovi, *Cleveland, OH*
Laurent Pierot, *Reims, France*
Jay J. Pillai, *Baltimore, MD*
Whitney B. Pope, *Los Angeles, CA*
Joana Ramalho, *Lisbon, Portugal*
Otto Rapalino, *Boston, MA*
Álex Rovira-Cañellas, *Barcelona, Spain*
Paul M. Ruggieri, *Cleveland, OH*
Amit M. Saindane, *Atlanta, GA*
Erin Simon Schwartz, *Philadelphia, PA*
Lubdha M. Shah, *Salt Lake City, UT*
Maksim Shapiro, *New York, NY*
Timothy Shepherd, *New York, NY*
Mark S. Shiroishi, *Los Angeles, CA*
Bruno P. Soares, *Baltimore, MD*
Maria Vittoria Spampinato, *Charleston, SC*
Khin Khin Tha, *Sapporo, Hokkaido, Japan*
Krishnamoorthy Thamburaj, *Hershey, PA*
Cheng Hong Toh, *Taipei, Taiwan*
Aquilla S. Turk, *Charleston, SC*
Anja G. van der Kolk, *Madison, WI*
Willem Jan van Rooij, *Tilburg, Netherlands*
Arastoo Vossough, *Philadelphia, PA*
Elysa Widjaja, *Toronto, Ontario, Canada*
Max Wintermark, *Stanford, CA*
Ronald L. Wolf, *Philadelphia, PA*
Kei Yamada, *Kyoto, Japan*
Carlos Zamora, *Chapel Hill, NC*
Vahe M. Zohrabian, *New Haven, CT*

EDITORIAL FELLOW

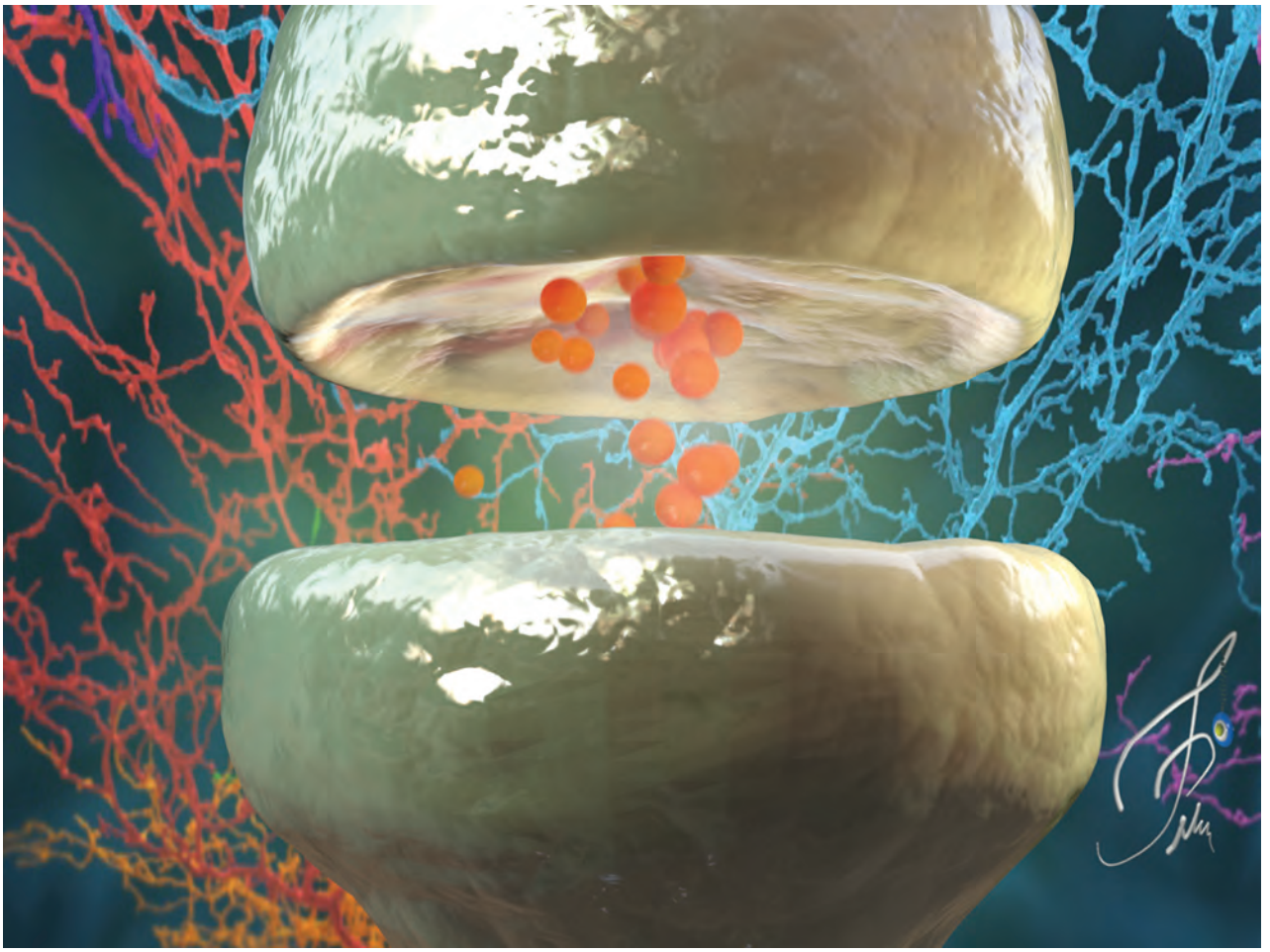
Alireza Radmanesh, *New York, NY*

SPECIAL CONSULTANTS TO THE EDITOR

AJNR Blog Editor
Neil Lall, *Denver, CO*
Case of the Month Editor
Nicholas Stence, *Aurora, CO*
Case of the Week Editors
Juan Pablo Cruz, *Santiago, Chile*
Sapna Rawal, *Toronto, Ontario, Canada*
Classic Case Editor
Sandy Cheng-Yu Chen, *Taipei, Taiwan*
Health Care and Socioeconomics Editor
Pina C. Sanelli, *New York, NY*
Physics Editor
Greg Zaharchuk, *Stanford, CA*
Podcast Editor
Wende N. Gibbs, *Los Angeles, CA*
Twitter Editor
Jennifer McCarty, *Houston, TX*

Founding Editor
Juan M. Taveras
Editors Emeriti
Mauricio Castillo, Robert I. Grossman, Michael S. Huckman, Robert M. Quencer

Managing Editor
Karen Halm
Assistant Managing Editor
Laura Wilhelm
Communications Coordinator
Rebecca Artz
Marketing and Social Media Manager
Kylie Mason
Executive Director, ASNR
Mary Beth Hepp



Title: Synaptic Cleft. This picture describes 3 components of a typical synapse. First, the presynaptic knob is located at the distal end of an axon from which neurotransmitters are released and then the synaptic cleft separates the presynaptic knob of the axon from the postsynaptic membrane of the dendrite. Finally, the postsynaptic membrane of the dendrite contains receptors with an affinity for the neurotransmitters. In this picture, all modelling and addition of texture, lighting, and visual effects were made in Cinema 4D software (Maxon Computer, Friedrichsdorf, Germany).

Tuncay Peker, MD, Gazi University Faculty of Medicine, Department of Anatomy, Ankara, Turkey

Women Rising to the Top: The Tipping Point for the ASNR

C.C. Meltzer, P.C. Sanelli, M.B. Hepp, and J.A. Bello

Like many leading professional academic societies, the American Society of Neuroradiology (ASNR) has enjoyed a rich history of visionary leadership. Yet, from its birth in 1962 through 2009, only 2 of the 46 ASNR Presidents were women (with 17 years separating their appointments). The striking change in demographics of the ASNR leadership during the past decade appears to demonstrate a clear tipping point. In this most recent decade, 5 of the 9 Presidents have been women (Figure). Given the establishment of the Women in Neuroradiology Leadership Award in 2010 and other efforts, ASNR stands out as a national radiology society that has reached a tipping point from which lessons may be learned.

The field of radiology has a long history of being male-dominated. Even as women have achieved equity in medical schools in recent years, radiology remains characterized by one of the lowest proportions of women among medical specialties.[1] With just more than one-quarter of radiologists being women, neuroradiologists especially lack female colleagues; the current ASNR membership includes less than 18% women (ASNR membership records 2018).

The advancement of women into leadership positions has been slow across medicine, with women representing a paucity of chairs, deans, and practice and society leaders.[1] Radiology is no exception. Thus, it was not until 2018 that the American College of Radiology (ACR) named its first female chair of the Board of Chancellors in its 95-year history. Recently, Ahmadi et al[2] reported a strong correlation between gender and academic leadership positions among neuroradiologists, with 87.5% of leadership ranks occupied by men.

Differential Support for Professional Advancement

Research accomplishments often serve as a gateway to professional advancement. There is mounting evidence that male physicians and scientists in science, technology, engineering and mathematic fields enjoy greater scholarship opportunities impacting consequential publications,[2,3] research grants,[4] awards and honors, and leadership opportunities relative to their female counterparts. Considering the grants and awards bestowed by the ASNR and the Foundation of the ASNR, the gender imbalance is evident (Table). Yet for some awards, there have been improvements in the past decade. For example, from 1992 to 2009, only 4 of the 38 Cornelius G. Dyke Awards (10.5%) were received by women; from 2010 to 2018, three of the 8 awards (37.5%) went to female applicants. Not unexpectedly, gender imbalance was greatest for the senior honorary awards such as the Gold Medal (bestowed to only 1 woman of 33 awardees [3.0%]) and Outstanding Research Contributions (bestowed to only 1 woman of 14 awardees [7.1%]).

Even in the most recent decade in which women were awarded more grants and honors than previously, the rates of awards still lagged behind the proportion of female applicants. Since the inception of the ASNR Comparative Effectiveness Awards in 2011, women have made up 41.3% of applicants yet were only 14.3% of awardees. Similarly, for the Research Scientist Award, 36.7% of applicants versus 12.5% of awardees were women.

Contributors to the Tipping Point

Tipping points in societal norms are an interesting phenomenon. Recent cultural examples that have received much attention are the rapid spread of public resistance to workplace harassment and mass gun violence as embodied by #MeToo and #NeverAgain, respectively. Centola et al[5] argue that tipping points in social convention may be explained by the theory of critical mass as posited by the evolutionary game theory. When a committed minority reaches a certain size, which they showed to be approximately 25% through experimental constructs, the social system crosses a tipping point at which a rapid change in attitude and behaviors favor the minority view.

Kim and Mauborgne[6] framed tipping point leadership as a change management that requires overcoming 4 types of hurdles: cognitive, political, resource, and motivational. This construct is applicable to examining factors that potentially contributed to the leadership tipping point of the ASNR.

Much has been written about cognitive barriers to the advancement of women's careers. Our implicit biases, particularly in male-dominated professions and organizations, reinforce our expectation that men will serve as leaders and women will occupy support roles.[7] The persistence of these deeply embedded but rarely spoken barriers is sometimes referred to as second-generation bias. In such a climate, women who express leadership interests may be criticized rather than supported.[8] Yet as women take on leadership positions and serve as counter-stereotype exemplars and role models for other women, cognitive barriers can be eroded with time.

Political hurdles may also be ingrained and not well-recognized. Coalitions of like-thinking individuals may exert undue influence on the strategy and direction of an organization and become largely responsible for the distribution of formal and informal positions of power.[9] This dynamic can be further solidified by organizational structures and by-laws that favor those in key positions assuming additional roles of power. One example might be the structure of the ASNR in which the President-Elect serves as the Chair of the Nominating Committee, a position that oversees the nomination process for the incoming leadership positions.

Resources that have been shown to support the professional advancement of women include mentorship and leadership development programs. In 2010, as the ASNR neared its 50th anniversary, the idea of an ASNR-sponsored award for promising midcareer female neuroradiologists began to develop. The intent was to support such women both through national recognition of their leadership promise and leadership skill-building. The following year, a call for nominations for the Women in Neuroradi-

http://dx.doi.org/10.3174/ajnr.A5893

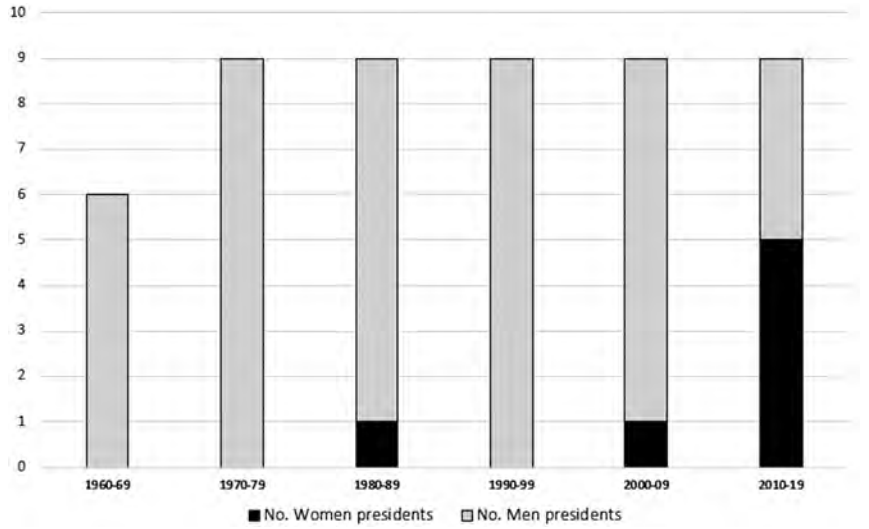


FIGURE. The number of male-versus-female ASNR Presidents by decade (note the first President was appointed in 1962 with an initial term of 2 years).

Gender balance of ASNR awards and grants

Award Name/Type	Years Awarded	All Years		Past Decade (2010–2018)	
		No. of Awardees (M/F)	% Female Awardees	No. of Awardees (M/F)	% Female Awardees
Cornelius G. Dyke Award/honor[a]	1992–2018[b]	38 (34:4)	10.5	8 (5:3)	37.5
Scholar Award/grant	1999–2018[c]	24 (14:10)	41.7	12 (8:4)	33.3
Gold Medal/honor	1995–2018	33 (32:1)	3.0	10 (10:0)	0.0
Outstanding Contributions/honor	2004–2018[d]	14 (13:1)	7.1	9 (8:1)	11.1
Research Scientist Award/grant	2011–2018	8 (7:1)	12.5	8 (7:1)	12.5
Comparative Effectiveness/grant[e]	2011–2018	7 (6:1)	14.3	7 (6:1)	14.3

[a] Established to honor Cornelius G. Dyke, one of the pioneers in neuroradiology, given to an assistant professor, fellow, or resident for excellence in original research.
[b] No award in 1976, 1980, 1983, 1985, 1987, 1998, 2001, 2003, 2005, 2008, and 2011.
[c] Two awards per year in 2010, 2017, and 2018.
[d] No award in 2006.
[e] No award in 2013.

ology Leadership Award was announced in partnership with the ACR and the American Association for Women Radiologists, and the first award was bestowed in 2012.[10] The annual awardees have been celebrated by all 3 organizations, and funding has been provided for the recipient to attend the ACR Radiology Leadership Institute Summit. Two of the 7 awardees to date have ascended to officer positions in the ASNR, including the current President, as well as attaining leadership roles in other professional organizations.

While the motivational forces responsible for an organizational tipping point are the most difficult to evaluate and quantify, it is likely that a critical mass of engaged women and male champions is largely responsible for the momentum achieved. While the leadership structure of the ASNR had been overwhelmingly male for most of its history, in this most recent decade women made up an average of nearly one-third of the executive committee (compared with an average of 12% during the prior decade). As a pipeline to the presidency, the appointment of increasing numbers of women to this body is further evidence of sustained change. While progress is to be celebrated and we believe the momentum for positive change is considerable, there is more work to do. Since its first issue in 1980, the premier journal in the field, the *American Journal of Neuroradiology*, has never had a female editor at its helm.

Summary

The ASNR is the premier professional society for neuroradiology, a field that persists in attracting and training a minority of women. While there has been only slow progress in the number of women entering radiology overall and particularly neuroradiology, in the past decade the ASNR has demonstrated dramatic positive change in the gender diversity of its leadership. Between 1962 and 2009, only 2 of the 46 ASNR Presidents were women, yet in this most recent decade, more than half of the Presidents appointed have been women. In this editorial, we attribute this tipping point change to a variety of factors.

Disclosures: Carolyn C. Meltzer—*UNRELATED*: *Board Membership*: GE-Association of University Radiologists Scientific Board, *Comments*: reimbursement from the Association of University Radiologists for travel to meetings; *Consultancy*: Wake Forest Health, University of Tennessee Medical Center, University of Massachusetts, University of Pennsylvania, Northwestern University; *Employment*: Emory University; *Expert Testimony*: Attorney General Nashville, Tennessee; Floyd Pflueger & Ringer, Seattle. Mary Beth Hepp—*UNRELATED*: *Employment*: ASNR, *Comments*: I am the Executive Director of the ASNR and have access to statistical data related to this article. Jacqueline A. Bello—*OTHER RELATIONSHIPS*: I am a recent past female President.

REFERENCES

1. AAMC Report: The State of Women in Academic Medicine: The Pipeline and Pathways to Leadership, 2013–2014. https://members.aamc.org/eweb/upload/The%20State%20of%20Women%20in%

20Academic%20Medicine%202013-2014%20FINAL.pdf. Accessed November 1, 2018
2. Ahmadi M, Khurshid K, Sanelli PC, et al. **Influences for gender disparity in academic neuroradiology.** *AJNR Am J Neuroradiol* 2018;39:18–23 CrossRef Medline
3. Shen YA, Webster JM, Shoda Y, et al. Persistent underrepresentation of women's science in high-profile journals. bioRxiv 2018 https://www.biorxiv.org/content/early/2018/03/08/275362. Accessed November 1, 2018
4. Magua W, Zhu X, Bhattacharya A, et al. **Are female applicants disadvantaged in National Institutes of Health peer review? Combining algorithmic text mining and qualitative methods to detect evaluative differences in R01 reviewers' critiques.** *J Womens Health (Larchmt)* 2017;26:560–70 CrossRef Medline
5. Centola D, Becker J, Brackbill D, et al. **Experimental evidence for tipping points in social convention.** *Science* 2018;360:1116–19 CrossRef Medline
6. Kim WC, Mauborgne R. Tipping Point Leadership**.** Harvard Business Review, April 2003. https://hbr.org/2003/04/tipping-point-leadership. Accessed June 29, 2018
7. Meltzer CC. **Women leaders: myths and challenges.** *J Am Coll Radiol* 2018. Jul 25. [Epub ahead of print] CrossRef Medline
8. Ibarra H, Ely RJ, Kolb DM. Women Rising: The Unseen Barriers. Harvard Business Review September 2013**.** https://hbr.org/2013/09/women-rising-the-unseen-barriers. Accessed June 2, 2018
9. Zaleznik A. Power and Politics in Organizational Life. Harvard Business Review, May 1970**.** https://hbr.org/1970/05/power-and-politics-in-organizational-life. Accessed September 2, 2018
10. Meltzer CC. **Three radiology societies come together to sponsor Women in Neuroradiology Leadership Award.** *AJNR Am J Neuroradiol* Blog. February 3, 2014. http://www.ajnrblog.org/2014/02/03/three-radiology-societies-come-together-sponsor-women-neuroradiology-leadership-award. Accessed September 2, 2018

Imaging of Surgical Free Flaps in Head and Neck Reconstruction

J.L. McCarty, A.S. Corey, M.W. El-Deiry, H.M. Baddour, B.M. Cavazuti, and P.A. Hudgins

ABSTRACT

SUMMARY: Head and neck surgical reconstruction is complex, and postoperative imaging interpretation is challenging. Surgeons now use microvascular free tissue transfer, also known as free flaps, more frequently in head and neck reconstruction than ever before. Thus, an understanding of free flaps, their expected appearance on cross-sectional imaging, and their associated complications (including tumor recurrence) is crucial for the interpreting radiologist. Despite the complexity and increasing frequency of free flap reconstruction, there is no comprehensive head and neck resource intended for the radiologist. We hope that this image-rich review will fill that void and serve as a go to reference for radiologists interpreting imaging of surgical free flaps in head and neck reconstruction.

ABBREVIATIONS: ALT = anterolateral thigh; CECT = contrast-enhanced CT; FF = free flap; H&N = head and neck; SCC = squamous cell carcinoma

Surgical free flaps (FFs) have become the preferred reconstruction method for most large head and neck (H&N) oncologic defects. They have better functional outcomes and cosmetic restoration than surgical grafts and other types of surgical flaps (local and regional).[1-4] Since free flaps were first used in the 1970s, surgeons have expanded their repertoire, fine-tuned techniques, and improved outcomes.[5] While FFs may be used to reconstruct defects from infection, trauma, and osteonecrosis, they are most often used following tumor extirpation. FFs are unfortunately frequent today as >675,000 patients worldwide are diagnosed with H&N cancer annually.[6]

Posttreatment H&N clinical and imaging follow-up is complex, even more so when the resection site is reconstructed. To provide accurate and useful H&N imaging reports, radiologists should have a basic understanding of the surgical options, ranging from skin grafts to surgical flaps. This review article focuses on free flaps, summarizing what defines an FF, those most commonly used, the expected postoperative imaging appearances, and associated complications.

Flap Versus Graft

Surgical flaps and surgical grafts are both used in H&N reconstruction. While the 2 are distinct entities, it is not uncommon to hear them incorrectly used interchangeably. Both are blocks of transferred tissue, but flaps have their own blood supply while grafts depend on angiogenesis.[5,7] Surgical flaps are transferred either with an intact vascular supply or the blood supply is reestablished at the recipient site using microvascular techniques. Conversely, the major graft vasculature is transected at the donor site and the tissue inset without vessel-to-vessel anastomosis.

Grafts may be autograft (from the patient), allograft (from a donor, often cadaveric), or alloplastic (man-made). Grafts are typically 1 or 2 tissue types, whereas surgical flaps are often more complex and contain several different tissue constituents. The most commonly used grafts in H&N reconstruction are the fairly straightforward skin grafts, which can be full thickness (complete segments of both epidermis and dermis) or split thickness (complete epidermis but incomplete varying-thickness dermis).[7,8] Bone grafts continue to be used for certain craniofacial and spine reconstructive procedures.[9] Often for large-volume composite defects of the head and neck, flaps are superior to grafts with respect to cosmesis because the bulk of flap tissue better fills defects and maintains its size and shape for the duration of the patient's life. Flaps also typically heal better, in a quicker and more predictable fashion with less contracture than grafts in the head and neck.[10,11]

Received April 13, 2018; accepted after revision June 18.

From the Department of Diagnostic and Interventional Imaging (J.L.M.), University of Texas Health Sciences Center, Houston, Texas; Departments of Radiology and Imaging Sciences (A.S.C., P.A.H.) and Otolaryngology (M.W.E.-D., H.M.B.), Emory University School of Medicine, Atlanta, Georgia; Quantum Radiology (B.M.C.), Marietta, Georgia; and Atlanta VA Healthcare System (A.S.C.), Atlanta, Georgia.

Paper previously presented as an educational electronic exhibit at: Annual Meeting of the American Society of Head and Neck Radiology, September 7–11, 2016; Washington, DC.

Please address correspondence to Jennifer L. McCarty, MD, Department of Diagnostic and Interventional Imaging, University of Texas Health Sciences Center, Houston, Texas; e-mail: jennifer.l.mccarty@uth.tmc.edu; @JMcCartyMD

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

http://dx.doi.org/10.3174/ajnr.A5776

Flap Types

Surgical flaps are typically classified by their pattern of vascularity and their proximity to the primary defect. In terms of blood supply, flaps can be typified as random or axial.[5,7,12,13] Random flaps are supplied by the subdermal plexus of the skin and are not supplied by distinct named vessels. In contrast, axial flaps are supplied by a specific arteriovenous system. Axial flaps are typically considered more reliable than random flaps on the basis of improved distal perfusion.

With respect to defect proximity, flaps can be classified as local, regional, or free (Fig 1). Local flaps, flaps generated adjacent to the primary defect, are examples of random flaps. Examples of local flaps include rotation, advancement, and transposition. Regional and free flaps typically have axial-based vascularity, wherein these flaps are dependent on a specific vascular pedicle for viability. An example of a regional flap is the pectoralis major myocutaneous flap, supplied by the pectoral branch of the thoracoacromial artery. Free tissue flaps are the most complex and technically challenging form of flap reconstruction. Each free flap is designed and harvested at a spatially distinct site from the primary defect known as the donor or harvest site. Each donor vascular pedicle is transected at the donor site, transferred along with the flap constituents to the primary defect, and inset at the primary defect. Then, with microsurgical techniques, the donor pedicle is anastomosed to the recipient vessels near the defect to re-establish the blood supply to the flap.[5,7,12,13]

An additional distinction with regional and free flaps is that they may be of simple or composite design. A simple flap is typically composed of skin and subcutaneous tissue. Composite flaps, like composite resection sites, can consist of multiple tissue types and often include bone and/or muscle.[5,7,13]

The decision of which flap to use is made on an individual case basis because some flap types are advantageous over others for certain patients, tumors, and reconstruction locations. For example, in patients with tongue reconstruction following glossectomy, those with FFs had superior speech intelligibility compared with those with pedicle flaps.[3,4]

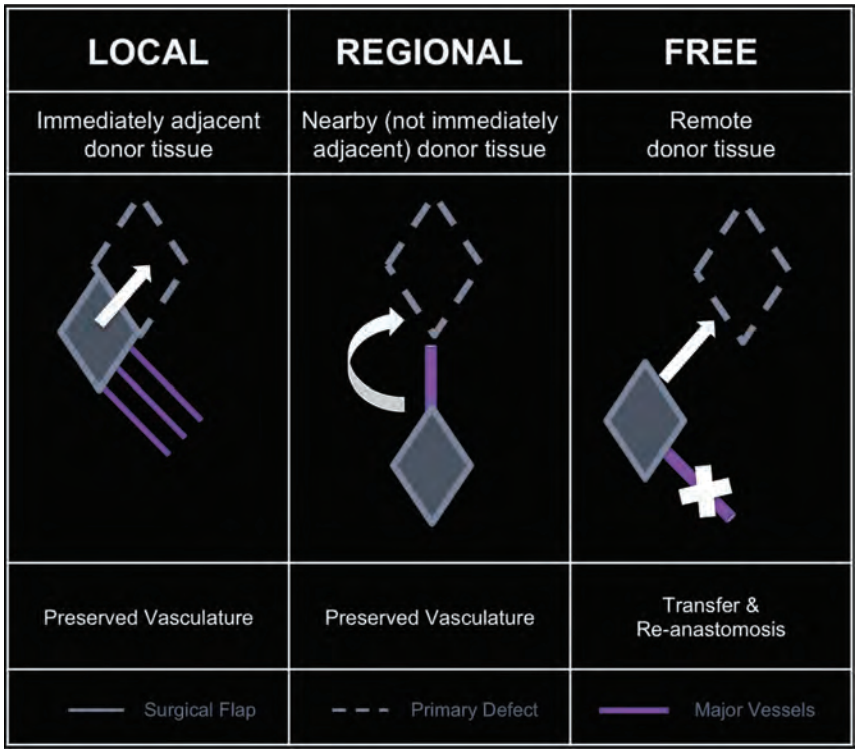


FIG 1. Flaps based on donor location with respect to the primary defect.

Free Flap Types

Surgeons innovatively use a wide variety of surgical free flaps (Table 1).[14] Most surgeons approach FF reconstruction by giving consideration to the oncologic surgical defect, the donor sites available, surgeon preference, and the available hospital and surgical resources (Table 2).[13,15,16] Because there is no "one size fits all" approach, the number and complexity of the many different types of surgical FFs are vast and can be overwhelming. Thus, for radiologists, it is useful to categorize free flaps by the dominant donor tissue constituents (Table 1), creating 4 main categories: muscular, fascial, osseous, and visceral FFs.[17]

This simplified approach has important facts to note. Although we are grouping these by the dominant tissue type, most of these FFs also contain skin paddles that reconstruct both skin and

Table 1: Surgical free flaps in head and neck reconstruction

Category	Free Flap	Reconstructs	Donor Artery
Muscular	Rectus abdominis	Skull base, orbit	Deep inferior epigastric
	Latissimus dorsi	Skull base, scalp	Thoracodorsal
Fascial	Radial forearm	Oral cavity, tongue, palate, nose, face, scalp, lip, pharynx, larynx	Radial
	Ulnar forearm	Oral cavity, tongue, palate, nose, face, scalp, lip, pharynx, larynx, cervical esophagus	Ulnar
	Lateral thigh	Oral cavity, tongue, palate, pharynx	Deep femoral
	Anterolateral thigh	Oral cavity, tongue, palate, pharynx, larynx, cervical esophagus	Descending branch, lateral circumflex femoral
	Scapula	Oral cavity, tongue, palate, nose, face, lip	Subscapular
Osseous	Fibula	Mandible	Peroneal
	Radius	Mandible & midface	Radial
	Scapula	Mandible & midface	Subscapular, thoracodorsal
	Iliac crest	Mandible & midface	Deep circumflex
Visceral	Jejunum	Pharynx, esophagus	Superior mesenteric branches
	Omentum	Scalp	Gastroepiploic

Table 2: Surgical approach to free flaps in H&N reconstruction

Factor	Question	Limitation
Defect	What needs to be replaced?	Constituents Function Size
Donor	What is available?	Body habitus Vascular integrity Vascular anomalies
Surgeon	What can be done?	Skills Support

mucosal surfaces because both should be reformed with epithelial tissue. Additionally, some FFs (rectus abdominis and latissimus dorsi) may be harvested as either myocutaneous or fasciocutaneous. Last, there is a subset of free flaps referred to as perforator flaps, defined by the surgical isolation of the small vessels perforating the muscles at the donor site, sparing the larger donor vessel and resulting in improved donor site morbidity but shorter FF vascular pedicles.[18]

Muscular. Muscle-containing FFs in H&N reconstruction are mostly myocutaneous (or musculocutaneous), composed of both muscle and skin as well as the adjacent subcutaneous tissue, vessels, and fascia. This category of FF is particularly useful to fill large defects (Fig 2), including skull base defects.[7,19] Two of the more commonly used myocutaneous FFs are the rectus abdominis and latissimus dorsi muscle FFs.

Rectus abdominis FFs include one of the paired vertically oriented ventral abdominal muscles.[5,20] They are especially advantageous because of the versatility of flap design and the length of the associated vascular pedicle, up to 10–15 cm long. This vascular pedicle length allows the surgeon to inset rectus FFs into almost any H&N defect, even those a distance away or on the opposite side of the head and neck from the vascular anastomoses. As with most cutaneous FFs, the skin paddle recreates the skin surface and mucosal surface (Fig 2). One drawback is that the thickness of the FF is particularly reflective of a patient's weight and may prove too bulky for some resection cavities in obese patients due to excessive subcutaneous fat.[5] Rectus abdominis FFs can also be harvested as fasciocutaneous free flaps.

Latissimus dorsi FFs are the largest muscle flaps, with a total area measuring up to 25 × 40 cm. However, the muscle itself is one of the thinnest in the body. The latissimus spans from the posterolateral thorax to the inferomedial back (Fig 3). Like rectus FFs, they can fill large defects. One or 2 skin paddles may be harvested, allowing a variety of uses, such as floor of mouth reconstruction (Fig 3) or large skull base defects.[5,7]

Fascial. Fascia-containing FFs in H&N are nearly exclusively fasciocutaneous, including a skin paddle in addition to the fascia, vessels, and subcutaneous tissue. Fasciocutaneous FFs can be used to restore skin or mucosal defects (Fig 4). Additionally, fasciocutaneous FFs can be tubed/rolled to recreate epithelial-lined conduits (Fig 5).[5,7,21] Two of the more commonly used fasciocutaneous FFs are the radial forearm and anterolateral thigh FFs.

The radial forearm FF has a rich vascular supply and may be harvested in a variety of sizes. It has been used more extensively and for a wider variety of reconstructions than any other flap.[5] This FF is particularly advantageous because the forearm skin is usually non-hair-bearing and the forearm is least influenced by obesity.[7] Primary indications for use include reconstruction of skin or mucosal lining defects, partial/hemiglossectomy defects, and pharyngeal defects. The radial forearm FF is harvested from the volar aspect of the forearm and includes the radial artery (Fig 4). Thus, performing an accurate preoperative Allen test is of the utmost importance to ensure an adequate supply of the hand via the ulnar artery and to avoid catastrophic ischemia of the hand.[5]

The anterolateral thigh (ALT) FF has a large, thin, pliable skin paddle and a long vascular pedicle, up to 15 cm. Similar to the

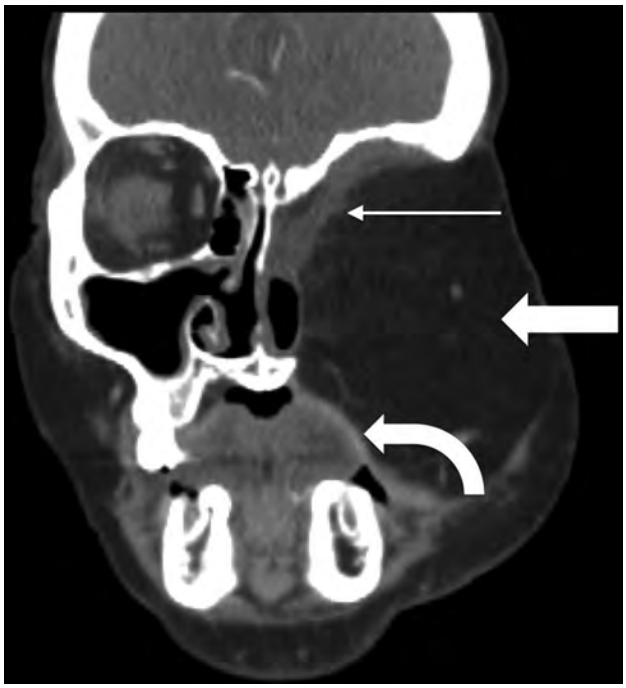


FIG 2. Rectus abdominis FF. Postoperative coronal CECT in a patient with T4bN0M0 basaloid carcinoma and recurrence after initial partial maxillectomy demonstrates the bulky rectus abdominis FF (*block arrow*) filling the left midface defect following orbital exenteration and total maxillectomy. The rectus muscle (*thin arrow*) lines the skull base defect, while a portion of the skin paddle recreates the oral cavity mucosal surface (*curved arrow*).

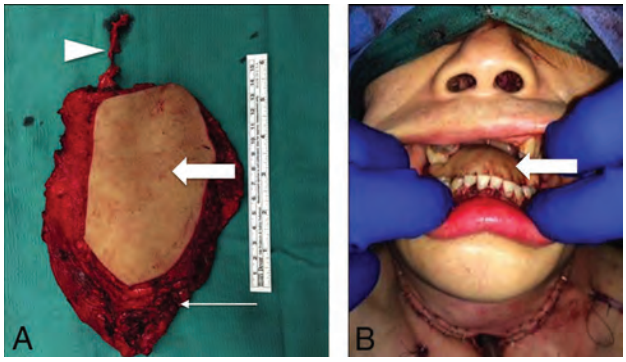


FIG 3. Latissimus dorsi FF. Following total glossectomy, right oropharyngectomy, and total laryngectomy for T4a squamous cell carcinoma (SCC) of the right oral tongue, the FF was harvested and set on the operating room back table (*A*) with the elongated vascular pedicle (*arrowhead*), skin (*block arrow*), and latissimus (*arrow*). The latissimus is fastened to secure the mandible and recreate the mylohyoid sling. Immediate postoperative picture (*B*) shows the skin paddle (*block arrow*) closing the glossectomy defect.

radial forearm FF, ALT reconstruction can be used for skin and mucosal lining defects, subtotal glossectomy defects (Fig 6), pharyngeal defects, and skull base defects. This FF can be tubed for pharyngoesophageal defects. Because the ALT FF is harvested from the anterior and proximal aspect of the lower extremity, there is relatively little morbidity at the donor site and the patient's clothing typically covers any postoperative scarring.[5,7]

Osseous. Osseous-containing FFs are some of the most complex FFs. They are mostly composite flaps, containing multiple different tissue types in addition to bone, including skin, subcutaneous tissue, fascia, and muscle (Fig 7). One of the surgical goals is to "replace like with like," meaning reconstructing surgical cavities with tissue constituents similar to those resected. Thus, osteocutaneous FFs have become popular solutions for many maxillary and mandibular defects.[5] Long segments of bone, up to 25 cm, may be harvested and crafted as a variety of surgical constructs using surgical plates and screws. Two of the most commonly used osteocutaneous FFs are the fibular and scapular FFs.

The tubular shape and thick cortical bone make the fibular FF particularly strong. This can be harvested with skin (free osteocutaneous) or without (free osseous). Up to 25 cm of the fibula may be resected with little effect on the gait as the fibula is a nonweight-bearing bone. The fibular FF is most commonly used for mandibular reconstruction (Fig 7). Most importantly preoperative evaluation must ensure that the peroneal artery does not supply the foot because this is the main supplying artery of the FF and will lead to foot ischemia if harvested.[5,22]

The scapular osteocutaneous FF has become a more popular option for complex midface reconstructions (Fig 8). The vascular pedicle is long, up to 14 cm, and the vessel diameters are large, 3–4.5 mm. Different segments of the bone can be harvested, including the scapular tip and up to 2 segments of the lateral border. Any of the segments can be fashioned to reconstruct the hard palate or orbital rim. One or 2 skin paddles can be obtained; unfortunately, they are often hair-bearing in male patients, resulting in an undesired postoperative cosmetic appearance if used in certain locations (ie, oral cavity mucosal reconstruction). The scapular skin paddles are particularly advantageous because they can be completely separate from the osseous component, providing the most freedom for 3D insetting of any composite FF.[5,7]

Visceral. Visceral FFs can be used in pharyngoesophageal reconstruction. The jejunal FF is especially useful given

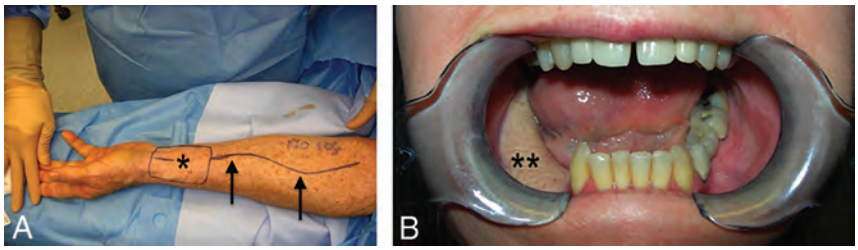


FIG 4. Radial forearm FF. Intraoperative photo (*A*) shows the radial artery (*arrows*) up to the skin paddle (*asterisk*) after Doppler mapping. Postoperative clinical picture (*B*) demonstrates the well-incorporated mature flap (*double asterisks*) following marginal mandibulectomy.

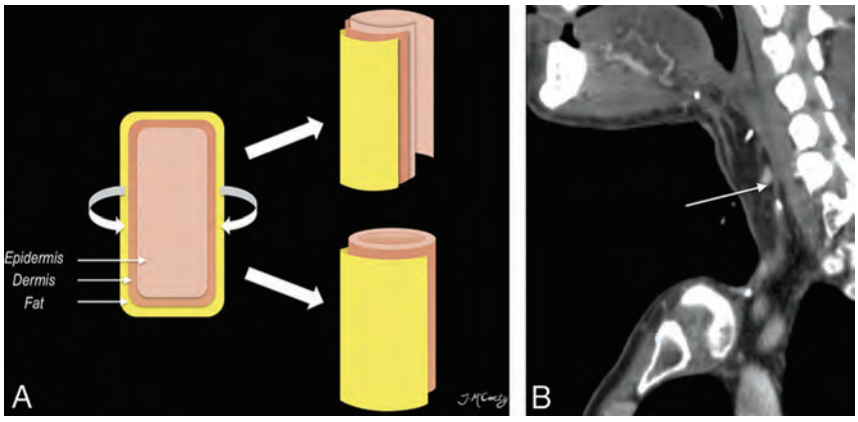


FIG 5. Tubed radial forearm FF. The diagram (*A*) shows that free flaps can be partially rolled (*upper right*) or completed tubed (*lower right*) to reconstruct the upper aerodigestive tract. Postoperative sagittal CECT (*B*) following base of tongue hemiglossectomy and laryngopharyngectomy shows the radial forearm FF reconstruction (*long arrow*).

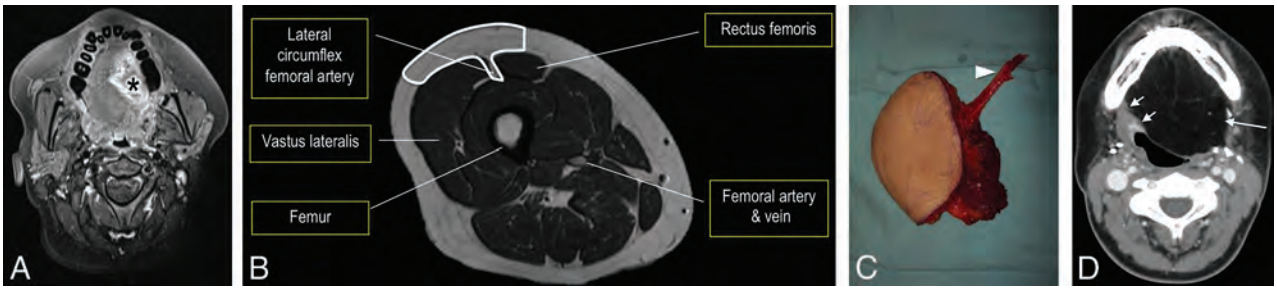


FIG 6. Anterolateral thigh FF. Axial fat-saturated T1 postcontrast MR imaging (*A*) demonstrates the T4a left lateral oral tongue SCC (*asterisk*), which also involved the extrinsic tongue muscles and floor of mouth. Line drawing (*B*) of the ALT donor site (*outlined*) axial anatomy includes the descending branch of the lateral circumflex femoral artery. The ALT has a large, thin, pliable skin paddle with relatively little morbidity at the donor site. ALT FF intraoperative image (*C*) shows the harvested FF on the operating room back table with an elongated vascular pedicle (*arrowhead*). Postoperative axial CECT (*D*), obtained 12 weeks after the operation, shows the inset homogeneous fatty tongue ALT FF (*arrow*) without induration or edema. The recipient site margins (*short arrow*) have no nodularity.

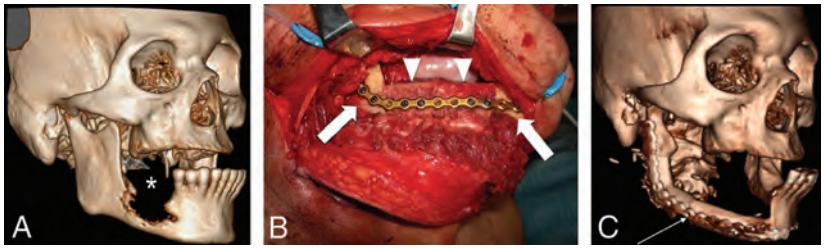


FIG 7. Fibular FF. Preoperative 3D-volume rendered CT (*A*) demonstrates the T4aN0 right oral cavity SCC (*asterisk*). Intraoperative photograph (*B*) (different patient but with a similar reconstruction) shows the surgical plate and bicortical screws used to fixate the inset fibular flap (*arrowheads*) to the residual native mandible (*block arrows*). Postoperative 3D-volume rendered CT (*C*) demonstrates the fibular FF reconstruction (*thin arrow*) following the right segmental mandibulectomy.

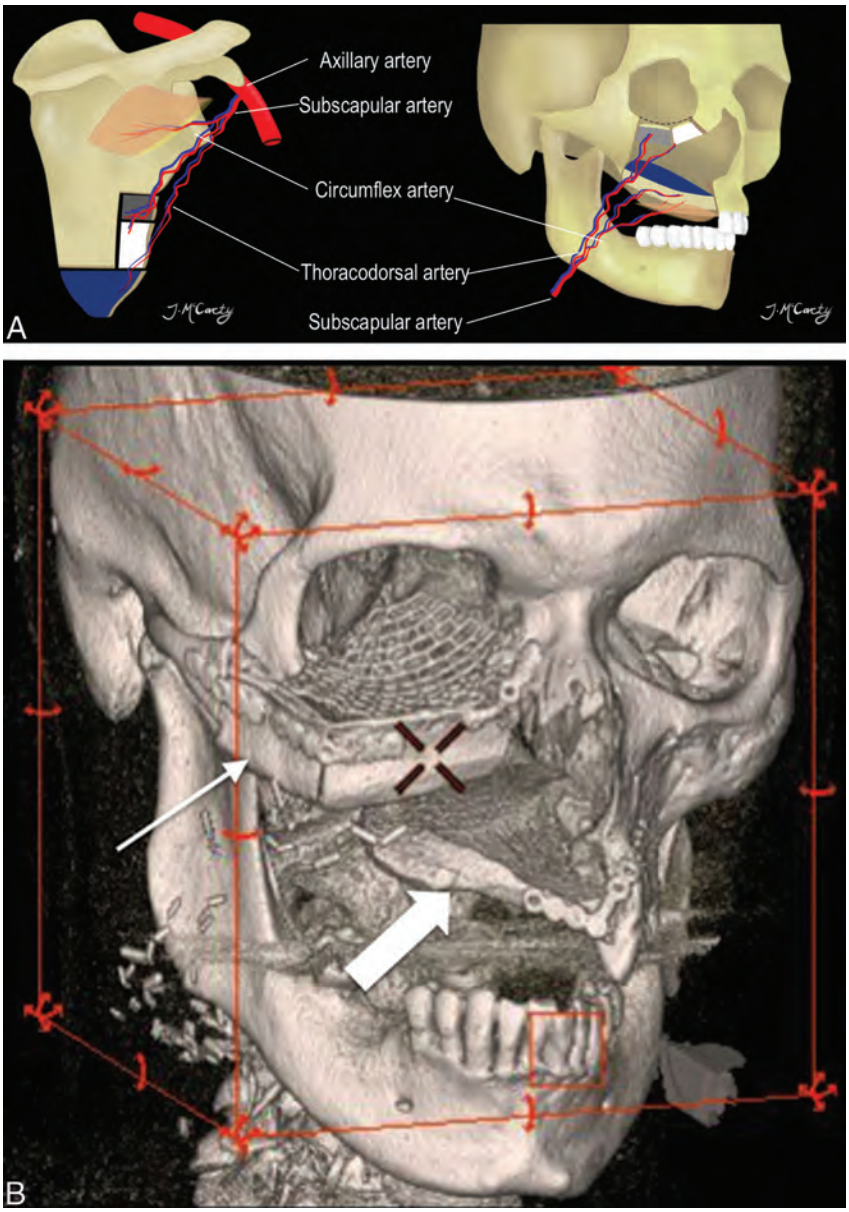


FIG 8. Scapular osteocutaneous FF. *A*, Diagram of the scapular donor site (*left*) and midface inset (*right*) demonstrates the versatile scapular osteocutaneous FF for complex midface defects. A variety of different bone shapes can be obtained depending on the contour of the defect. One or 2 skin paddles may be harvested; one usually recreates the oral mucosal surface as shown in the diagram. 3D-volume rendered CT (*B*) in a patient who is status post right maxillectomy shows the lateral scapular border reconstructing the right maxilla and orbital rim (*thin arrow*), while the scapular tip reconstructs the hard palate (*block arrow*).

its similar in caliber to the esophagus (Fig 9).[16] The pharynx is a larger caliber than the jejunal FF; thus, the cephalad border of the flap may be opened along the antimesenteric border to achieve a more suitable anastomosis.[5] For circumferential pharyngeal defects, jejunal and ALT FFs have slightly better functional outcomes than radial forearm FFs.[23]

Radiologic Evaluation following FF Reconstruction

Preoperative imaging should always be reviewed when the posttreatment scan is interpreted as it helps to understand what anatomic structures were resected and where the primary tumor was located. The CT or MR imaging appearance of the FF reconstruction reflects the flap components.

CT should always be performed with intravenous contrast administered in a standard amount and timing of the acquisition so that serial or surveillance scans can be compared. In our practice, the first posttreatment contrast-enhanced CT (CECT) is performed with PET to increase detection of persistent tumor, nodal, or distant metastases and to differentiate residual tumor from non-neoplastic postoperative changes. Imaging protocols include combined PET/CT from the skull vertex through the midthigh 1 hour after intravenous administration of 10–14 mCi of FDG. Helical noncontrast CT is performed before PET for attenuation correction and anatomic localization. A CECT of the neck with the arms down is performed following PET, using a split-bolus technique with 110 mL of intravenous iopamidol (Isovue-370; Bracco, Princeton, New Jersey), with 55 mL injected first at 2.5 mL/s, a 40-second delay, then another 55 mL at the same rate, and a total scan delay of 90 seconds. Axial images are acquired from the frontal sinuses through the mediastinum at a 1.25-mm section thickness and are sent to the PACS. Multiplanar reformations are also sent to the PACS.

The initial baseline posttreatment PET/CECT is performed 10–12 weeks after the end of radiation treatment, or after the operation, to allow posttreatment changes to resolve. Surveillance imaging timing has not been universally

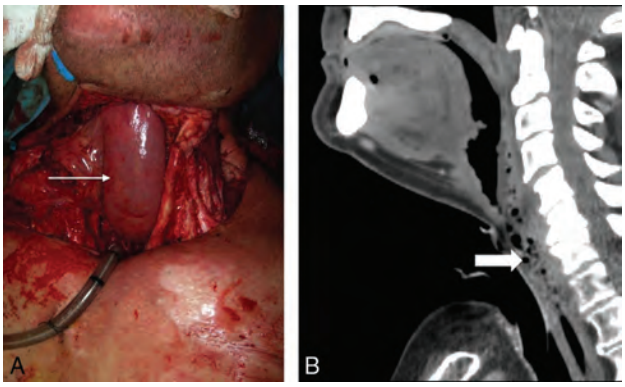


FIG 9. Jejunal FF. Following laryngopharyngectomy, the intraoperative photograph (*A*) shows a segment of the jejunum (*arrow*) interposed between the oropharynx and esophagus, anastomosed end to end to allow patients to eat and swallow. Sagittal postoperative CECT (*B*) shows the jejunal FF neopharynx (*block arrow*).

Table 3: Interpretation checklist

Checklist	What to Look for
Clinical note	Primary tumor and stage Type and date of FF reconstruction Most recent physical examination New symptoms
Flap appearance	Inspect the flap itself, ensure expected attenuation/signal with smooth non-nodular and non-mass-like enhancement
Surgical bed	Next inspect the FF margins; again, ensure expected attenuation/signal with smooth non-nodular and non-mass-like enhancement
Osteotomies	Nonunited or healed? If nonunited, ensure that the margins remain sharp

standardized, but our H&N cancer multidisciplinary group has a surveillance algorithm, and we are researching the optimal protocol.

MR imaging may be useful for evaluation of perineural tumor, intracranial extension, cartilaginous involvement, and other troubleshooting. MR imaging protocols vary depending on the location of the abnormality. Our H&N MR imaging protocols generally have the following sequences in common: 3-plane non-fat-saturated precontrast T1-weighted, axial fat-saturated precontrast T2-weighted, and axial and coronal postcontrast fat-saturated T1-weighted imaging, following intravenous administration of 0.1 mmol/kg of gadobenate dimeglumine (MultiHance; Bracco). Again, a standardized protocol involving the same timing and sequences is essential.

An organized approach to postreconstruction imaging interpretation helps make a complex study easier to understand (Table 3). The first steps are to determine the location and appearance of the primary malignancy, which tissues were removed at the time of the oncologic operation, and what type of FF was used to reconstruct the resulting defect. Then, the FF itself is evaluated. Free flaps contain a combination of muscle, skin, fascia, fat, and bone. The bone should be well-corticated without erosion or destruction. The osseous interface with native bone in the mandible, maxilla, or orbital walls should be assessed for bridging new bone (Fig 10). If a plate and screws have been placed at the flap–native bone interface, there may be diastasis bridged by the plate, but the

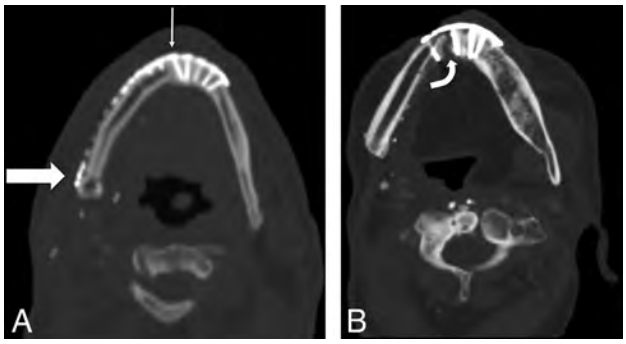


FIG 10. Osseous FF margins. Axial noncontrast CT shows the expected postoperative appearance (*A*) following right segmental mandibulectomy and fibular FF reconstruction, with a healing symphyseal interface (*thin arrow*) and a nonunited-but-sharp osteotomy underlying the posterior mandibular body surgical plate (*thick arrow*). Axial noncontrast CT of a different patient with a fibular FF shows a complicated postoperative appearance (*B*), with a nonunited, diastatic, irregular symphyseal margin with a periosteal reaction in this patient with osteonecrosis of the fibular FF and native mandible.

cut end bone margins should be smooth. Be sure the plate has not elevated from the bone and that the screws maintain the plate, without periscrew lucency, which would imply loosening or infection.

The fatty portion of the flap should be relatively homogeneous without induration, nodularity, or abnormal focal enhancement (Fig 11). The interface of the flap and resection cavity, known as the recipient bed, is the most critical area to examine because this is the site of local disease recurrence. Nodularity, a mass, or focal discrete enhancement is a characteristic imaging appearance of a recurrence (Fig 12). These findings are especially important if the patient has new pain, dysphagia, or any symptom that would suggest recurrent malignancy. Multiple clips are usually present at the vascular pedicle, denoting the anastomosis between the flap and resection cavity, but they are small and rarely degrade image quality.

Muscular flap components are usually striated, thin, and relatively flat (Fig 11). On CECT, the flap muscle is isodense to striated muscle elsewhere. The MR signal intensity and enhancement pattern have been described and are predictable, with moderate-to-intense enhancement.[13,24-26] In the early posttreatment period, the flap may be edematous, hypointense on T1-weighted images (Fig 13), hyperintense on T2-weighted images, and enhance with gadolinium contrast. Later, the muscular portion of the flap decreases in bulk and becomes heterogeneous on T1- and T2-weighted MR images and relatively more hypoattenuating on CT as the denervated muscle becomes fatty.[21,27]

Complications

Radiologists will encounter cross-sectional imaging studies performed to evaluate postoperative complications in patients with FF. Complications have been divided into early or late,[28] but in actuality, there is overlap between the 2 categories.

The survival rate of a flap is around 95%, but early complications soon after a reconstructive FF operation include ischemia, infection, bleeding, and dehiscence.[29] These are closely moni-

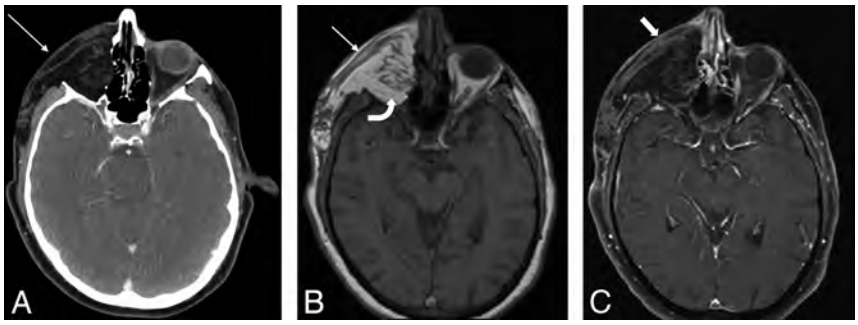


FIG 11. Expected FF fat and muscle appearance. The postoperative axial CECT (*A*) following orbital exenteration and latissimus FF reconstruction shows the normal thin musculature (*arrow*) and fat of the flap. Axial T1 precontrast MR image (*B*) shows the FF with muscular striations (*arrow*). The FF fat (*curved arrow*) deep to the muscular component is homogeneously hyperintense. Axial postcontrast fat-saturated T1 MR image (*C*) shows the FF muscular thin, non-nodular enhancement (*block arrow*), similar to that in other muscles in the H&N.

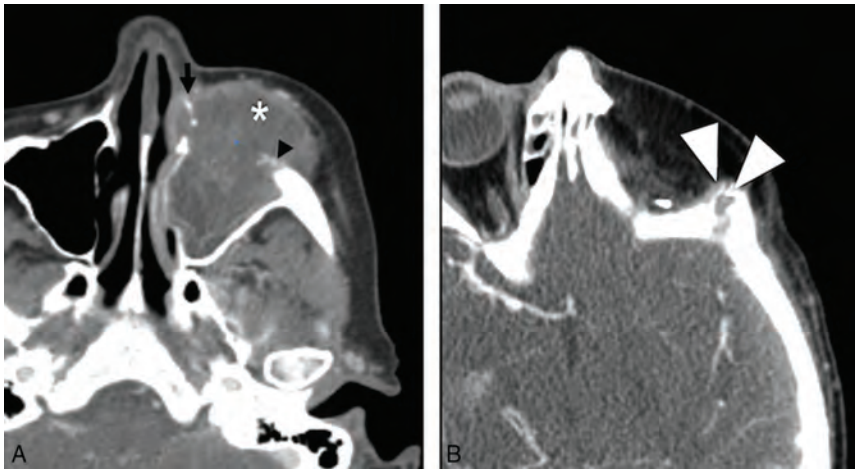


FIG 12. Tumor recurrence. Preoperative axial CECT (*A*) shows a T4a maxillary sinus SCC. Following maxillectomy and orbital exenteration with ALT FF reconstruction, the patient had a biopsy-proved recurrence (*white arrowheads*) at the margins of the ALT FF on postoperative CECT (*B*) several months later.

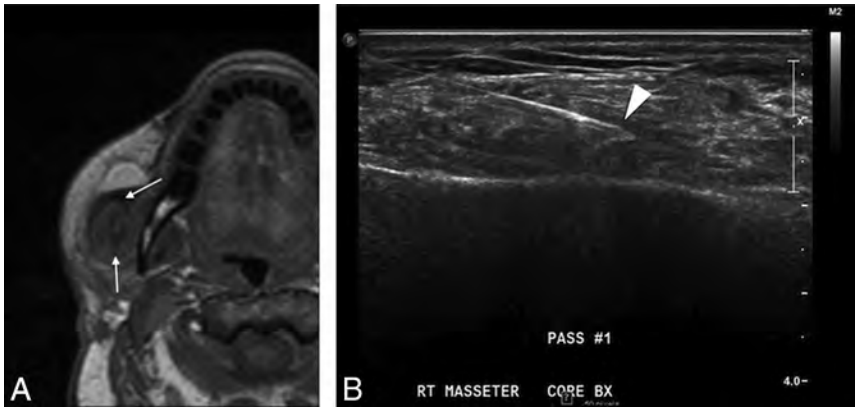


FIG 13. Inflammation. Postoperative T1 axial MR image (*A*) after reconstruction with a latissimus FF shows hypointense abnormal signal (*arrows*) and enlargement of the right masseter. Considerations included myositis, denervated muscle, or recurrent intramuscular tumor. Ultrasound-guided biopsy (*B*) shows good positioning of the needle tip in the muscle (*arrowhead*). Final pathology results were benign skeletal muscle and fibroadipose tissue, consistent with focal inflammation; no malignant cells were present.

tored clinically at the bedside and rarely require cross-sectional imaging. Ischemia of FFs may be due to either venous or arterial thrombosis.[8] The surgical team is on high alert for signs of FF ischemia in the immediate postoperative period. Patients undergo frequent inspection of tissue color, capillary refill, turgor, and temperature.[12,30] Various other techniques, including Doppler monitoring and needle pricks, are also used in the immediate postoperative period. When recognized and surgically treated promptly, compromised FFs have a salvage rate of 50%–75%.[8] Management includes re-exploration of the site, with possible thrombectomy and anastomosis revision.

More commonly imaged complications occur later in the postoperative period and include infection, fistulas, hardware exposure, and osteonecrosis. Infection and fistulas can present with nonspecific imaging features, such as soft-tissue swelling and stranding, loss of fat planes, and collections of fluid and air. In patients with H&N cancer with a history of radiation therapy or a recent operation, unless baseline postoperative studies are available, it may be impossible to distinguish infection from treatment-related changes solely on the basis of imaging alone. However, some cases may be more obvious, with new rim-enhancing fluid collections or areas of frank dehiscence (Fig 14).

Dehiscence of FFs overlying surgical hardware can result in exposure of the surgical construct (Fig 15). Hardware exposure and extrusion are the most commonly cited flap complications and occur in around 15% of patients.[29,31] This complication is often seen in association with continued tobacco use.

Osteonecrosis is primarily a clinical diagnosis and is seen in patients with exposed bone.[32] CT features of cortical destruction, trabecular disorganization, periosteal reaction (Fig 10*B*), and associated soft-tissue abnormality overlap findings of osteomyelitis and tumor recurrence.[33,34] CT is usually performed not to differentiate etiologies but to determine the extent of disease. Osteonecrosis is treated surgically, while osteomyelitis will usually be treated medically. Biopsy is generally avoided in cases of bisphosphonate osteonecrosis because it may cause progression and further damage but can be definitive in differentiating osteoradionecrosis from tumor recurrence.[34,35] One useful distinguishing imaging finding is that abnormalities at a site distant

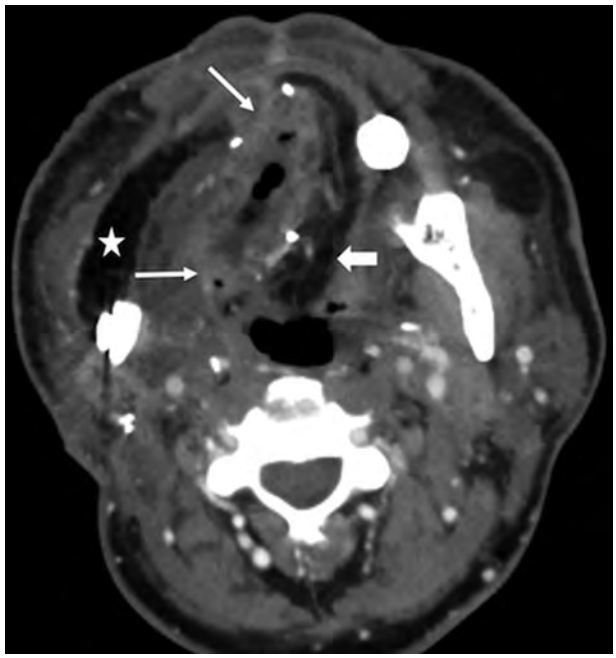


FIG 14. FF Infection. This patient had fever, elevated white blood cell count, and purulent left neck drainage following radial forearm FF and fibular FF reconstruction for T4aN2c SCC of the right oral tongue. Axial CECT shows an abscess with flap induration, irregular enhancement, and pockets of gas in the right floor of mouth (*arrows*). Note stranding in the fatty flap (*block arrow*), compared with the homogeneous fat density in the lateral flap that reconstructed the buccal mucosa (*star*).

FIG 15. Hardware exposure. A patient with T4aN0M0 left floor of mouth SCC status post pectoralis rotational flap and surgical bar reconstruction. A fibular FF reconstruction was originally planned but abandoned due to severe peripheral vascular disease. The patient was lost to follow-up for 2 years and then presented with a 1-month history of hardware exposure with a large area of exposed mandibular hardware (*arrow*) on axial CECT (*A*) and clinical examination (*B*).

or contralateral to the primary tumor are more likely osteoradionecrosis.[34]

Ossification of the vascular pedicle, while not a surgical complication, is an imaging pitfall. This entity presents in patients with fibular FF as a thin, linear, or curvilinear ossific density following the course of the vascular pedicle (Fig 16), the result of ossification of a strip of periosteum inset with the vascular pedicle. This ossification can be seen in up to 50% of patients as soon as 1 month after the operation and may present clinically as a palpable mass.[36]

Surveillance for tumor recurrence is the main focus of follow-up imaging. Primary site recurrences most often occur at the

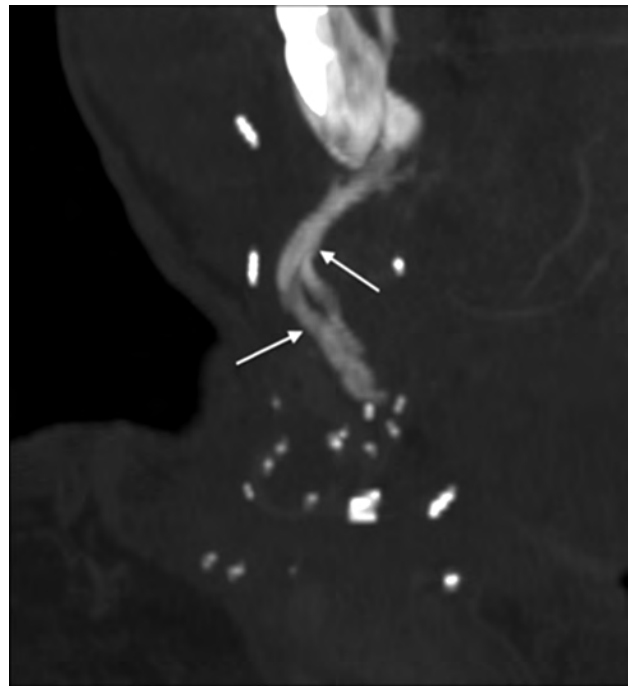


FIG 16. Vascular pedicle ossification. A patient with fibular FF following right mandibulectomy for T2N1 SCC of the mandibular gingiva returned 8 weeks after the operation with a palpable right-neck mass. Oblique coronal MIP reconstruction from CECT, bone windows, shows linear ossification (*arrows*) along the course of vascular pedicle, corresponding to the palpable abnormality.

margins of the resection at the flap–native tissue interface and have nodular or masslike enhancement with signal characteristics similar to those of the original tumor (Fig 12).[37-39] Nodal recurrence may have the typical expected regional distribution of the original tumor or may vary from the norm due to surgical alteration of drainage pathways.[28] The third recurrence pattern is perineural spread of disease. If one is not actively assessing perineural tumor, the findings may be subtle enough to evade detection.

CONCLUSIONS

Postoperative imaging in patients with H&N reconstruction is challenging. An organized approach and thorough understanding of FF appearances and complications will help interpreting radiologists provide accurate, useful imaging reports for both the patients and their referring clinicians.

Disclosures: Amanda S. Corey—*UNRELATED*: *Consultancy*: Rad MC LLC; *Payment for Lectures Including Service on Speakers Bureaus*: The Osler Institute, *Comments*: Continuing Medical Education lectures. Patricia A. Hudgins—*UNRELATED*: *Royalties*: Elsevier/Amirsys, *Comments*: medical textbook authorship.

REFERENCES

1. Urken ML, Weinberg H, Buchbinder D, et al. **Microvascular free flaps in head and neck reconstruction: report of 200 cases and review of complications.** *Arch Otolaryngol Head Neck Surg* 1994;120:633–40 CrossRef Medline
2. Blackwell KE. **Unsurpassed reliability of free flaps for head and neck reconstruction.** *Arch Otolaryngol Head Neck Surg* 1999;125:295–99 CrossRef Medline
3. Corbitt C, Skoracki RJ, Yu P, et al. **Free flap failure in head and neck reconstruction.** *Head Neck* 2014;36:1440–45 CrossRef Medline

4. Su WF, Hsia YJ, Chang YC, et al. **Functional comparison after reconstruction with a radial forearm free flap or a pectoralis major flap for cancer of the tongue.** *Otolaryngol Head Neck Surg* 2003;128: 412–18 CrossRef Medline
5. Urken ML. *Atlas of Regional and Free Flaps for Head and Neck Reconstruction.* New York: Raven Press; 1995
6. Ferlay J, Soerjomataram I, Dikshit R, et al. **Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012.** *Int J Cancer* 2015;136:E359–86 CrossRef Medline
7. Conley J, Patow C, eds. *Flaps in Head & Neck Surgery.* 2nd ed. New York: Thieme; 1989
8. Goldenberg D, Goldstein BJ, eds. *7.3 Facial Reconstruction.* Stuttgart: Thieme; 2011
9. Resto VA, Deschler DG. **Complications of skin, cartilage, and bone grafts.** In: Eisele DW, Smith RV. *Complications in Head and Neck Surgery.* 2nd ed. Philadelphia: Elsevier Mosby; 2009:803–11
10. Weiland AJ, Phillips TW, Randolph MA. **Bone grafts: a radiologic, histologic, and biomechanical model comparing autografts, allografts, and free vascularized bone grafts.** *Plast Reconstr Surg* 1984;74: 368–79 CrossRef Medline
11. Sparks DS, Wagels M, Taylor GI. **Bone reconstruction: a history of vascularized bone transfer.** *Microsurgery* 2018;38:7–13 CrossRef Medline
12. Wester DJ, Whiteman ML, Singer S, et al. **Imaging of the postoperative neck with emphasis on surgical flaps and their complications.** *AJR Am J Roentgenol* 1995;164:989–93 CrossRef Medline
13. Hudgins PA. **Flap reconstruction in the head and neck: expected appearance, complications, and recurrent disease.** *Semin Ultrasound CT MR* 2002;23:492–500 CrossRef Medline
14. Cummings CW, Haughey BH, Thomas JR. *Cummings Otolaryngology: Head and Neck Surgery.* St. Louis: Mosby; 2005
15. Chim H, Salgado CJ, Seselgyte R, et al. **Principles of head and neck reconstruction: an algorithm to guide flap selection.** *Semin Plast Surg* 2010;24:148–54 CrossRef Medline
16. Wong CH, Wei FC. **Microsurgical free flap in head and neck reconstruction.** *Head Neck* 2010;32:1236–45 CrossRef Medline
17. Syed F, Spector ME, Cornelius R, et al. **Head and neck reconstructive surgery: what the radiologist needs to know.** *Eur Radiol* 2016;26: 3345–52 CrossRef Medline
18. Wei FC, Jain V, Suominen S, et al. **Confusion among perforator flaps: what is a true perforator flap?** *Plast Reconstr Surg* 2001;107: 874–76 CrossRef Medline
19. Howard BE, Nagel TH, Barrs DM, et al. **Reconstruction of lateral skull base defects: a comparison of the submental flap to free and regional flaps.** *Otolaryngology Head Neck Surg* 2016;154:1014–18 CrossRef Medline
20. Netter FH. *Atlas of Human Anatomy.* Philadelphia: Elsevier; 2010
21. Saito N, Nadgir RN, Nakahira M, et al. **Posttreatment CT and MR imaging in head and neck cancer: what the radiologist needs to know.** *RadioGraphics* 2012;32:1261–82; discussion 1282–94 CrossRef Medline
22. Kelly AM, Cronin P, Hussain HK. **Preoperative MR angiography in free fibula flap transfer for head and neck cancer: clinical application and influence on surgical decision making.** *AJR Am J Roentgenol* 2007;188:268–74 CrossRef Medline
23. Welkoborsky HJ, Deichmüller C, Bauer L, et al. **Reconstruction of large pharyngeal defects with microvascular free flaps and myocutaneous pedicled flaps.** *Curr Opin Otolaryngol Head Neck Surg* 2013; 21:318–27 CrossRef Medline
24. Chong J, Chan LL, Langstein HN, et al. **MR imaging of the muscular component of myocutaneous flaps in the head and neck.** *AJNR Am J Neuroradiol* 2001;22:170–74 Medline
25. Hudgins PA, Burson JG, Gussack GS, et al. **CT and MR appearance of recurrent malignant head and neck neoplasms after resection and flap reconstruction.** *AJNR Am J Neuroradiol* 1994;15:1689–94 Medline
26. Ginsberg LE. **Imaging pitfalls in the postoperative head and neck.** *Semin Ultrasound CT MR* 2002;23:444–59 CrossRef Medline
27. Fox MG, Bancroft LW, Peterson JJ, et al. **MRI appearance of myocutaneous flaps commonly used in orthopedic reconstructive surgery.** *AJR Am J Roentgenol* 2006;187:800–06 CrossRef Medline
28. Garcia MR, Passos UL, Ezzedine TA, et al. **Postsurgical imaging of the oral cavity and oropharynx: what radiologists need to know-erratum.** *RadioGraphics* 2015;35:1624 CrossRef Medline
29. Haughey BH, Wilson E, Kluwe L, et al. **Free flap reconstruction of the head and neck: analysis of 241 cases.** *Otolaryngol Head Neck Surg* 2001;125:10–17 CrossRef Medline
30. Kubo T, Yano K, Hosokawa K. *Management of Flaps with Compromised Venous Outflow in Head and Neck Microsurgical Reconstruction.* Vol 22. Wiley Online Library; 2002:391–95. https://onlinelibrary.wiley.com/doi/abs/10.1002/micr.10059. Accessed February 20, 2018
31. Day KE, Desmond R, Magnuson JS, et al. **Hardware removal after osseous free flap reconstruction.** *Otolaryngol Head Neck Surg* 2014; 150:40–46 CrossRef Medline
32. Estilo CL, Van Poznak CH, Wiliams T, et al. **Osteonecrosis of the maxilla and mandible in patients with advanced cancer treated with bisphosphonate therapy.** *Oncologist* 2008;13:911–20 CrossRef Medline
33. Morag Y, Morag-Hezroni M, Jamadar DA, et al. **Bisphosphonate-related osteonecrosis of the jaw: a pictorial review.** *RadioGraphics* 2009;29:1971–84 CrossRef Medline
34. Chong J, Hinckley LK, Ginsberg LE. **Masticator space abnormalities associated with mandibular osteoradionecrosis: MR and CT findings in five patients.** *AJNR Am J Neuroradiol* 2000;21:175–78 Medline
35. Ruggiero SL, Fantasia J, Carlson E. **Bisphosphonate-related osteonecrosis of the jaw: background and guidelines for diagnosis, staging and management.** *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 2006;102:433–41 CrossRef Medline
36. Glastonbury CM, van Zante A, Knott PD. **Ossification of the vascular pedicle in microsurgical fibular free flap reconstruction of the head and neck.** *AJNR Am J Neuroradiol* 2014;35:1965–69 CrossRef Medline
37. Lall C, Tirkes TA, Patel AA, et al. **Flaps, slings, and other things: CT after reconstructive surgery–expected changes and detection of complications.** *AJR Am J Roentgenol* 2012;198:W521–33 CrossRef Medline
38. Offiah C, Hall E. **Post-treatment imaging appearances in head and neck cancer patients.** *Clin Radiol* 2011;66:13–24 CrossRef Medline
39. Tomura N, Watanabe O, Hirano Y, et al. **MR imaging of recurrent head and neck tumours following flap reconstructive surgery.** *Clin Radiol* 2002;57:109–13 CrossRef Medline

The ASNR-ACR-RSNA Common Data Elements Project: What Will It Do for the House of Neuroradiology?

A.E. Flanders and J.E. Jordan

ABSTRACT

SUMMARY: The American Society of Neuroradiology has teamed up with the American College of Radiology and the Radiological Society of North America to create a catalog of neuroradiology common data elements that addresses specific clinical use cases. Fundamentally, a common data element is a question, concept, measurement, or feature with a set of controlled responses. This could be a measurement, subjective assessment, or ordinal value. Common data elements can be both machine- and human-generated. Rather than redesigning neuroradiology reporting, the goal is to establish the minimum number of "essential" concepts that should be in a report to address a clinical question. As medicine shifts toward value-based service compensation methodologies, there will be an even greater need to benchmark quality care and allow peer-to-peer comparisons in all specialties. Many government programs are now focusing on these measures, the most recent being the Merit-Based Incentive Payment System and the Medicare Access Children's Health Insurance Program Reauthorization Act of 2015. Standardized or structured reporting is advocated as one method of assessing radiology report quality, and common data elements are a means for expressing these concepts. Incorporating common data elements into clinical practice fosters a number of very useful downstream processes including establishing benchmarks for quality-assurance programs, ensuring more accurate billing, improving communication to providers and patients, participating in public health initiatives, creating comparative effectiveness research, and providing classifiers for machine learning. Generalized adoption of the recommended common data elements in clinical practice will provide the means to collect and compare imaging report data from multiple institutions locally, regionally, and even nationally, to establish quality benchmarks.

ABBREVIATIONS: ACR = American College of Radiology; AI = artificial intelligence; ASNR = American Society of Neuroradiology; BI-RADS = Breast Imaging Reporting and Data System; CDE = common data element; EHR = Electronic Health Record; IT = information technology; LGG = low grade glioma; PQRS = physician quality reporting system; RSNA = Radiological Society of North America; TCGA = The Cancer Genome Atlas; TCIA = The Cancer Imaging Archive; VASARI = Visually AcceSsible Rembrandt Images

The ASNR-ACR-RSNA Common Data Elements (CDEs) Project represents a collaboration between the American Society of Neuroradiology (ASNR), the American College of Radiology (ACR), and the Radiological Society of North America (RSNA). The function of this workgroup is to develop a catalog of CDEs for neuroradiology that are both practical and useful for clinical practice, with the goals of unifying practice standards by improving consistency in reporting and developing human- and machine-interpretable features that can be used to measure quality in our specialty. There are numerous secondary benefits in comparative effectiveness research, precision medicine, radiomics, registry participation, machine learning, communication, and public health. This joint committee was formed to catalog, unify, and codify known existing neuroradiology findings, observations, and concepts commonly used in neuroradiology. The following is a brief overview of the rationale and processes behind this collaborative effort and the potential benefits to our profession and patient care.

Despite the advances in information technology that are ubiquitous in our profession, the process for composing the radiology report has changed little in the past 100 years; reports are largely prose descriptions of findings.[1] The consumer of the prose report is obligated to extract its concepts to drive clinical decision-making. There is no author obligation to use consistent terminology when generating a report. This situation creates myriad problems, including challenges in comparing studies or aggregating reports of the same type produced by different authors. Re-review of the original images is often the only practical solution in either case. While there have been several notable efforts to create consistency

Received May 24, 2018; accepted after revision June 28.

From the Department of Radiology/Neuroradiology (A.E.F.), Thomas Jefferson University Hospital, Philadelphia, Pennsylvania; and Standards and Guidelines Committee for the American Society of Neuroradiology (J.E.J.), Rancho Palos Verdas, California.

Please address correspondence to Adam E. Flanders, MD, Thomas Jefferson University Hospital, Suite 1080B Main Building, 132 S. Tenth Street, Philadelphia, PA 19107; e-mail: adam.flanders@jefferson.edu; @BFlanksteak

http://dx.doi.org/10.3174/ajnr.A5780

in reporting styles through structured or standardized reporting (RSNA Informatics Reporting: RadReport; radreport.org) and by the creation of a vendor-neutral standard for a report template data schema and exchange (Management of Radiology Report Templates; https://docs.imphub.org/display/ITMT/MRRT+Documentation),[2,3] little attention has been paid to developing consistent representation of the intrinsic concepts contained in the report that drive clinical decisions. This deficiency is at the core of the common data elements effort.

What is a Common Data Element?

Fundamentally, a CDE is a question/concept combined with a set of expected responses. A CDE is the most granular statement or observation that one can provide in a report. It is a single accepted concept with a response. The important characteristic is that both the concept and the response are consistent whenever it is used. CDEs can record properties of imaging findings such as anatomic location, shape, image number, image coordinates, and dimensions and can store computed values such as texture metrics.[4] Machine-generated values that are subsequently inserted into a radiology report could (eg, from a sonography device or postprocessing workstation) also represent CDEs. The response could be Boolean (eg, yes or no), quantitative (eg, 1, 2.3, 5.01), ordinal (eg, A, B, C1, D6), or a list of consistent terms/phrases. A report might contain many CDEs or sets of CDEs that are relevant to a specific disease. A brain MR imaging evaluation for multiple sclerosis, for example, might include a CDE set related to specific plaque characteristics (eg, number, location, features, size, enhancement). Sets of CDEs could be incorporated into a report on approval of the radiologist based on specific circumstances. For example, reporting an incidental laryngeal mass on a neck CTA could be improved by automatically importing a laryngeal mass CDE subset into a CTA report template. CDE sets can be used once or reused in other clinical contexts.

Use of a controlled response creates uniformity in activities such as clinical reporting for the human consumer, but it also creates an environment that facilitates computable consumption of concepts that can drive downstream actionable processes.[4] Additional benefits include diminished ambiguity, increased acceptance by clinicians, modular authoring, and modification of report templates. Examples might include an ASPECTS for stroke (http://www.aspectsinstroke.com/) (integer range: 0–10), Pfirrmann grade for disc degeneration (https://www.researchgate.net/publication/5840284_Modified_Pfirrmann_Grading_System_for_Lumbar_Intervertebral_Disc_Degeneration) (integer range: 1 – 8), or foraminal stenosis (text: normal, mild, moderate, severe). There are many examples of CDEs in our literature that correlate to outcomes, therapeutic response, and disease state. In most instances, CDEs are concepts that are already familiar to the practicing radiologist and clinician; they need not be obscure, complex, or uncommon. CDEs can also be used in both prose reporting and structured reporting.

The concept of CDEs should sound familiar because radiologists have been using them in various forms for years. The Breast Imaging Reporting and Data System (ACR BI-RADS Atlas 5th Edition; https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Bi-Rads) is the first clinical progenitor of CDEs. BI-RADS is focused on the probability of malignancy (0–6) using a global assessment category (eg, shape, margin, density of masses, calcifications, and so forth, which are part of a controlled terminology). Paramount to the generation of a BI-RADS global assessment score is the dependency on the component features/observations. The "RADS" construct has increased in popularity in recent years in other areas such as LI-RADS (Liver Reporting and Data System; https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/LI-RADS), PI-RADS (Prostate Imaging Reporting and Data System; https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/PI-RADS), TI-RADS (Thyroid Imaging Reporting and Data System; https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/TI-RADS), NI-RADS (Neck Imaging Reporting and Data System; https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/NI-RADs), and HI-RADS (Head Injury Imaging Reporting and Data System; https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/HI-RADS).[4] Compliance with a single terminology facilitates aggregation of data from multiple facilities and increases the value of our reports, including at points of care.[4] Related initiatives that are tied to compliance and payment include the Physician Quality Reporting System (https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/index.html) reporting measures for the Centers for Medicare and Medicaid Services.

Many published CDE sets originated through clinical trials research and have reasonable interrater agreement. The National Institute of Neurological Disorders and Stroke, for example, maintains a catalog of imaging-based CDEs for research purposes (https://www.commondataelements.ninds.nih.gov). The National Institute of Neurological Disorders and Stroke collection has imaging CDEs relevant to traumatic brain injury, stroke, multiple sclerosis, spinal cord injury, Parkinson disease, and others. The Visually AcceSsible Rembrandt Images (VASARI; https://radiopaedia.org/articles/vasari-mri-feature-set) collection of The Cancer Imaging Archive (https://www.cancerimagingarchive.net/) is the most comprehensive set of visual features that have been used to describe human gliomas on baseline MR imaging studies. The multicenter, federated The Cancer Genome Atlas (TCGA; https://wiki.cancerimagingarchive.net/display/Public/TCGA+Glioma+Phenotype+Research+Group) Glioma Phenotype Research Group collected and validated the most useful imaging features culled from the known literature on gliomas. The group developed the VASARI feature set using a large set of baseline glioblastoma and low-grade glioma (LGG) imaging studies stored in The Cancer Imaging Archive (TCIA). These phenotypic imaging data were successfully correlated with gene-expression data derived from the tumors in TCGA. The 25 features contained in the VASARI collection are all concepts familiar to neuroradiologists (eg, cyst, necrosis, enhancement, and so forth).[5] A subset of the VASARI features that demonstrates value in predicting tumor genomics or survival could be incorporated into a CDE module for clinical reporting.

While substantial effort by domain experts has gone into cataloging and validating these collections for research, there has hardly been any adoption of these very valuable observations into clinical reporting until very recently. Moreover, most of the CDEs

in existence took initial form as part of research initiatives and subsequently were never used once the research was completed. It is now well-recognized that there is substantial value in resurrecting many of these visually derived imaging features that were originally applied to address a specific research question and adapt them for clinical use. Some practices have taken on the task of incorporating CDEs to enhance the quality of local reporting practices. Mamlouk et al[6] reported on their very successful collaborative effort to disseminate consistent contextual reporting templates for neuroradiology examinations in a large multicenter practice. Over 50 specific use-case neuroradiology reporting templates were created. They describe a formal process in which templates are proposed and adjudicated by a panel that includes clinical input before dissemination to all radiologists.[6]

Why Do Common Data Elements Matter Now?

US health care is at a crossroads in which each specialty is being asked to define its inherent value in the patient care continuum. Health care organizations and subspecialty provider organizations are being asked to develop, benchmark, and comply with specific quality standards. Pay-for-performance initiatives, meaningful use, and Physician Quality Reporting System programs are now being wrapped up into the new value-based programs under the Merit-Based Incentive Payment System and the Medicare Access and Children's Health Insurance Program Reauthorization Act of 2015 (MIPS/MACRA: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/MACRA-MIPS-and-APMs.html). The radiology value chain considers the importance of clear and accurate reporting and report communications to physicians, patients, and other stakeholders. The ACR has also identified these areas as potential value-based payment metrics. In addition to clarity of reporting, structured reporting, standard lexicon, and language are key elements of the value-based payment metrics proposal.[7] Because referring physicians have shown a preference for structured reporting in contrast to conventional radiology reports, CDEs will likely play a key role in service to the goals of value-based reporting.[8]

The dissemination of electronic medical records created by the American Recovery and Reinvestment Act of 2009 stimulus package has facilitated the capability of collecting, sharing, and disseminating data.[9] Thus, various clinical subspecialties (eg, cardiology, gastroenterology, pathology, ophthalmology, oncology) have been very active in defining clinical concepts and reporting elements for the electronic medical records that can be readily mined to establish quality parameters and benchmark quality, thereby demonstrating the value of the services delivered. Clinical use of CDEs fosters participation in data registries, which, in turn, are being used to benchmark practice performance. The field of radiology has led the way in health care information technology (IT), interoperability, and information exchange, yet our field remains behind in standardizing quality measures for radiology reporting. With the exception of mammography, most of our quality metrics have focused on service delivery, workflow metrics, and payment policy and less on the content of our reports. Nevertheless, accuracy of reporting is a quintessential value metric of what radiologists offer, and CDEs should be viewed as a powerful tool to enhance the quality of our reports and actionable information.[10]

How the Currently Available CDEs Were Created

At an Imaging Informatics summit of the Radiological Society of North America Radiology Informatics Committee and the American College of Radiology Imaging Technology and Informatics Committee, discussions focused on the relative absence of codified observations/findings in radiology and a structure for representing them in our IT platforms. A collaboration was started to help fill this void. The initial objectives of this collaboration were to generate a common data model and syntax for representing reporting concepts that could interoperate with existing reporting technologies and extend their capabilities. An on-line repository (RadElement.org; http://www.radelement.org/) was built to house some of the limited existing content (eg, LI-RADS, PI-RADS). Each concept and controlled response stored in the repository is assigned a unique identifier used in downstream IT systems to retain the fidelity of the concept and response. Having set the stage, groups of domain experts could begin to create content to populate the repository, validate the concepts, and develop modifications.

Similar to the related efforts in terminology (RadLex; http://www.rsna.org/RadLex.aspx) and reporting (RadReport), cataloguing CDEs requires enlisting the knowledge of domain experts to ensure that relevant content is being included. For both the RadLex and RadReport efforts, the American Society of Neuroradiology was the first subspecialty organization to volunteer to help the RSNA to create repositories of neuroradiology-/ear, nose, throat–specific terminology and report templates, respectively. The ASNR has again volunteered to be the first subspecialty society to lend their expertise to this new CDE effort under the auspices of the ASNR Standards and Guidelines Committee. The neuroradiology effort is taking a pragmatic approach by developing CDEs for common clinical use cases that a neuroradiologist encounters every day rather than attempting to encompass all diseases in our specialty. This will help to inform us on how to design a process for authoring, vetting, editing, and publishing content in an efficient manner. The workgroup's charge is to compile only the most essential concepts for each clinical use case and to avoid making the lists comprehensive or exhaustive. By limiting the sets to the most essential concepts, the CDE sets become more practical, modular, and easier to use in practice and to incorporate into a report. The intent is to replicate what is taught in the training environment, whereby a neuroradiology attending physician might recite to a trainee the few key concepts that must be conveyed in a clinically useful report. Ultimately, the goal of this initiative is to empower the domain experts in the ASNR to develop the criteria on the basis of experience, evidence, and clinical consultation.

Twenty-five neuroradiologists and staff from the ASNR, ACR, and RSNA participate in the workgroup activities. There is neuroradiology subspecialty representation from the American Society of Spine Radiology, American Society of Functional Neuroradiology, American Society of Pediatric Neuroradiology, and American Society of Head and Neck Radiology. A group e-mail account and a collaborative workspace were set up to support

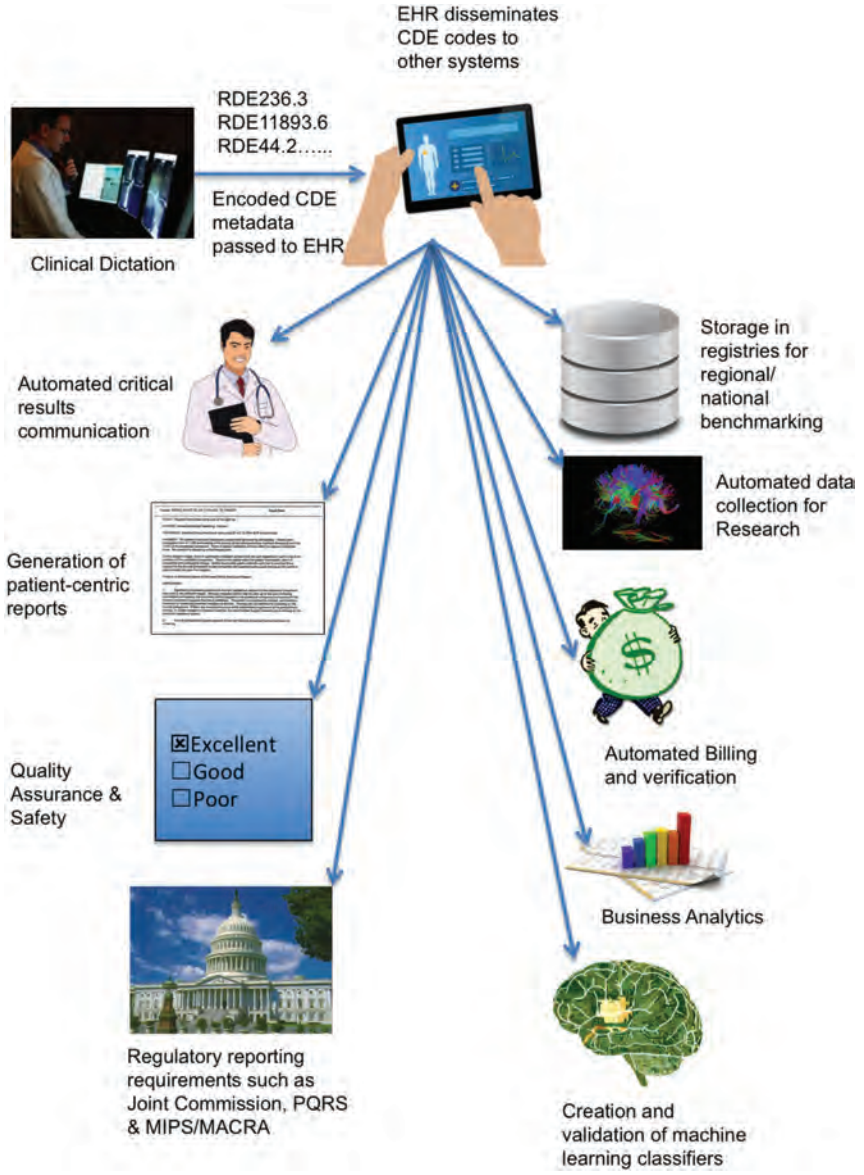


FIGURE. Concepts/responses captured through report CDEs are used in downstream IT systems. Concepts, features, and measurements from CDEs are encoded with unique identifiers (eg, RDE236.3) by the reporting system, which are passed across interfaces and received by various systems programmed to act on specific values. The unique identifiers can trigger other events or be recoded/translated to provide discrete data that drive additional value-based health care processes. PQRS indicates Physician Quality Reporting System; EHR, indicates Electronic Health Record.

asynchronous communication and for all members to have access to all work products and artifacts. The group meets monthly by teleconference with a preplanned agenda, action items, and minutes. Ideas for new CDE nominations are brought to the entire group. A single subspecialty volunteer then takes ownership of the first draft of the CDE set, which is authored directly on a spreadsheet visible to all workgroup members. The workgroup is free to revise or comment on the draft. Corrections or modifications are made on the basis of exchange through group e-mails or via discussion at the monthly teleconference. The final version of the CDE is then handed off to the ACR-RSNA CDE subcommittee to catalog and number in the RadElement.org CDE repository. A Neuro-CDE master list is used to track CDE progress from proposal to final draft. Twelve of the initial CDE sets or modules were converted into PowerScribe 360 macros (https://www.nuance.com/healthcare/medical-imaging/powerscribe-360-reporting.html) and posted on the ASNR Web site for public view/download (https://www.asnr.org/resources/cde/) and were featured in a public demonstration at the ASNR 2018 Annual Meeting (Vancouver, British Columbia, Canada).

What Are the Potential Benefits of CDEs?

A number of other potential tangible benefits and incentives for radiologists to embrace CDE models and reporting exist, and there is growing evidence that inclusion of CDEs in clinical reporting can be performed efficiently, will augment communication, and is preferred by clinicians.[11,12] While current vendor offerings of reporting products are limited in their ability to fully support CDEs, there is a movement underway to address these limitations for the next generation of reporting tools that will include CDEs and radiology decision support content (Computed Assisted Radiology Decision Support [CAR/DS]).[13] Artificial intelligence (AI) and natural language processing cannot ultimately solve the problem of converting heterogeneous prose reports into homogeneous concepts. A combination of solutions that includes new reporting tools that aid radiologists in image interpretation and dictation will ultimately provide the ideal balance between quality and efficiency. While vendors can encode picklists and insertion fields into templates, the current commercial offerings lack the ability to incorporate triggers and logic into the reporting workflow that enhance efficiency, mitigate reporting errors, and augment quality. There is the capability today, however, to dynamically modify a report on the basis of the content that has already been created. For example, mention of a mass on brain MR imaging might invoke a CDE set that cues the radiologist with a list of ASNR-recommended brain mass features. The ACR-Assist (https://www.acr.org/Practice-Management-Quality-Informatics/Informatics/Structured-Content) technology is a radiologist decision support framework that uses the spoken or transcribed concepts in a report as a "trigger" to instantaneously provide consistent and useful supplemental recommendations in a report.[4] The software behind radiology decision support has "awareness" of key concepts/findings/observations (eg, CDEs) and can use this knowledge base to automatically suggest other supplemental features that should be included or to provide

recommendations based on the individual features of a finding. A TI-RADS score could automatically be calculated and inserted into a report on the basis of feature descriptions of a thyroid nodule. The automatic insertion of a macro containing the essential imaging features of laryngeal cancer could follow after describing an incidental laryngeal mass on a neck CT angiogram. Information collected from that macro could generate staging information for the electronic medical record that would be valuable to the oncologist/otolaryngologist. Paramount to the development and deployment of these new software tools is expert review of the inherent concepts and potential enhancements (eg, calculations, assessment scores).

Inclusion of CDEs in reports creates a multitude of opportunities for the concepts in the report to improve other downstream processes. The unique identifier associated with each CDE concept/response can be encoded and transmitted with the text report by a reporting system and can be used to trigger downstream events in other disparate IT systems that have been programmed to comprehend and respond to specific concepts and values. New events could include automatic notification of care team members for critical results communication while the report is still in process. Automated generation of quality-assurance data for a number of clinical use cases such as acute stroke turnaround times and notification could be more accurately collected. Payment denials could be mitigated at the time of report generation by checking for appropriate terminology and concepts in reports that are critical for approval. In the electronic medical record, encoded CDEs could be used to supplement the problem list, progress notes, recommendations, and discharge summaries of the patient. Patient-centric versions of radiology reports could be generated for consumption on patient portals. The concepts from CDEs could be used to collect vast quantities of data for quality assurance and benchmarking in registries. Local, regional, and national registries containing imaging features for specific clinical use cases could be created and could be used for large-scale imaging-based comparative effectiveness research for population health and high-profile health care initiatives. These all have an additive effect of augmenting the value of every radiology report (Figure).

Medical imaging AI research and development could also be accelerated by the inclusion of CDEs in clinical reporting. While close to one-half billion unique imaging studies are generated annually in the United States, only a small portion of these examinations are "AI ready" for training and validation of AI algorithms. The lack of relevant annotations is often cited as the principal reason for shortages of suitable AI training datasets. Investigators have attempted to mobilize natural language–processing applications to retrospectively extract the needed concepts from prose reports with varied success. Additional expert resources are usually required to re-review the original imaging data to create the annotations for a specific disease entity (eg, stroke, glioma, fracture). The annotations and anatomic locations of the features on the images are used to create AI classifiers of disease. Imaging concepts encoded in CDEs make it simpler to create the annotations and subsequently the AI classifiers. The inclusion of CDEs in reports makes it easier to prospectively generate needed annotations. Moreover, universal adoption of CDEs for stroke, cerebral neoplasia, multiple sclerosis, and so forth makes it easier to aggregate data from multiple sites for AI development.

CONCLUSIONS

There are clearly a large number of benefits to be derived from adopting the general practice of using a singular set of concepts, observations, and features in radiology reporting. Codifying the content with neuroradiology domain experts is critical to the success of the process. The joint collaboration among ASNR, ACR, and RSNA is to develop a continual process by which CDE content is proposed, authored, reviewed, approved, and validated for all of neuroradiology. The effort can provide a single clearinghouse of neuroradiology CDEs that can be directly used by the commercial and research sectors to improve product offerings. There is a "symbiosis" between the product development and content creation for CDEs, with each relying on the deliverables of the other. The hope is that the processes being set forth by the ASNR in collaboration with the RSNA and ACR will serve as a pilot for content creation of the other radiology subspecialties. We encourage all practitioners in the "House of Neuroradiology" to contribute and provide guidance for the construction of this collection.

REFERENCES

1. Langlotz CP. *The Radiology Report: A Guide to Thoughtful Communication for Radiologists and Other Medical Professionals*. CreateSpace Independent Publishing Platform, https://www.createspace.com/; 2015
2. Chen JY, Sippel-Schmidt TM, Carr CD, et al. **Enabling the next-generation radiology report: description of two new system standards.** *Radiographics* 2017;37:2106–12 CrossRef Medline
3. Kahn CE Jr, Langlotz CP, Burnside ES, et al. **Toward best practices in radiology reporting.** *Radiology* 2009;252:852–56 CrossRef Medline
4. Rubin DL, Kahn CE. **Common data elements in radiology.** *Radiology* 2017;283:837–44 CrossRef Medline
5. The TCGA Glioma Phenotype Research Group. https://wiki.cancerimagingarchive.net/display/Public/TCGA+Glioma+Phenotype+Research+Group. Accessed May 15, 2018
6. Mamlouk MD, Chang PC, Saket RR. **Contextual radiology reporting: a new approach to neuroradiology structured templates.** *AJNR Am J Neuroradiol* 2018 Jun 14. [Epub ahead of print]. CrossRef Medline
7. US Department of Health and Human Services, Centers for Medicare and Medicaid Services. 42 CFR Parts 414 and 495. Medicare program; Merit-Based Incentive Payment System (MIPS) and Alternative Payment Model (APM) incentive under the Physician Fee Schedule, and criteria for physician focused payment models. Available at: https://www.gpo.gov/fdsys/pkg/FR-2016-05-09/pdf/2016-10032.pdf. Accessed August 20, 2016
8. Schwartz LH, Panicek DM, Berk AR, et al. **Improving communication of diagnostic radiology findings through structured reporting.** *Radiology* 2011;260:174–81 CrossRef Medline
9. H.R.1 - American Recovery and Reinvestment Act of 2009 https://www.congress.gov/bill/111th-congress/house-bill/1/text. Last accessed August 8, 2018.
10. Wibmer A, Vargas HA, Sosa R, et al. **Value of a standardized lexicon for reporting levels of diagnostic certainty in prostate MRI.** *AJR Am J Roentgenol* 2014;203:W651–57 CrossRef Medline
11. Weinberg BD, Gore A, Shu HG, et al. **Management-based structured reporting of posttreatment glioma response with the brain tumor reporting and data system.** *J Am Coll Radiol* 2018;15:767–71 CrossRef Medline
12. Bink A, Benner J, Reinhardt J, et al. **Structured reporting in neuroradiology: intracranial tumors.** *Front Neurol* 2018;9:32 CrossRef Medline
13. Alkasab TK, Bizzo BC, Berland LL, et al. **Creation of an open framework for point-of-care computer-assisted reporting and decision support tools for radiologists.** *J Am Coll Radiol* 2017;14:1184–89

Bias in Neuroradiology Peer Review: Impact of a "Ding" on "Dinging" Others

P. Charkhchi, B. Wang, B. Caffo, and D.M. Yousem

ABSTRACT

BACKGROUND AND PURPOSE: The validity of radiology peer review requires an unbiased assessment of studies in an environment that values the process. We assessed radiologists' behavior reviewing colleagues' reports. We hypothesized that when a radiologist receives a discrepant peer review, he is more likely to submit a discrepant review about another radiologist.

MATERIALS AND METHODS: We analyzed the anonymous peer review submissions of 13 neuroradiologists in semimonthly blocks of time from 2016 to 2018. We defined a discrepant review as any one of the following: 1) detection miss, clinically significant; 2) detection miss, clinically not significant; 3) interpretation miss, clinically significant; or 4) interpretation miss, clinically not significant. We used random-effects Poisson regression analysis to determine whether a neuroradiologist was more likely to submit a discrepant report during the semimonthly block in which he or she received one versus the semimonthly block thereafter.

RESULTS: Four hundred sixty-eight discrepant peer review reports were submitted; 161 were submitted in the same semimonthly block of receipt of a discrepant report and 325 were not. Receiving a discrepant report had a positive effect on submitting discrepant reports: an expected relative increase of 14% (95% CI, 8%–21%). Notably, receiving a clinically not significant discrepant report (coefficient = 0.13; 95% CI, 0.05–0.22) significantly and positively correlated with submitting a discrepant report within the same time block, but this was not true of clinically significant reports.

CONCLUSIONS: The receipt of a clinically not significant discrepant report leads to a greater likelihood of submitting a discrepant report. The motivation for such an increase should be explored for potential bias.

Peer review is one form of evaluation of a radiologist's performance, mostly targeting the diagnostic accuracy of interpretation.[1] The 2007 medical staff standards of The Joint Commission (https://www.jointcommission.org/) have strengthened the peer review process by explicitly requiring focused and ongoing professional practice evaluations. These standards evaluate a practitioner's knowledge, skill, and behavior. Focused professional practice evaluations involve an intense assessment of a practitioner's credentials and current competence at the initial appointment in a practice. Ongoing professional practice evaluations are the routine monitoring of current physician competency, which includes but is not limited to assessment of a practitioner's ongoing interpersonal and communication skills, professional behavior, practice competency, and behavior as a team member.[2] To address these standards, most radiology practices use some form of peer review to assess radiologists' accuracy and performance.[3,4]

The primary goal of radiology peer review is to reduce diagnostic errors, educate radiologists to their blind spots and areas for improvement, and improve patient safety. In addition to evaluating the radiologist's technical performance, peer review can evaluate communication skills, interpersonal relationships, team cooperation, and responsiveness.[5]

The American College of Radiology currently recommends that medical centers participate in physician peer review to obtain and maintain accreditation. Many radiology groups have committed to using a peer review system due to hospital requirements, The Joint Commission standards, or recommendations from specialty societies.[6] There are different types of peer review systems, including RADPEER, implemented by American College of Radiology (https://www.acr.org/Clinical-Resources/RADPEER). Our department is currently using a home-grown peer review system with the advantage of an integrated information technol-

Received July 10, 2018; accepted after revision October 24.

From the Department of Radiology and Radiological Science (P.C., D.M.Y.), Division of Neuroradiology, Johns Hopkins Medical Institutions, Baltimore, Maryland; and Department of Biostatistics (B.W., B.C.), Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland.

Please address correspondence to David M. Yousem, MD, MBA, Johns Hopkins Bloomberg School of Public Health, 600 N Wolfe St, Baltimore, MD 21287; e-mail: dyousem1@jhu.edu; @dyousem1

http://dx.doi.org/10.3174/ajnr.A5908

ogy solution that allows the review of cases within 24 hours of completion, thereby catching errors early, rather than several months later. The process is anonymized so that neither the reviewer nor the original reader knows the author of the reports or the peer reviews. Both RADPEER and our internal system are scoring-based peer review systems.

As Kaewlai and Abujudeh[5] indicated, 2 critical areas for success in peer review are a positive peer review culture and a committed team. Larson et al[7] indicated that scoring-based systems tend to drive radiologists inward, against each other and against practice leaders. Our aim in this communication was to critically assess the radiologists' behavior in the setting of reviewing colleagues' reports. We hypothesized that when a radiologist receives a discrepant review, he or she would be more likely to report a discrepant review (colloquially referred to as a "ding") on another person's report within 2–4 weeks.

MATERIALS AND METHODS

Because this project dealt with quality-improvement processes, it was deemed by the Johns Hopkins institutional review board to be exempt from review. The data collected were independent of protected health information; therefore, the study was Health Insurance Portability and Accountability Act–compliant.

Data Source

This is a retrospective study. We used peer review data from the division of neuroradiology because of the early implementation of the internal system by this team. The peer review system is completely anonymous: The radiologists are aware of neither who has reviewed their reports nor whose report they are reviewing. The program randomly selects colleagues' reports from the previous 24 hours and assigns them to the peer reviewer, providing the report without the author of the report being identified. It opens the images on the case. The radiologist reviewing the case must choose whether he or she concurs with the interpretation, identifies a detection miss, or identifies an interpretation miss. If a miss is identified, it is scored as clinically significant or clinically not significant. If such a miss is identified, whether significant or not, the peer reviewer fills in a text box identifying the miss. The radiologist submits the case; and if a miss has been identified, the original reader receives an e-mail immediately thereafter notifying him or her that there is a discrepancy and to review the case.

We collected data from January 1, 2016, to January 30, 2018, in increments from the first of the month to the 15th day of the month and from the 16th day of the month to the last day of the month (ie, twice a month for 25 months) for all 13 neuroradiologists who were practicing in our facility for the full duration of the study. The system provides data on the number of cases that are read by the radiologist, the number of peer reviews that the radiologist completed, the number of discrepant reviews that the radiologist submitted, the number of the radiologist's cases that were reviewed by neuroradiology colleagues, the number of discrepant cases in which a lesion was not detected and whether it was clinically significant, and the number of cases in which a lesion was appropriately detected but its etiology was misinterpreted and whether that misinterpretation was clinically significant. We defined a discrepant review in our study as any one of the following: 1) detection miss clinically significant, 2) detection miss clinically not significant, 3) interpretation miss clinically significant, and 4) interpretation miss clinically not significant.

The members of the neuroradiology division are encouraged to perform peer review each day that a neuroradiologist has clinical duties, and all members must review cases at a rate equaling at least 3% of the total number of cases they read each month (ie, if they read 600 cases, they have to peer review at least 18 cases from colleagues). Fulfilling this participation rate is part of their end of year bonus "quality and safety" calculation.

Study Variables

We defined the independent variable as the receipt of a discrepant report. We defined 1 dependent variable as submitting a discrepant review within the semimonthly time block of receiving a discrepancy and a second dependent variable as submitting a discrepant review in the semimonthly block after receiving the discrepant review. As an example, if someone received a discrepant report on day 7 of the month, we surveyed for discrepant reports within the block that included the first 15 days of the month (first dependent variable) and the last half of the month (second dependent variable) for a discrepant submission by that person. If the discrepancy was received, for example, on the 18th day of the month, then the second half of the month becomes the first dependent variable and the first 15 days of the next month were the second dependent variable. We did not include more than 4 weeks because we assumed that the likelihood of a reflexive response diminished after a 2- to 4-week interval.

By virtue of collecting the data twice a month for 2 years of practice from January 1, 2016, to December 31, 2017, and including the 4 weeks of follow-up extending to January 31, 2018, we had 50 data points. We also assessed the association between the type of discrepant report (clinically significant versus clinically not significant) and submitting the discrepant report on others.

Data Analysis

We used random-effects Poisson regression models to assess the effect of receiving a discrepant review on submitting a discrepant report within the received block and the next block (with doctor as the random effect). Also, we included a multivariate regression model in our analysis using clinically significant and non-clinically significant reports as covariates to assess the association of receiving different types of discrepant reviews (clinically significant versus clinically not significant) and submitting one. All analyses were performed with R statistical and computing software (Version 3.4.3; www.r-project.org).

We ran 2 sensitivity analyses by including and excluding outliers. In the first analysis, we excluded 1 radiologist with the largest number of submitted discrepant reports. This radiologist reported, on average, 3.26 discrepant reports, while the mean of all radiologists was 0.75 ± 1.6. We thought this radiologist may influence the results because this radiologist submitted the highest number of discrepant reviews in the study time period. In the second sensitivity analysis, we checked for any extreme observations and excluded greater than 5 discrepancy reports received or submitted in a block and repeated the analysis.

RESULTS

The overall distribution of submitted reports for each neuroradiologist is presented in Fig 1.

In the 2-year period, 486 discrepant peer review reports were submitted, of which 161 were submitted by individuals in the same 2-week block in which they received notice of a discrepant report; 325 were not. There was a positive effect (coefficient = 0.13; 95% confidence interval, 0.08–0.19) between submitting a discrepant report within the block of receiving one. The relative rate was 14% (95% CI, 8%–21%). In other words, according to the model, for every 5 discrepancy reports received, the number of discrepancies submitted will be doubled (times 1.93). If one does not receive any discrepant reports, then he or she will submit 0.47 discrepancy reports (on average) in a 2-week block.

There was no statistically significant effect between receiving a discrepant report in one block and then submitting one in the following 2-week time block (coefficient, −0.09; 95% CI, −0.19–0.02) (Table and Fig 2).

If one ran a multivariate regression analysis to assess the effect of different discrepant reports on submitting a ding on others, there was a significant association between receiving a not clinically significant (coefficient = 0.13; 95% CI, 0.05–0.22) report and submitting a discrepant report, while there was no statistically significant association between receiving a clinically significant (coefficient = 0.26; 95% CI, −0.04–0.55) discrepant report and submitting a discrepancy in the same time block. There was no significant association between receiving clinically significant or not clinically significant discrepant reports and submitting a ding on others in the next following 2-week time block (Table).

After excluding 1 outlier radiologist who submitted the most discrepancies, there were no changes in our results: For the same block analysis, the coefficient changed from 0.13 to 0.12 (95% CI, 0.03–0.21) and remained statistically significant; for the next time block, the impact remained nonsignificant (coefficient = −0.05; 95% CI, −0.18–0.07).

The coefficient for each radiologist is reported in Fig 3. If one compared the 2 time blocks, there was no significant association between receiving and submitting a discrepant review for any of the radiologists in the next time block (Fig 3).

Removing outliers (more than 5 discrepant or submitted reports) kept the

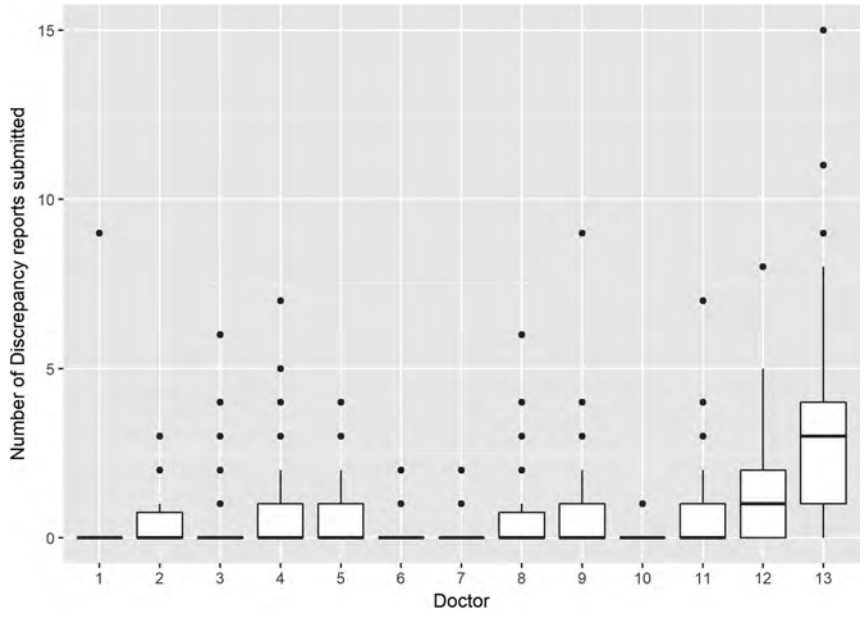


FIG 1. Mean number of discrepancy reports submitted for each neuroradiologist.

The association between receiving and submitting discrepant reports in the same and next time block in a clinical radiology peer review system

Received a Discrepant Report	Submitting a Discrepant Report within Same Time Block (Coefficient) (95% CI)	Submitting a Discrepant Report in the Next Time Block (Coefficient) (95% CI)
Any kind of discrepant report	0.13 (0.08–0.19)	−0.09 (−0.19–0.02)
Clinically significant discrepant report	0.26 (−0.04–0.55)	−0.30 (−0.81–0.20)
Clinically not significant discrepant report	0.13 (0.05–0.22)	−0.05 (−0.18–0.07)

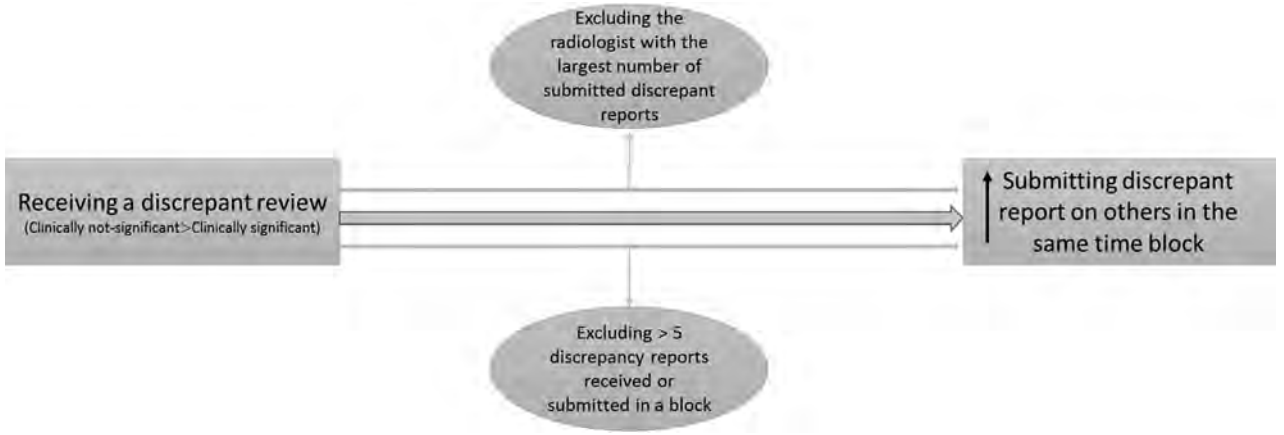


FIG 2. The association between receiving and submitting a discrepant report and 2 sensitivity analyses.

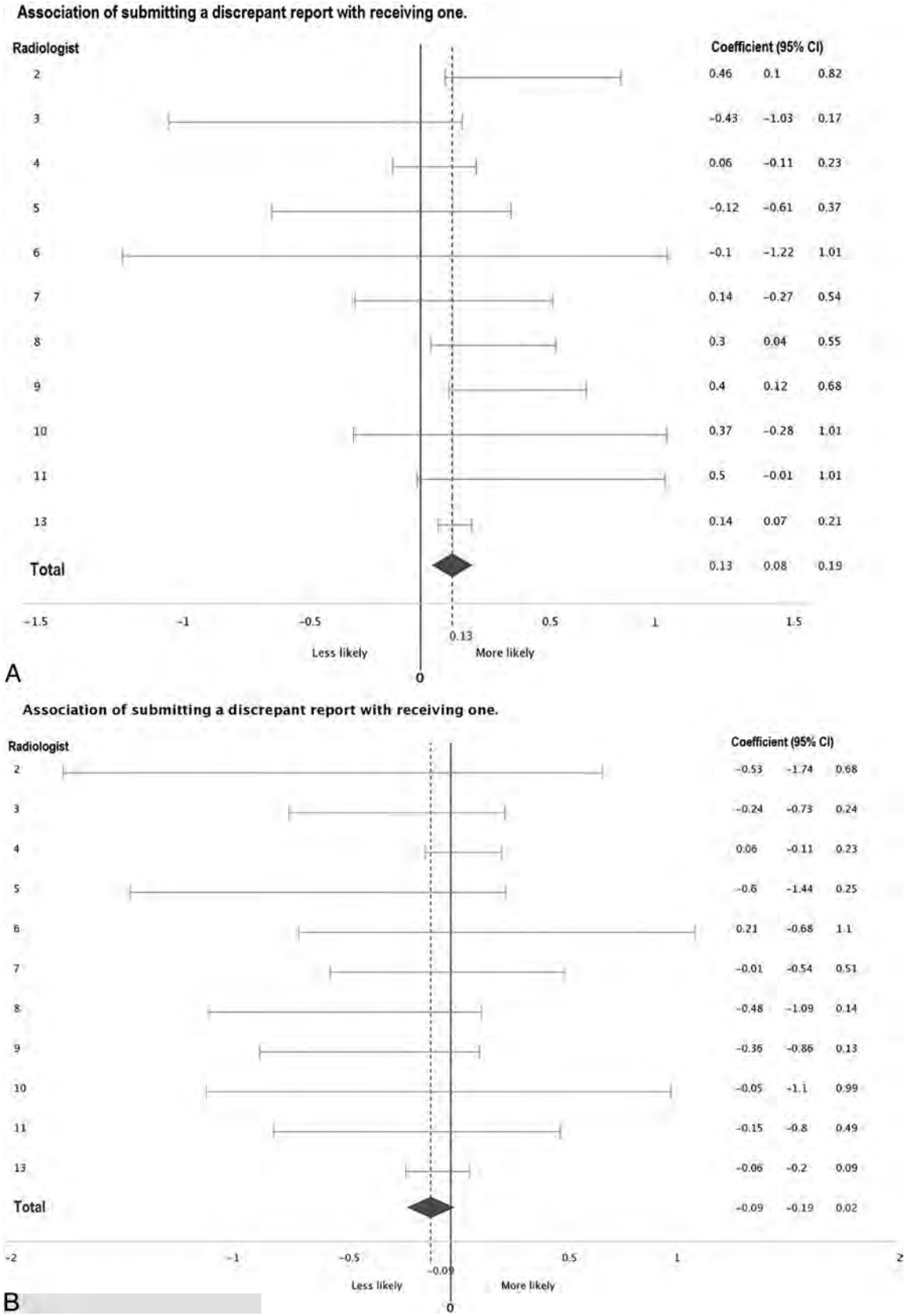


FIG 3. Association of receiving a discrepant review with submitting one for each radiologist. Plot A (top) indicates the current time block; plot B (bottom), the next time block.

current time block significant (coefficient = 0.32; 95% CI, 0.12–0.34) and the next time block nonsignificant (coefficient = 0.06; 95% CI, −0.08–0.19).

DISCUSSION

We found that when a radiologist in our study received a discrepant report, he or she was more likely to submit a discrepant peer review report within the 2-week time block of receiving it. The observed effect was not seen in the following 2-week block of time, suggesting an immediate reaction to the ding rather than a delayed or sustained effect. Receiving a clinically not significant report and submitting a discrepant report on others are significantly positively correlated compared with receiving a "clinically significant" report.

To the best of our knowledge, this is the first article studying physicians' reactions to a discrepant report in the clinical setting. Data are well-published on causes of discrepancies in radiology and also strategies to prevent them.[8-10] However, none of these articles qualitatively or quantitatively studied the radiologists' behavior on receiving a discrepant report.[7,11]

Generally, there is a negative attitude toward the peer review system among radiologists. In an American College of Radiology survey assessing the RADPEER program, most radiologists opined that the peer review system is only performed to meet accreditation and hospital credentialing requirements.[12] Nearly half believed that their practice patterns had not changed as a result of peer review. One-third of respondents admitted that there was underreporting of disagreements in the peer review process at their practice.[12] This underreporting highlights the current peer review systems deficits. Peer review may elicit anxiety, shame, humiliation, and fear, leading to a reluctance to report disagreements.[7] These factors may lead to the behavior demonstrated in our study. If the peer review system is converted into a retributive instrument among colleagues, it becomes worse than meaningless; it becomes destructive.

On the other hand, the positive effect of receiving a discrepant report on submitting discrepant reports may illustrate a positive bias rather than a negative reaction. While previous studies have shown that radiologists tend to underreport discrepancies on peer review,[12,13] in contrast, our data suggest that receiving a discrepant report may motivate the radiologist to review their colleagues' reports more diligently and potentially identify errors that might otherwise be overlooked. However according to our findings, participants tend to submit more discrepancy reports on their colleagues when they receive a not clinically significant report compared with a clinically significant one. We posit that this result may be in favor of a motivation for a retributive reaction rather than motivating the reviewer to be more conscientious. When a radiologist receives a discrepant report that is clinically significant, she or he may react with gratitude and not negatively react to it, but if she or he receives a clinically not significant ("nuisance") discrepant report, the radiologist may be more likely to respond by submitting a reciprocal discrepant report on a colleague.

There are a few limitations associated with this study. First, the peer review system we use is unlike most peer review programs that use historical studies for review. In other words, most peer review systems require the radiologist to review a comparison study from months to years earlier. In that gap, the diagnosis may become clear and, for example, growth of a missed cancer can be readily detected. By limiting our peer review system to reviews within the previous 24 hours, we identify discrepancies earlier, but a final diagnosis may not be clear at that point. Second, if the reviewing radiologist wanted to enter the Radiological Information System (RIS) or Electronic Medical Record (EMR), he or she could break the anonymity of the self-contained peer review program and identify who read each study. Third, we used semimonthly time intervals because our peer review system data are collected this way. We cannot determine whether the radiologist immediately submitted a discrepant report the same hour or day that he or she received a discrepant report because we do not have the data on the exact time of receiving and submitting reports. We do not monitor the peer review system at such a granular level. On a similar note, the reporting function of the program is able to document that a dinged physician submitted a discrepant report but not the type (detection versus interpretation/significant or not) of report. Finally, if a discrepancy is challenged by the receiver, the division chief then adjudicates the 2 reviews, which could change the initial discrepant designation by the dinged physician.

How can we address this potential bias in the peer review system? We could write code to the program that after receiving notice of a discrepant report and reviewing it, that individual is "frozen" from submitting any peer review reports for 7 days. Thus, the more immediate "gut" reaction could be assuaged. Education and re-education continuously on the purpose of peer review may also be helpful. Providing data showing the overall results and how well individuals perform may decrease the psychological impact of a solitary discrepant review. In addition, department leadership support to keep peer review results completely anonymous, blinded to leadership, and accessible only by individual physicians can improve rate of participation in the peer review system.

CONCLUSIONS

When a radiologist in our study received a discrepant report, he or she was more likely to submit a discrepant report within the semimonthly block of time of receiving it. The observed effect was not seen in the following block of time, suggesting an immediate reaction to the ding rather than a delayed or sustained effect. The impact was maximal after receiving a clinically Radiological Information System (RIS) discrepant peer review.

Disclosures: Brian Caffo—*UNRELATED*: *Consultancy*: for d8alab. *Comments*: I do personal consulting from time to time; *Grants/Grants Pending*: National Institutes of Health*; *Royalties*: Leanpub Publishing. *Comments*: book royalties; *Payment for Development of Educational Presentations*: Becton Dickison. *Comments*: on-line courses. David M. Yousem—*UNRELATED*: *Expert Testimony*: medicolegal work; *Payment for Lectures Including Service on Speakers Bureaus*: American College of Radiology Education Center speaker; *Royalties*: Elsevier for 5 books; *Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed*: Radiological Society of North America 2018 as an awardee. *Money paid to the institution.

REFERENCES

1. Alpert HR, Hillman BJ. **Quality and variability in diagnostic radiology.** *J Am Coll Radiol* 2004;1:127–32 CrossRef Medline
2. Freedman S. **How 2007 joint commission standards expand hospital peer review.** 2007. https://www.psqh.com/analysis/peer-review-how-2007-joint-commission-standards-expand-hospital-peer-review/. Accessed November 17, 2018

3. Donnelly LF, Strife JL. **Performance-based assessment of radiology faculty: a practical plan to promote improvement and meet JCAHO standards.** *AJR Am J Roentgenol* 2005;184:1398–401 CrossRef Medline
4. Donnelly LF. **Performance-based assessment of radiology practitioners: promoting improvement in accordance with the 2007 Joint Commission standards.** *J Am Coll Radiol* 2007;4:699–703 CrossRef Medline
5. Kaewlai R, Abujudeh H. **Peer review in clinical radiology practice.** *AJR Am J Roentgenol* 2012;199:W158–62 CrossRef Medline
6. Kruskal J, Eisenberg R. **Focused professional performance evaluation of a radiologist: a Centers for Medicare and Medicaid Services and Joint Commission requirement.** *Curr Probl Diagn Radiol* 2016; 45:87–93 CrossRef Medline
7. 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
8. Brady AP. **Error and discrepancy in radiology: inevitable or avoidable?** *Insights Imaging* 2017;8:171–82 CrossRef Medline
9. Bruno MA, Walker EA, Abujudeh HH. **Understanding and confronting our mistakes: the epidemiology of error in radiology and strategies for error reduction.** *Radiographics* 2015;35:1668–76 CrossRef Medline
10. Brady A, Laoide RÓ, McCarthy P, et al. **Discrepancy and error in radiology: concepts, causes and consequences.** *Ulster Med J* 2012;81: 3–9 Medline
11. FitzGerald R. **Radiological error: analysis, standard setting, targeted instruction and teamworking.** *Eur Radiol* 2005;15:1760–67 CrossRef Medline
12. Abujudeh H, Pyatt RS, Bruno MA, et al. **RADPEER peer review: relevance, use, concerns, challenges, and direction forward.** *J Am Coll Radiol* 2014;11:899–904 CrossRef Medline
13. Moriarity AK, Hawkins CM, Geis JR, et al. **Meaningful peer review in radiology: a review of current practices and potential future directions.** *J Am Coll Radiol* 2016;13:1519–24 CrossRef Medline

Deep Learning–Based Detection of Intracranial Aneurysms in 3D TOF-MRA

T. Sichtermann, A. Faron, R. Sijben, N. Teichert, J. Freiherr, and M. Wiesmann

ABSTRACT

BACKGROUND AND PURPOSE: The rupture of an intracranial aneurysm is a serious incident, causing subarachnoid hemorrhage associated with high fatality and morbidity rates. Because the demand for radiologic examinations is steadily growing, physician fatigue due to an increased workload is a real concern and may lead to mistaken diagnoses of potentially relevant findings. Our aim was to develop a sufficient system for automated detection of intracranial aneurysms.

MATERIALS AND METHODS: In a retrospective study, we established a system for the detection of intracranial aneurysms from 3D TOF-MRA data. The system is based on an open-source neural network, originally developed for segmentation of anatomic structures in medical images. Eighty-five datasets of patients with a total of 115 intracranial aneurysms were used to train the system and evaluate its performance. Manual annotation of aneurysms based on radiologic reports and critical revision of image data served as the reference standard. Sensitivity, false-positives per case, and positive predictive value were determined for different pipelines with modified pre- and postprocessing.

RESULTS: The highest overall sensitivity of our system for the detection of intracranial aneurysms was 90% with a sensitivity of 96% for aneurysms with a diameter of 3–7 mm and 100% for aneurysms of >7 mm. The best location-dependent performance was in the posterior circulation. Pre- and postprocessing sufficiently reduced the number of false-positives.

CONCLUSIONS: Our system, based on a deep learning convolutional network, can detect intracranial aneurysms with a high sensitivity from 3D TOF-MRA data.

ABBREVIATIONS: CNN = convolutional neural network; DSC = Dice similarity coefficient; FPs/case = false-positives per case

Unruptured intracranial aneurysms are common among the general population. It is estimated that approximately 3% of healthy adults have an intracranial aneurysm.[1] These aneurysms often remain undiagnosed unless they become symptomatic (eg, by compression of adjacent neural structures or rupture into the subarachnoid space).[2] Rupture of an intracranial aneurysm is a serious incident with high fatality and morbidity rates.[3] Identification of factors contributing to the risk of intracranial aneurysm development, growth, and rupture is an active field of investigation. Apart from several disorders like polycystic kidney disease or Marfan syndrome, elements such as genetic factors, family history, female sex, and age are linked to an increased risk of aneurysm development. Intracranial aneurysm site, size, and shape are further strongly associated with the risk of rupture.[4-6] Detection of an intracranial aneurysm before it becomes symptomatic allows endovascular or surgical treatment of the aneurysm before it ruptures and may thus prevent death or morbidity.

DSA is still considered the criterion standard in evaluating intracranial vessels and detection of intracranial aneurysms[7]; however, it is inconvenient for primary diagnoses because it is invasive and time-consuming. CTA and MRA are noninvasive methods widely used in clinical routine. Unlike DSA and CTA, which are based on x-ray imaging, MRA does not cause radiation exposure. It is therefore the preferred technique for screening asymptomatic patients for intracranial pathology. The number of radiology examinations performed for diagnoses is steadily in-

Received June 5, 2018; accepted after revision October 29.

From the Department of Diagnostic and Interventional Neuroradiology (T.S., A.F., R.S., N.T., J.F., M.W.), University Hospital RWTH Aachen, Aachen, Germany; Department of Radiology (A.F.), University Hospital Bonn, Bonn, Germany; and Department of Diagnostic and Interventional Neuroradiology (N.T.), University Hospital Düsseldorf, Düsseldorf, Germany.

Thorsten Sichtermann and Anton Faron contributed equally to this work.

This work was supported by Nvidia Corporation with the donation of a Titan XP GPU.

Please address correspondence to Thorsten Sichtermann, MD, Klinik für Diagnostische und Interventionelle Neuroradiologie, Universitätsklinikum Aachen, Pauwelsstr 30, 52074 Aachen, Germany; e-mail: tsichtermann@ukaachen.de

http://dx.doi.org/10.3174/ajnr.A5911

creasing.[8,9] Given the growing workload of radiology departments, physician fatigue with the inherent risk of missed diagnosis of potentially significant findings is a relevant concern. Hence, a reliable method for automated detection of intracranial aneurysms from routine diagnostic imaging would be of great utility in clinical routine.

Rapid advances in the field of computing and a growing amount of data prompted the rise of convolutional neural networks (CNNs), a specific type of deep learning network architecture, for segmentation, classification, and detection tasks in medical imaging.[10-12] The training process of a CNN is straightforward to implement because the features for discrimination of the desired output classes are not designed but learned in an automated fashion from the input data.[13] Several approaches for automated detection of intracranial aneurysms from noninvasive imaging have been proposed in the literature.[14-17] However, a deep learning–based method for sufficient detection of intracranial aneurysms from 3D TOF data has not yet been reported, to our knowledge. The aim of this study was to investigate the potential of a deep learning algorithm for automated detection of intracranial aneurysms from 3D TOF-MRA clinical data.

MATERIALS AND METHODS

Dataset

This retrospective study was approved by the Independent Ethics Committee at the RWTH Aachen Faculty of Medicine. The requirement for informed consent was waived. From an internal data base belonging to our department, we incorporated data from all patients with a 3D TOF-MRA examination of at least 1 previously untreated intracranial aneurysm. Images were obtained for clinical purposes between 2015 and 2017. After we removed protected patient information and substituted subject identifiers, examinations were retrieved from the local PACS. The dataset consisted of 85 examinations. Of those, 72 image sets originated from our department. Sixty of these examinations were performed on a 3T scanner (Magnetom Prisma; Siemens; Erlangen, Germany). Twelve examinations were performed on a 1.5 scanner (Magnetom Aera; Siemens).

The following parameters were used for the 3D TOF-MRA: Magnetom Prisma (3T)—TR, 21 ms; TE, 3.42 ms; flip angle, 18°; FOV, 200 mm; section thickness, 0.5 mm; matrix, 348 × 384; acquisition time, 5 minutes 33 seconds; 20-channel head/neck coil; Magnetom Area (1.5T)—TR, 28 ms; TE, 7 ms; flip angle, 25°; FOV, 200 mm; section thickness, 0.5 mm; matrix, 256 × 320; acquisition time, 5 minutes 52 seconds; 20-channel head/neck coil.

Thirteen examinations included in this dataset originated from external departments and were performed on different scanners.

We included all TOF acquisitions with at least 1 previously untreated aneurysm, irrespective of etiology, symptomatology, and configuration (saccular, fusiform, and dissecting). The aneurysms were located in the internal carotid arteries, the anterior cerebral arteries (including the anterior communicating artery), the middle cerebral arteries, or the posterior circulation (including the vertebral, basilar, posterior, cerebral, and posterior communicating arteries). One patient had polycystic kidney disease, while the remainder had incidental findings. Exclusion criteria were previous treatment (coil embolization or surgical clipping) or pronounced motion artifacts, preventing accurate segmentation.

The DeepMedic (Version .6.1; https://biomedia.doc.ic.ac.uk/software/deepmedic/) CNN was used[18] with an application of required preprocessing on the dataset[19]: voxel size resampling ($0.5 \times 0.5 \times 0.5$ mm^3) and intensity normalization to a zero-mean, unit-variance space. To evaluate the impact of preprocessing on the performance of the CNN, we modified our dataset using different BET2 skull-stripping (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET)[20] and performing N4 bias correction.[21]

The ground truth segmentation was performed by a neuroradiology resident experienced in cranial diagnostic imaging. On the basis of radiologic reports, anonymized TOF data were critically reviewed, and aneurysms were manually annotated in a voxelwise manner using the manual segmentation tool of ITK-SNAP (www.itksnap.org).[22] Intrarater reliability was studied using the Pearson correlation coefficient.

After evaluation of the dataset, we trained DeepMedic and performed inference to segment aneurysms. Remarkably, 2 aneurysms that had been previously overlooked were detected by the CNN in this early stage. Consequently, the dataset was validated by another radiologist who was blinded to the radiology reports. Complete ground truth was evaluated once again and adjusted accordingly.

The dataset needed division into training, test, and validation sets, to run the CNN and assess its performance. The training set was used for learning, which describes the process of fitting the parameters of the network to learn features for discriminating the output classes. The validation set was used during training to reduce overfitting to the training data. This is done by comparing the Dice similarity coefficient (DSC) (a measure indicative of segmentation accuracy) of the training samples with the DSC of the unknown validation samples and adjusting the learning rate of the network. The test set is used for evaluation of the trained model.[18] Training the model took about 20 hours; inference per case was about 50 seconds on a Titan XP GPU (Nvidia, Santa Clara, California).

Five-fold cross-validation was performed. For each split, the whole dataset was randomly divided into the 3 subsets as explained earlier: training set (58 cases, 68%), validation set (10 cases, 12%), and test set (17 cases, 20%).

DeepMedic and Evaluation

Segmentation of the aneurysms was executed with the DeepMedic framework, a CNN for voxelwise classification of medical imaging data after training with 3D patches at multiple scales. DeepMedic was developed and evaluated for the segmentation of brain lesions.[23]

The network consists of 2 pathways with 11 layers. Both pathways are identical, but the input of the second pathway is a subsampled version of the first (see the full architecture in Fig 1). Parameters were set as proposed by Kamnitsas et al[18]: An initial learning rate of 10^{-3} was used and gradually reduced. For optimization, a Nesterov Momentum of 0.6 was set. For better regularization, drop-out and L1 =

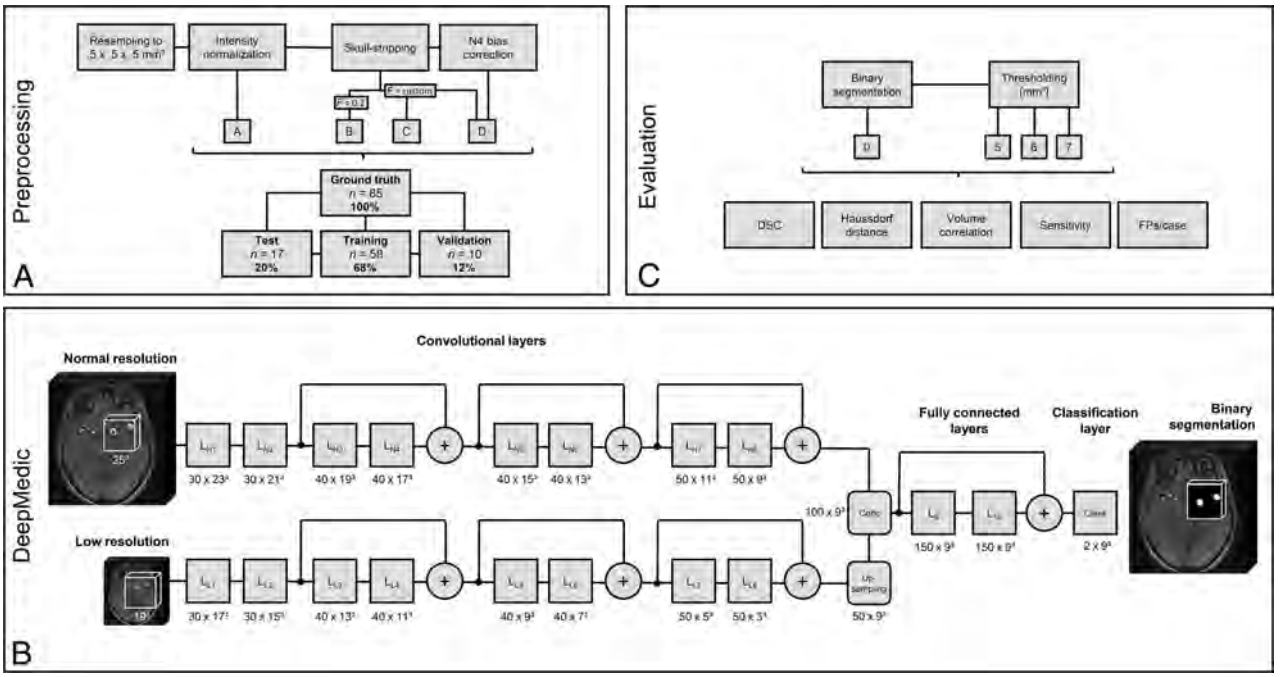


FIG 1. Flowchart of the pipeline. *A*, Preprocessing is performed with 4 different models. The dataset is split into test, training, and validation sets. *B*, Inference is performed with the convolutional neural network DeepMedic with a 2-pathway architecture. The number of feature maps and their size is depicted as number × size. The + depicts the addition of the 2 preceding layers, which adds an additional nonlinearity and reduces the number of weights.[18] The diagram is based on the depiction in the DeepMedic documentation. (Modified from Kamnitsas K, Ledig C, Newcombe VFJ, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis* 2017;36:61–78 under CC-BY4 license).[32] *C*, Thresholding is applied to the resulting segmentation and evaluated with different metrics. L_N indicates layers in the normal resolution pathway, L_L indicates layers in the low resolution pathway.

10^{-6} and L2 = 10^{-4} regularization was performed. To accelerate the convergence, we used Rectified Linear Unit activation functions and batch-normalization as implemented in the DeepMedic framework.[23] We used the proposed DeepMedic hybrid sampling scheme. In this strategy, image segments larger than the neural network's receptive field are given as an input to the network. A training batch is built by extracting segments with 50% probability centered on the foreground or background voxels, facilitating an automatic method for balancing the distribution of training samples regarding the size of the desired class in the segment and therefore preventing class imbalance by adjusting to the true distribution of background and aneurysm voxels.[18] With a probability of 50%, the training images were mirrored on the coronal axis to increase the diversity of the training set.

We used the EvaluateSegmentation Tool (https://github.com/Visceral-Project/EvaluateSegmentation)[24] to analyze the segmentation results by determining Hausdorff distances and the DSC. For methodologic reasons, each segmented voxel or connected component of voxels in the output binary segmentation was considered a positive detection. Each positive detection that corresponded to an aneurysm in ground truth was considered a true-positive finding, while each positive detection that did not correspond to an aneurysm in ground truth was considered a false-positive finding. In preliminary studies, this approach led to a very high rate of false-positive detections. Because we observed that compared with true-positive detections, false-positives tended to be rather small, we further examined whether the integration of a detection threshold as a postprocessing step, removing connected components smaller than a given volume, would improve our results. On the basis of the composition of our dataset, detection thresholds of 5, 6, and 7 mm^3 were studied (Fig 1). To further reduce the number of false-positives, we fine-tuned the network using a modified training strategy in which 90% of the input samples corresponded to background class; and 10%, to aneurysm class, reflecting a more realistic distribution of aneurysms. The learning rate was lowered to 10^{-4} and the pretrained weights of the last 3 layers were changed while the training weights of the other layers were kept constant. To study the reliability of true-positive detections and the capability of the system in predicting aneurysm size, we compared the volume segmented by the algorithm with the manually examined volume of the ground truth.

To assess the impact of preprocessing, we evaluated 4 models (A–D). In model A, only the necessary steps to obtain reasonable results from DeepMedic, resampling to isotropic voxel size and intensity normalization, were performed. Additional skull-stripping is advised in the DeepMedic documentation.[18] We used the well-established BET2 skull-stripping method. Skull-stripping in model B was performed with a fixed fractional intensity threshold of 0.2. In model C, the parameter was adjusted manually in each case to receive an optimal brain outline, without nonbrain structures such as skull or parts of the ocular muscles and nerves. For model D, we used the skull-stripping masks from model C and performed an additional N4 bias correction[25] to evaluate whether low-frequency intensity inhomogeneities in the acquisitions would have an impact on the performance of the algorithm (Fig 1). In this work, each model is depicted as a preprocessing model identifier (A–D), followed by the detection threshold (0, 5, 6, 7).

Full preprocessing per case took about 5 minutes on a Corei7–8700K CPU (Intel, Santa Clara, California). Individual creation of a skull-stripping mask was performed by an experienced user and took about 8 minutes for each sample.

Statistical analysis was performed using SPSS software, Version 25.0 (Released 2017; IBM Armonk, New York). We used the Shapiro-Wilk test to test for normality. Significance values of normality tests are only reported for cases in which the normality assumption was violated. A Kruskal-Wallis test was used for the split-validation of maximum diameters.

Comparisons among the Models

We hypothesized the 4 different levels of preprocessing to each be improvements over the previous version. Therefore, sensitivity values of each preprocessing model were compared only with those of its closest neighbor by testing for differences in the proportions of hits and misses, using McNemar tests. These tests were chosen over χ^2 tests because the values obtained from each model were not independent of one another. Comparing each model with its closest neighbor yielded 3 comparisons (A0 versus B0, B0 versus C0, C0 versus D0); thus, significance levels were corrected for 3 comparisons using a Bonferroni correction.

False-positives per case (FPs/case) were compared for each preprocessing model using a Friedman test. Post hoc tests were run using Wilcoxon signed rank tests for each closest neighbor.

DSCs of each preprocessing model were compared using Friedman tests. Post hoc tests were run using Wilcoxon signed rank tests for each closest neighbor. Missing values, caused by the inability of the evaluation tool to analyze volumes with no segmented voxels, were set to zero.

Hausdorff distances of each preprocessing model were analyzed using a linear mixed model, which included a random subject factor, and "model" as the sole fixed dependent variable. This linear mixed model was chosen over a repeated-measures ANOVA because the linear mixed model can analyze missing values better; unlike DSCs, a Hausdorff distance of zero would not accurately describe the inability of the tool to analyze a volume with no segmented voxels.

Comparisons within the Models

We hypothesized that each of the postprocessing models reduces the number of false-positives sequentially. Thus, sensitivity values for each detection threshold were compared with those of the closest neighbor within each model by testing for a difference in the proportions of hits and misses using McNemar tests. This yielded 3 comparisons per preprocessing model (0 versus 5, five versus 6, and 6 versus 7).

FPs/case were compared for each detection threshold using a Friedman test. Post hoc tests were run using Wilcoxon signed rank tests comparing each closest neighbor.

Size and Location

Increased aneurysm size embodies an increased rupture risk.[4] However, consented classifications of aneurysms based on aneurysm size are missing. To study the impact of aneurysm size on the detection rate, we classified aneurysms on the basis of maximum diameter as follows: In the literature, aneurysms with a maximum diameter of ≤3 mm are generally considered tiny.[26] For simplification, we termed these findings small aneurysms. A distinct increased risk of rupture was identified for aneurysms with a diameter of >7 mm.[6] We therefore defined aneurysms of >3 but ≤7 mm as medium, and those of >7 mm as large. Additionally, aneurysms were categorized on the basis of their location.

Sensitivity values of these categories were compared for both categorizations using Fisher exact tests rather than χ^2 tests because the cases numbered below 5 for certain cells. Spearman rank correlation coefficients were calculated between ground truth and predicted volumes because the normality assumption was violated in all samples.

RESULTS

Dataset

In 85 patients (58 women, 68%; 23–84 years of age; mean, 56 ±13 years), 115 untreated aneurysms with a mean volume of 214.6 ± 480.9 mm^3 (range, 6.4–4518.0 mm^3) and a mean maximum diameter of 7.1 ± 4.4 mm (range, 2.1–37.0 mm) were identified as the ground truth. Intrarater reliability for manual aneurysm segmentation was excellent ($r = 0.998$; 95% CI, 0.988–0.999; $P < .0001$). In the dataset, large-sized aneurysms accounted for 39%; medium-sized aneurysms, for 50%; and small-sized aneurysms, for 11%.

The locational proportion of aneurysms was as follows: Forty-two percent of all aneurysms were located in internal carotid arteries; 17%, in the anterior cerebral arteries, including the anterior communicating artery; 23%, in the middle cerebral arteries; and 19%, in the posterior circulation, including the vertebral, basilar, posterior, cerebral, and posterior communicating arteries.

For cross-validation, the dataset was split into 5 subgroups in a randomized fashion. Normality was violated for the diameter distributions in the splits ($P < .001$). The mean maximum diameter values of the splits did not differ significantly (χ^2 [4] = 6.195, $P = .19$). The mean maximum diameters of the 5 splits were 7.6 ± 3.7 mm, 7.9 ± 6.8 mm, 5.3 ± 2.4 mm, 7.6 ±4.1 mm, and 7.0 ±3.1 mm, respectively.

Sensitivity among Models

Comparing sensitivity values of the nearest neighbors' preprocessing models (A0, B0, C0, and D0) yielded no significant differences ($P = 1$, binomial distribution used for all comparisons). Even the models showing the largest difference (A0 versus D0) did not approach significance ($P = .29$, binomial distribution used, uncorrected for multiple comparisons).

False-Positives per Case among Models

Analyses of false-positive rates between the preprocessing models revealed a significant difference among models (χ^2 [3] = 136.144, $P < .001$). Pair-wise comparisons indicated a significant difference between models A0 and B0 ($z = 7.425$, $P < .001$), but not B0 and C0 or C0 and D0 ($z = 1.878$, $P = .18$ and $z = 0.991$, $P = .97$, respectively).

For each preprocessing model, the impact of detection thresholds on sensitivity, FPs/case, and positive predictive value was studied (Fig 2).

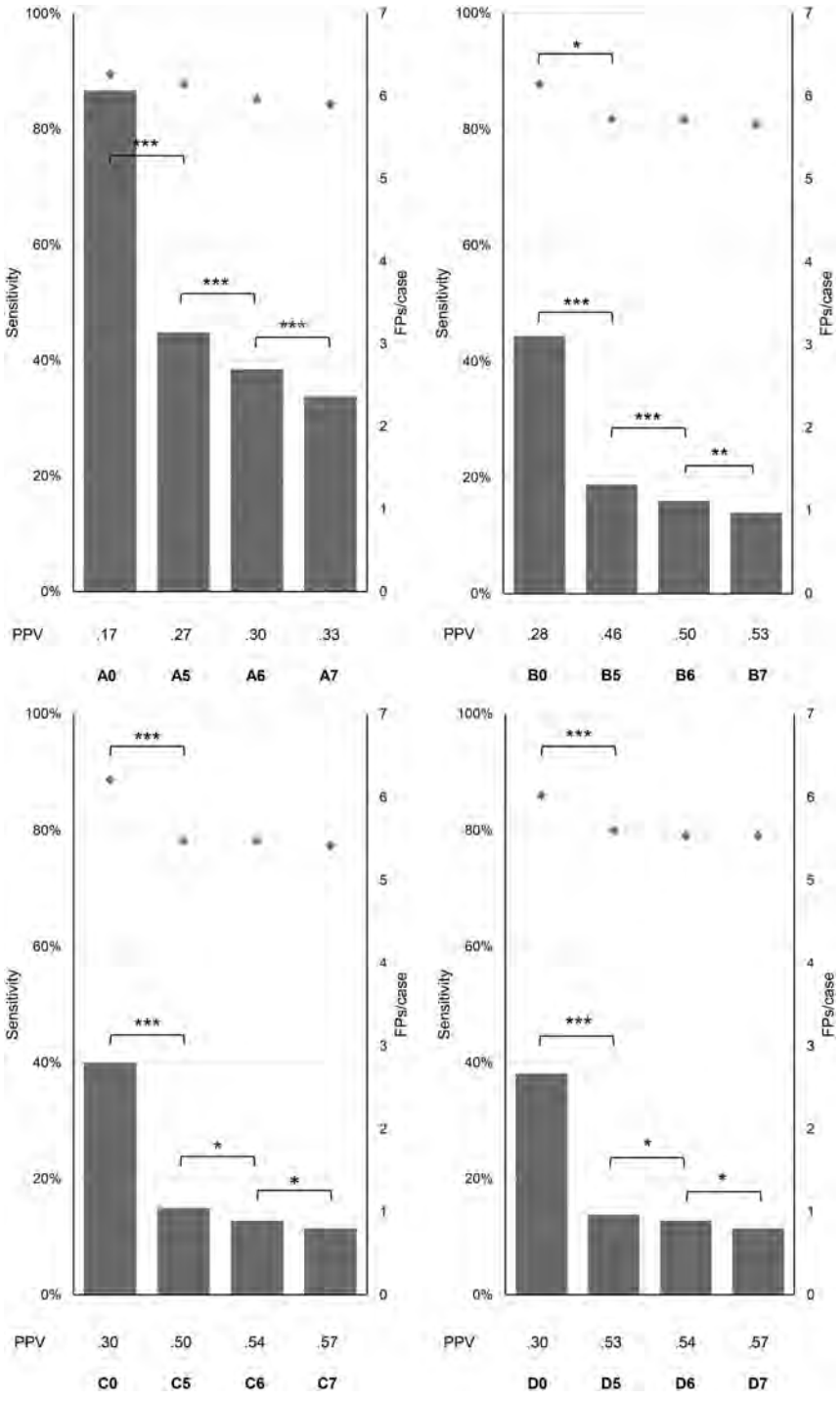


FIG 2. Impact of detection thresholds on sensitivity, the number of false-positives, and the positive predictive value (PPV). Versions A, B, C, D without detection thresholds (A0, B0, C0, D0) and with thresholds of 5 mm^3 (A5, B5, C5, D5), 6 mm^3 (A6, B6, C6, D6), and 7 mm^3 (A7, B7, C7, D7). Depicted as *bars* are the FPs/case; depicted as *diamonds* are the sensitivities. PPV is shown below the diagrams for each model. The *asterisk* indicates $P < .05$; *double asterisks*, $P = .001$; *triple asterisks*, $P < .001$.

Sensitivity within Models

For model A, no significant changes in sensitivity were found between detection thresholds 0, 5, 6, and 7 mm^3 ($P = 1$, binomial distribution used for all comparisons). For model B, a significant decrease in sensitivity was found between thresholds B0 and B5 ($P = .05$, binomial distribution used). Sensitivity did not differ between thresholds B5 and B6, or B6 and B7 ($P = 1$, binomial distribution used for both comparisons). For model C, a significant decrease in sensitivity was found between thresholds C0 and C5 ($P < .001$, binomial distribution used). Sensitivity did not differ between thresholds C5 and C6 or C6 and C7 ($P = 1$, binomial distribution used for both comparisons). For model D, a significant decrease in sensitivity was found between thresholds D0 and D5 ($P < .001$, binomial distribution used). Sensitivity did not differ between thresholds D5 and D6 or D6 and D7 ($P = 1$, binomial distribution used for both comparisons). A consecutive decrease in sensitivity ranged between 2% (version A) and 10% (version C).

False-Positives per Case within Models

Normality was violated for all models without thresholding applied ($P = .008$ for A0, $P < .001$ for all other models).

For model A, significant changes were found in the number of FPs/case among detection thresholds A0 and A5 ($z = 8.14$, $P < .001$), A5 and A6 ($z = 6.16$, $P < .001$), and A6 and A7 ($z = 5.12$, $P < .001$). For model B, significant changes were found in the number of FPs/case among detection thresholds B0 and B5 ($z = 7.89$, $P < .001$), B5 and B6 ($z = 4.12$, $P < .001$), and B6 and B7 ($z = 3.46$, $P = .001$). For model C, significant changes were found in the number of FPs/case among detection thresholds C0 and C5 ($z = 7.62$, $P < .001$), C5 and C6 ($z = 3.61$, $P < .001$), and C6 and C7 ($z = 2.83$, $P = .005$). For model D, significant changes were found in the number of FPs/case among detection thresholds D0 and D5 ($z = 7.64$, $P < .001$), D5 and D6 ($z = 2.45$, $P = .01$), and D6 and D7 ($z = 2.83$, $P = .005$).

Impact of Aneurysm Size

To evaluate the impact of aneurysm size on sensitivity, we divided aneurysms into 3 categories based on maximum diameter, as described above. Detection sensitivity was found to be dependent on aneurysm size (test statistics are shown in Table 1).

The Shapiro-Wilk test revealed that in all cases, normality assumption was violated by the ground truth volumes and/or the predicted volumes of the models. The ground truth volume showed a negative correlation with the predicted volume of each preprocessing

model for the group of small aneurysms. The highest correlation was found in preprocessing model A0 for large aneurysms. The correlation values for all aneurysm sizes combined were, in all models, similar to those of large aneurysms (Table 2).

Impact of Aneurysm Location

Sensitivity values among locations did not show a significant difference (test statistics are shown in Table 3).

Accuracy of Segmentation: DSC and Hausdorff Distance

The distribution of DSCs violated normality for all models and thresholds ($P \leq .001$ for all models). DSCs differed significantly among preprocessing models A0, B0, C0, and D0 (χ^2 [3] = 50.228, $P < .001$). Pair-wise comparisons between nearest neighbors indicated that this difference originated from the difference between A0 and B0 ($z = 5.44$, $P < .001$). DSCs did not differ among sessions B0, C0, and D0.

Hausdorff distances differed significantly among preprocessing models A0, B0, C0, and D0 (F[3, 255] = 56.44, $P < .001$). Pair-wise comparisons between nearest neighbors indicated that this difference originated from the difference between A0 and B0 ($P < .001$). B0 and C0 did not differ significantly ($P = .07$), nor did C0 and D0 ($P = .13$).

DSC and Hausdorff distance values of the different preprocessing models are shown in Table 4. After we fine-tuned model A0, the DSC increased significantly from 0.47 ± 0.28 to 0.50 ± 0.30 ($P < .001$), and the Hausdorff distance changed from 90.16 ± 22.25 to 85.6 ± 22.69 ($P = .004$) without significant changes in sensitivity or the number of FPs/case.

Visual Inspection

Two examples of our dataset are shown in Fig 3. The model was able to detect aneurysms of small-to-large size, location, and regional intensity distribution in the 2 displayed volumes. By means of a postprocessing step, false-positive components were removed.

DISCUSSION

Machine learning applications, in particular deep learning, have recently gained increased attention in the domain of medical imaging. These types of algorithms, specifically CNNs, are top performers in most medical-image analysis competitions. The ease of implementation of CNNs in processing pipelines[13] makes them accessible to a broad range of researchers. Machine learning is becoming a tool of growing importance in radiology and will probably change the way radiologists work.

In this study, we demonstrated the great potential of a CNN for reliable detection of intracranial aneurysms from 3D TOF-MRA. Demand for radiologic imaging is constantly growing; therefore, the steadily increasing workload must be managed by radiology departments.[27] Computer-aided detection tools may assist in preventing diagnostic errors that could occur due to a physician's fatigue or lack of concentration. In a clinical setting, cranial imaging is performed for several diagnostic purposes. However, potentially relevant findings are often missed if a conspicuity corresponding to the primary diagnostic purpose of an examination is found.[28] This phenomenon termed "satisfaction of search" is frequently observed in radiologic practice and could potentially be reduced by sufficient computer-aided detection tools. To evaluate a realistic scenario, we included unspecified and therefore rather heterogeneous images (ie, different scanners, different field strengths) with varying image quality (signal-to-noise ratio, motion artifacts).

Solely in terms of overall sensitivity, the best model was A0, without application of skull-stripping or bias correction, with a sensitivity of 90%. However, this model also had a FPs/case value of 6.1, which is rather high. The highest positive predictive value of 0.57 was achieved with model D7, consisting of customized skull-stripping and N4 bias correction. A sensitivity of 79% was achieved with a FPs/case rate of 0.8 ± 1.3. The amount of preprocessing had a significant impact on the rate of false-positives. In terms of sensitivity, no significant differences between preprocessing models were detected. Using a thresholding method that removes segmentation components below a distinct volume, we were able to further decrease the rate of false-positives.

Aneurysm size had a distinct impact on the performance of the CNN: For small aneurysms, a lower sensitivity value was measured. These missed detections resulted in low correlation values between ground truth volumes and the model-predicted volumes for small aneurysm sizes. This correlation increased for medium-sized aneurysms, which were detected with a higher certainty but in some cases lacked segmentation precision. The correlation for large aneurysms and the overall correlation were high, the latter mainly due to a good segmentation capability for medium and large aneurysms. The DSC could be improved significantly by skull-stripping from 47% ± 28% to 53% ± 29%. The Hausdorff distance likewise improved from a value of 90 ± 22 to 70 ± 17.

Small aneurysms were underrepresented in the dataset; increasing this number would possibly improve the ability of the model to segment those aneurysms and predict their size better. A larger dataset would also decrease a possible overfitting of the model to the training data. We endeavored to address this issue using 5-fold cross-validation and flipping the image as a data augmentation concept.

The ground truth segmentation is subjective and may differ among radiologists. A similar study showed that intra- and interoperator variability of 20% ± 15% and 28% ± 12% was reported for the segmen-

Table 1: Sensitivity depending on aneurysm size and preprocessing model

	≤3 mm (Small) (n = 13)	>3 and ≤7 mm (Medium) (n = 57)	>7 mm (Large) (n = 45)	Fisher Exact Test Statistic
A0	.38	.93	1	29.00, $P < .001$
B0	.38	.91	.98	25.93, $P < .001$
C0	.23	.96	.98	38.43, $P < .001$
D0	.08	.95	.98	49.89, $P < .001$

Table 2: Correlation between ground truth volume and model volume prediction depending on aneurysm size and preprocessing model

	≤3 mm (Small) (n = 13)	>3 and ≤7 mm (Medium) (n = 57)	>7 mm (Large) (n = 45)	Overall
A0	$r_s = -.28$ ($P = .36$)	$r_s = .46$ ($P < .001$)	$r_s = .91$ ($P < .001$)	$r_s = .90$ ($P < .001$)
B0	$r_s = -.03$ ($P = .91$)	$r_s = .45$ ($P < .001$)	$r_s = .87$ ($P < .001$)	$r_s = .87$ ($P < .001$)
C0	$r_s = -.09$ ($P = .78$)	$r_s = .47$ ($P < .001$)	$r_s = .89$ ($P < .001$)	$r_s = .88$ ($P < .001$)
D0	$r_s = -.31$ ($P = .31$)	$r_s = .43$ ($P = .001$)	$r_s = .89$ ($P < .001$)	$r_s = .88$ ($P < .001$)

Note:—r_s indicates the Spearman correlation coefficient.

tation of brain tumors.[29] We attempted to overcome this issue by evaluating our dataset through another radiologist.

Several approaches for automated detection of intracranial aneurysms from noninvasive cranial imaging have been reported previously.[14,15,17] However, most were limited by either the use of conventional computer-aided diagnosis algorithms or being applicable only on 2D images. For instance, Miki et al[14] increased the number of detections of 2 radiologists using a computer-aided diagnosis tool for MRA images. Their system is based on different handcrafted features[30] and reached a sensitivity of 82% in source and reconstructed images of a 3T MR imaging device. Štepán-Buksakowska et al[15] used a computer-aided diagnosis algorithm that applies global thresholding and region-growing schemes. They achieved a mean sensitivity of 83.6% by combining radiologists' examinations with their tool. Nakao et al[17] used a CNN for detecting aneurysms in 2D MIPs. Their tool detected aneurysms with a sensitivity of 94.2% with 2.9 FPs/case. However, their work is limited to 2D projections.

The main limitation of the presented algorithm is poor specificity. We acknowledge that this issue currently limits clinical utility. However, we demonstrated that an algorithm that was originally developed for segmentation tasks is able to detect aneurysms reliably from noninvasive cranial imaging, and this requires only a very limited number of training samples. We observed that several, easily applicable postprocessing steps allow distinct reduction of the number of false-positives. Because data augmentation is already included, we assume that for further improvement of specificity, enlargement of the sample size would be necessary. Given the low number of untreated aneurysms in MRA, this would require a multi-institutional approach. Fine-tuning the network on a larger dataset with a modified training strategy for a more realistic distribution of classes might improve not only the DSC and Hausdorff distance but also sensitivity and specificity.

In this study, the performance of DeepMedic was validated in a clinical dataset, which was based on radiology reports. To further investigate whether our approach might contribute to an improvement of aneurysm detection in a clinical setting, the performance of DeepMedic should be compared with that of human readers. Another limitation is that the algorithm was trained solely on cases that had intracranial aneurysms. Because DeepMedic works as a voxelwise classifier, this was done for methodologic reasons. The algorithm learns to differentiate between physiologic vessel anatomy and aneurysms by classifying each voxel within a volume as a positive (aneurysm) or negative (no aneurysm) prediction. Every dataset includes not only aneurysms but also physiologic vessels. Hence, every aneurysm-free voxel of a brain vessel could be considered a negative finding in a voxelwise classifier; therefore, one could argue that the algorithm can also learn to separate aneurysms from normal vessel anatomy using only pathologic cases. However, given the relatively low prevalence of intracranial aneurysms in the general population, this approach might lead to overprediction, which explains, to some extent, the relatively high number of false-positive cases observed in our study.

To obtain a highly autonomous system, a robust and automated skull-stripping algorithm for TOF sequences is necessary to obtain a reliable brain mask comprising all relevant vessels without extracranial or nonbrain tissues. Most skull-stripping methods perform best with T1-weighted images and need to be adjusted manually for different acquisition sequences.[31] Finally, in further research, it would be advantageous to compare the performance of DeepMedic in terms of aneurysm detection with that of other CNN architectures.

Table 3: Sensitivity of the different models depending on aneurysm location and preprocessing model

	ICA (*n* = 48)	MCA (n = 26)	A (*n* = 19)	P (*n* = 22)	Fisher Exact Test Statistic
A0	.90	.92	.84	.91	.98, $P = .86$
B0	.88	.88	.79	.95	2.52, $P = .48$
C0	.85	.92	.84	.95	2.09, $P = .59$
D0	.83	.92	.79	.91	2.27, $P = .53$

Note:—A indicates the anterior cerebral arteries (including the anterior communicating artery); P, posterior circulation (including vertebral, basilar, posterior, cerebral and posterior communicating arteries).

Table 4: Mean DSC and mean Hausdorff distance depending on the preprocessing model

	DSC (SD)	Hausdorff Distance (SD)
A0	.47 (.28)	90.16 (22.25)
B0	.53 (.29)	70.20 (16.58)
C0	.53 (.30)	65.40 (18.89)
D0	.53 (.31)	69.67 (19.08)

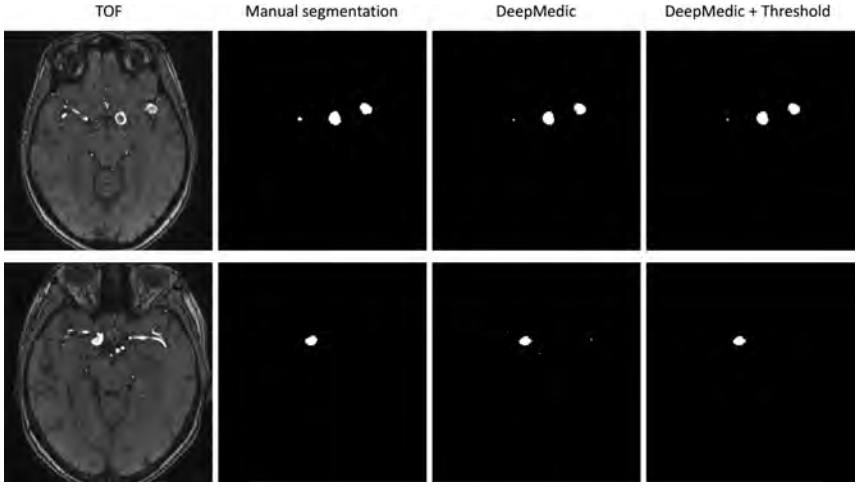


FIG 3. Results of the DeepMedic inference and thresholding method. Illustrated are 2 different subjects (top/bottom). In these volumes, aneurysms of different sizes with heterogenic and homogeneous intensity distributions are detected. After we remove small components below a certain volume, false-positives are removed sufficiently.

CONCLUSIONS

This study demonstrates that our CNN-based system can detect intracranial aneurysms with high sensitivity in a 3D TOF-MRA dataset. The dataset, comprising acquisitions of different field strengths and variable image quality, was created to evaluate a scenario similar to clinical reality. Adequate pre- and postprocessing significantly reduced the number of false-posi-

tives. The predicted aneurysm volume correlated well with the ground truth volume for medium- and large-sized aneurysms; hence, the system could also serve as a tool to predict aneurysm size.

ACKNOWLEDGMENT

We thank Marguerite Müller for providing patient data to our dataset.

Disclosures: Anton Faron—*UNRELATED*: *Employment*: University Hospital Bonn Germany, *Comments*: employed as resident physician, Department of Radiology. Nikolas Teichert—*UNRELATED*: *Employment*: University Hospital of Aachen, *Comments*: work as a doctor in the Department of Diagnostic and Interventional Neuroradiology. Martin Wiesmann—*UNRELATED*: *Consultancy*: Stryker Neurovascular; *Payment for Lectures Including Service on Speakers Bureaus*: Bracco Imaging, Medtronic, Siemens, Stryker Neurovascular; *Payment for Development of Educational Presentations*: Abbott, ab medica, Acandis, Bayer Healthcare AG, Bracco Imaging, B. Braun Medical, Codman Neurovascular, Kaneka Pharma, Medtronic, Dahlhausen, MicroVention, Nvidia, Penumbra, phenox, Philips Healthcare, Route 92 medical, Siemens, Silk Road Medical, St. Jude, and Stryker Neurovascular*. Money paid to the institution.

REFERENCES

1. Vlak MH, Algra A, Brandenburg R, et al. **Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis.** *Lancet Neurol* 2011;10:626–36 CrossRef Medline
2. Kapsalaki EZ, Rountas CD, Fountas KN. **The role of 3 Tesla MRA in the detection of intracranial aneurysms.** *Int J Vasc Med* 2012;2012: 792834 CrossRef Medline
3. Nieuwkamp DJ, Setz LE, Algra A, et al. **Changes in case fatality of aneurysmal subarachnoid haemorrhage over time, according to age, sex, and region: a meta-analysis.** *Lancet Neurol* 2009;8:635–42 CrossRef Medline
4. Lindgren AE, Koivisto T, Björkman J, et al. **Irregular shape of intracranial aneurysm indicates rupture risk irrespective of size in a population-based cohort.** *Stroke* 2016;47:1219–26 CrossRef Medline
5. Thompson BG, Brown RD Jr, Amin-Hanjani S, et al; American Heart Association Stroke Council, Council on Cardiovascular and Stroke Nursing, and Council on Epidemiology and Prevention; American Heart Association; American Stroke Association. **Guidelines for the management of patients with unruptured intracranial aneurysms: a guideline for healthcare professionals from the American Heart Association/American Stroke Association.** *Stroke* 2015;46:2368–400 CrossRef Medline
6. Wiebers DO, Whisnant JP, Huston J 3rd, et al; International Study of Unruptured Intracranial Aneurysms Investigators. **Unruptured intracranial aneurysms: natural history, clinical outcome, and risks of surgical and endovascular treatment.** *Lancet* 2003;362:103–10 CrossRef Medline
7. Leffers AM, Wagner A. **Neurologic complications of cerebral angiography: a retrospective study of complication rate and patient risk factors.** *Acta Radiol* 2000;41:204–10 Medline
8. Wang S, Summers RM. **Machine learning and radiology.** *Med Image Anal* 2012;16:933–51 CrossRef Medline
9. McDonald RJ, Schwartz KM, Eckel LJ, et al. **The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload.** *Acad Radiol* 2015;22:1191–98 CrossRef Medline
10. Erickson BJ, Korfiatis P, Akkus Z, et al. **Toolkits and libraries for deep learning.** *J Digit Imaging* 2017;30:400–05 CrossRef Medline
11. Larson DB, Chen MC, Lungren MP, et al. **Performance of a deep learning neural network model in assessing skeletal maturity on pediatric hand radiographs.** *Radiology* 2018;287:313–322 CrossRef Medline
12. Krizhevsky A, Sutskever I, Hinton GE. **ImageNet classification with deep convolutional neural networks.** *Adv Neural Inf Process Syst* 2012:1–9 https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf. Accessed March 20, 2018
13. Litjens G, Kooi T, Bejnordi BE, et al. **A survey on deep learning in medical image analysis.** *Med Image Anal* 2017;42:60–88 CrossRef Medline
14. Miki S, Hayashi N, Masutani Y, et al. **Computer-assisted detection of cerebral aneurysms in MR angiography in a routine image-reading environment: effects on diagnosis by radiologists.** *AJNR Am J Neuroradiol* 2016;37:1038–43 CrossRef Medline
15. Štepán-Buksakowska IL, Accurso JM, Diehn FE, et al. **Computer-aided diagnosis improves detection of small intracranial aneurysms on MRA in a clinical setting.** *AJNR Am J Neuroradiol* 2014;35: 1897–902 CrossRef Medline
16. Yang X, Blezek DJ, Cheng LT, et al. **Computer-aided detection of intracranial aneurysms in MR angiography.** *J Digit Imaging* 2011;24: 86–95 CrossRef Medline
17. Nakao T, Hanaoka S, Nomura Y, et al. **Deep neural network-based computer-assisted detection of cerebral aneurysms in MR angiography.** *J Magn Reson Imaging* 2018;47:948–53 CrossRef Medline
18. Kamnitsas K, Ledig C, Newcombe VF, et al. **Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation.** *Med Image Anal* 2017;36:61–78 CrossRef Medline
19. Kamnitsas K. GitHub-Kamnitsask/deepmedic. Efficient Multi-Scale 3D Convolutional Neural Network for Brain Lesion Segmentation. Published 2017. https://github.com/Kamnitsask/deepmedic. Accessed March 20, 2018
20. Jenkinson M, Pechaud M, Smith S. **BET2: MR-based estimation of brain, skull and scalp surfaces.** In: *Eleventh Annual Meeting of the Organization for Human Brain Mapping*, Toronto, Ontario, Canada, June 12–16, 2005
21. Tustison NJ, Avants BB, Cook PA, et al. **N4ITK: improved N3 bias correction.** *IEEE Trans Med Imaging* 2010;29:1310–20 CrossRef Medline
22. 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
23. Kamnitsas K, Chen LC, Ledig C, et al. **Ischemic stroke lesion segmentation.** In: *Proceedings of MICCAI-ISLES*, Munich, Germany. October 5, 2015. 2015:13–16. www.isles-challenge.org. Accessed March 19, 2018
24. Taha AA, Hanbury A. **Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool**. *BMC Med Imaging* 2015; 15:29 CrossRef Medline
25. Tustison NJ, Avants BB, Cook PA, et al. **N4ITK: improved N3 bias correction.** *IEEE Trans Med Imaging* 2010;29:1310–20 CrossRef Medline
26. Li M, Zhu Y, Song H, et al. **Subarachnoid hemorrhage in patients with good clinical grade: accuracy of 3.0-T MR angiography for detection and characterization.** *Radiology* 2017;284:191–99 CrossRef Medline
27. MacDonald SL, Cowan IA, Floyd R, et al. **Measuring and managing radiologist workload: application of lean and constraint theories and production planning principles to planning radiology services in a major tertiary hospital.** *J Med Imaging Radiat Oncol* 2013;57: 544–50 CrossRef Medline
28. Berbaum KS, Franken EA Jr, Dorfman DD, et al. **Satisfaction of search in diagnostic radiology.** *Invest Radiol* 1990;25:133–40 CrossRef Medline
29. Mazzara GP, Velthuizen RP, Pearlman JL, et al. **Brain tumor target volume determination for radiation treatment planning through automated MRI segmentation.** *Int J Radiat Oncol Biol Phys* 2004;59: 300–12 CrossRef Medline
30. Nomura Y, Masutani Y, Miki S, et al. **Performance improvement in computerized detection of cerebral aneurysms by retraining classifier using feedback data collected in routine reading environment.** *J Biomed Graph Comput* 2014;4:12–21 CrossRef
31. Kalavathi P, Prasath VB. **Methods on skull stripping of MRI head scan images: a review.** *J Digit Imaging* 2016;29:365–79 CrossRef Medline
32. Kamnitsas K. DeepMedic GitHub documentation. Documentation about pretraining/finetuning. Published 2017. https://github.com/Kamnitsask/deepmedic/blob/master/documentation/dmRes.png. Accessed July 30, 2018

Automated ASPECTS on Noncontrast CT Scans in Patients with Acute Ischemic Stroke Using Machine Learning

H. Kuang, M. Najm, D. Chakraborty, N. Maraj, S.I. Sohn, M. Goyal, M.D. Hill, A.M. Demchuk, B.K. Menon, and W. Qiu

ABSTRACT

BACKGROUND AND PURPOSE: Alberta Stroke Program Early CT Score (ASPECTS) was devised as a systematic method to assess the extent of early ischemic change on noncontrast CT (NCCT) in patients with acute ischemic stroke (AIS). Our aim was to automate ASPECTS to objectively score NCCT of AIS patients.

MATERIALS AND METHODS: We collected NCCT images with a 5-mm thickness of 257 patients with acute ischemic stroke (<8 hours from onset to scans) followed by a diffusion-weighted imaging acquisition within 1 hour. Expert ASPECTS readings on DWI were used as ground truth. Texture features were extracted from each ASPECTS region of the 157 training patient images to train a random forest classifier. The unseen 100 testing patient images were used to evaluate the performance of the trained classifier. Statistical analyses on the total ASPECTS and region-level ASPECTS were conducted.

RESULTS: For the total ASPECTS of the unseen 100 patients, the intraclass correlation coefficient between the automated ASPECTS method and DWI ASPECTS scores of expert readings was 0.76 (95% confidence interval, 0.67–0.83) and the mean ASPECTS difference in the Bland-Altman plot was 0.3 (limits of agreement, −3.3, 2.6). Individual ASPECTS region-level analysis showed that our method yielded $\kappa = 0.60$, sensitivity of 66.2%, specificity of 91.8%, and area under curve of 0.79 for 100 × 10 ASPECTS regions. Additionally, when ASPECTS was dichotomized (>4 and ≤4), $\kappa = 0.78$, sensitivity of 97.8%, specificity of 80%, and area under the curve of 0.89 were generated between the proposed method and expert readings on DWI.

CONCLUSIONS: The proposed automated ASPECTS scoring approach shows reasonable ability to determine ASPECTS on NCCT images in patients presenting with acute ischemic stroke.

ABBREVIATIONS: AIS = acute ischemic stroke; AUC = area under curve; ICC= intraclass correlation coefficient; IQR= interquartile range

Management of patients with acute ischemic stroke (AIS) relies heavily on an assessment of the extent of irreversibly injured brain at baseline. Patients with extensive early ischemic changes at presentation are unlikely to benefit from thrombolysis or thrombectomy procedures. Moreover, such patients may also be at higher risk of developing complications of treatment such as intracerebral hemorrhage. The Alberta Stroke Program Early CT Score was devised as a systematic method of assessing the extent of early ischemic change on noncontrast CT in patients with AIS.[1,2] Across the years, ASPECTS has gained credence and is now used the world over for this purpose,[3-7] though it has not been proved useful for selecting patients for treatment.[8,9]

Although conceptually, the ASPECTS is a simple method, scoring early ischemic change on NCCT scans continues to be a challenge, especially for readers with less experience.[10-12] Technical factors such as peak x-ray energy (kiloelectron volt/megaelectron volt) image processing and display procedures; patient factors such as old infarcts, brain atrophy, and leukoaraiosis; and reader factors such as experience, training, and specialty, all potentially affect ASPECTS interpretation.[11,12] A solution to improve ASPECTS reading is training readers to recognize these issues while providing them with strategies that can help improve the reliability and validity of these reads. Another solution is to use novel technologies such as machine

Received June 20, 2018; accepted after revision October 8.

From the Calgary Stroke Program (H.K., W.Q., M.N., D.C., N.M., M.G., M.D.H., A.M.D., B.K.M.), Department of Clinical Neurosciences, Department of Radiology (M.D.H., A.M.D., M.G., B.K.M.), and Department of Community Health Sciences (M.D.H., B.K.M.), University of Calgary, Calgary, Alberta, Canada; Hotchkiss Brain Institute, Calgary, Alberta, Canada (M.D.H., A.M.D., M.G., B.K.M.); and Department of Neurology (S.I.S.), Keimyung University, Daegu, South Korea.

This study was funded through an operating grant from the Canadian Institute of Health Research.

Please address correspondence to Wu Qiu, PhD, Department of Clinical Neurosciences, University of Calgary, Room 1079, 10th Floor Foothills Medical Centre, 1403 29th Street NW, Calgary, AB, Canada T2N 2T9; e-mail: qiu.wu.ch@gmail.com.

Indicates article with supplemental on-line appendix and table.

http://dx.doi.org/10.3174/ajnr.A5889

learning and feature extraction to develop automated solutions to ASPECTS interpretation.[13-17]

In recent years, evidence that automated ASPECTS scoring methods based on machine learning are comparable with expert reading of ASPECTS is accumulating.[18-24] In this study, we developed an automated ASPECTS scoring system based on machine learning and feature engineering and compared it with expert ASPECTS readings on acute DWI. We introduced multiple high-order computational textural features into our machine learning model and hypothesized that this automated method can determine ASPECTS scores accurately and reliably compared with expert ASPECTS readings on acute DWI.

MATERIALS AND METHODS

Data are from the Keimyung Stroke Registry, an ongoing single-center prospective cohort study of patients with acute ischemic stroke presenting to the Keimyung University Hospital in Daegu, South Korea. Two hundred fifty-seven patients with acute ischemic stroke presenting within 8 hours of last known well who had baseline NCCT (slice thickness, ≤ 5 mm) followed by DWI performed within 1 hour of NCCT were included in the study. An expert scored ASPECTS on DWI; any individual region with diffusion restriction occupying >20% of that region was considered affected. To assess the reliability of expert-reading DWI ASPECTS, another expert was asked to score 60 DWI scans randomly selected from the 257 patients with AIS.

Of the 257 patient images, 157 were randomly selected for training a machine learning model, while the remaining 100 images were used to evaluate the trained model. Specifically, a NCCT template with ASPECTS regions manually contoured was nonlinearly registered onto all NCCT images (Fig 1). During the training stage, 376 texture features (details of texture features are shown in the On-line Appendix) such as high-order statistics and image textural features were extracted from each ASPECTS region from the 157 patient images bilaterally after median filtering. Note that the feature extraction and classification for each ASPECTS region were performed in 3D. Information on the side of the brain affected by ischemic stroke was used as an additional input to compute difference features between ischemic and normal brain tissue. Specifically, observers first determined the ischemic hemisphere based on imaging and clinical parameters. Feature differences were then obtained by subtracting regional level values on the ischemic side from the those on the contralateral side. Sixty patients (38 in the training dataset and 22 in the testing dataset) with posterior circulation strokes were included intentionally to reflect clinical reality. In patients with posterior circulation strokes, the left side was regarded as the default ischemic side.

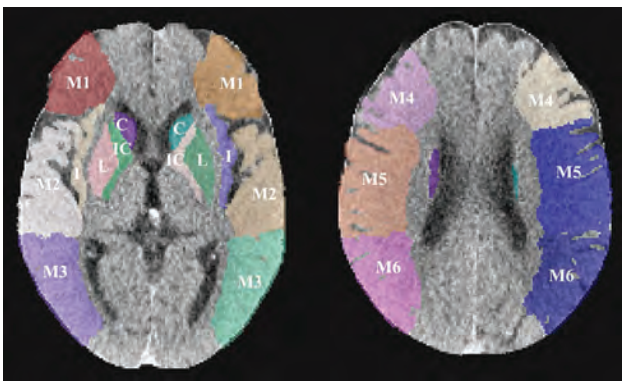


FIG 1. Examples of each ASPECTS region. L indicates lentiform; I, insula; C, caudate; IC, internal capsule; M, MCA.

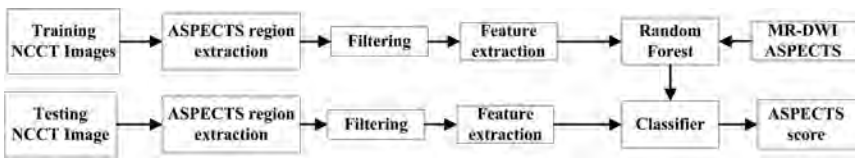


FIG 2. A flowchart of the training and testing processes used in the study for each ASPECTS region.

The computed features were first ranked using linear discriminant analysis. The ranked features were input into a random forest model using the expert-assessed ASPECTS on DWI as a class label. We used 5-fold cross-validation on the training samples to select training hyperparameters including the number of trees in the forest, the maximum depth of trees, and also the number of ranked features. Class weight was set to deal with the imbalanced data distribution on the basis of the ratio of abnormal and normal samples in the training data. The detailed parameter settings are shown in the On-line Table. We trained a classifier for each ASPECTS region. The random forest training and testing were implemented using Scikit-learn in Python (http://scikit-learn.org/stable/). The trained random forest classifier was then validated on the remaining 100 test patient images. A flowchart of the training and testing process of each ASPECTS region is shown in Fig 2.

Statistical Analysis

Expert ASPECTS readings on DWI of the 100 test images were used as the ground truth to evaluate the automated ASPECTS obtained by our method. Agreement on the total ASPECTS score was measured using the intraclass correlation coefficient (ICC). Boxplots and Bland-Altman plots were used to illustrate differences in the assessment of total ASPECTS between the automated method and the ground-truth (expert-read DWI ASPECTS). The ICC analysis was also stratified by stroke onset-to-CT time (≤90 minutes, $n = 69$; 90–270 minutes, $n = 21$; and >270 minutes, $n = 10$). Because physicians use the presence or absence of extensive early ischemic changes to make clinical decisions on treatment in patients with acute ischemic stroke, we also assessed agreement on the ASPECTS interpretation between the automated method and DWI using κ statistics on a dichotomized ASPECTS threshold (>4 versus ≤4).[25] κ statistics were also used to assess agreement between the automated method and expert-read DWI at each individual ASPECTS region.

Receiver operating characteristics based on the MedCalc for Windows software (MedCalc Software, Mariakerke, Belgium) were used to report the area under the curve (AUC) for the dichotomized ASPECTS (>4 versus ≤4), and individual region-level ASPECTS analysis using automated ASPECTS, as an independent variable and expert-read DWI ASPECTS, as a dependent variable. A clustered receiver oper-

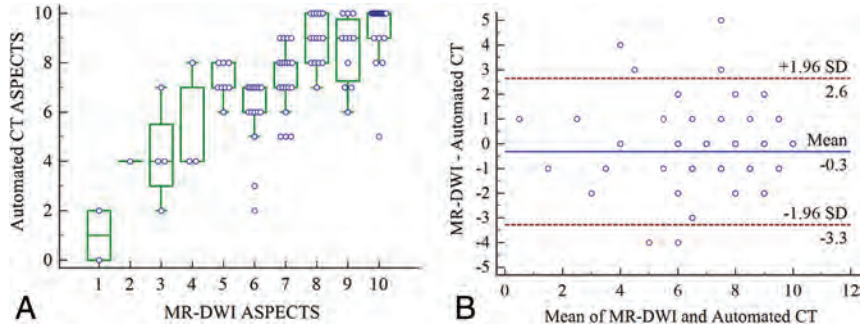


FIG 3. *A*, Boxplot with a scatterplot showing the distribution of the automated CT ASPECTS at each individual ASPECTS on DWI. *B*, A Bland-Altman plot illustrating agreement between a total automated ASPECTS score and ASPECTS scores on DWI. Random jitter has been added to illustrate the number of measurements at each ASPECTS point. The *horizontal black line* represents the mean difference in the ASPECTS score between the 2 methods, while the *dotted lines* represent a 1.96 SD around the difference.

ating characteristic method in R statistical and computing software (http://www.r-project.org) was used to report the AUC for grouped ASPECTS regions.[26] In addition, accuracy defined as the ratio of accurately classified and total samples, sensitivity, and specificity was also calculated to further measure the performance of our proposed ASPECTS method.

A linear-weighted κ of the trichotomized ASPECTS (0–4, 5–7, 8–10) was computed. A sensitivity analysis was performed by varying threshold involvement of each ASPECTS region on expert-read DWI as >0% and >50% involvement compared with the >20% involvement used for primary analyses. Additionally, to demonstrate the efficacy of the developed automated ASPECTS method, we compared the ASPECTS reading of a stroke expert on the 100 test images with the automated ASPECTS and the expert-assessed ASPECTS on DWI. All statistical analyses were performed by using MedCalc 17.8 and Matlab (MathWorks, Natick, Massachusetts). A 2-sided $\alpha < .05$ was considered statistically significant.

RESULTS

Of 157 patients included in the training dataset (median age, 69 years; interquartile range [IQR], 62–76 years; 54.8% male), baseline NCCT was performed within a median time of 46.5 minutes (IQR, 27–117 minutes) from last known well compared with a median baseline NCCT to baseline MR imaging time of 39.5 minutes (IQR, 30–51 minutes). Of 100 patients included in the test dataset (median age, 70 years; IQR, 64–77 years; 56% male), baseline NCCT was performed within a median time of 49 minutes from last known well (IQR, 23.8–95.5 minutes) compared with a median baseline NCCT to baseline MR imaging time of 39 minutes (IQR, 29–50.3 minutes). The median baseline ASPECTS on the training dataset using DWI was 8 (IQR, 6–9).

The κ values for regional and dichotomized ASPECTS between the 2 expert-read DWI ASPECTSs were 0.86 (95% CI, 0.81–0.91) and 0.85 (95% CI, 0.56–1), respectively. The ICC for total ASPECTS between the 2 expert-read DWI ASPECTSs was 0.90 (95% CI, 0.84–0.94).

When patients were stratified by stroke onset-to-CT time (≤90 minutes, $n = 69$; 90–270 minutes, $n = 21$; and >270 minutes, $n = 10$), the ICCs between the automated CT ASPECTS and the DWI ASPECTS for these 3 subgroups were 0.80 (95% CI, 0.69–0.87), 0.77 (95% CI, 0.52–0.90), and 0.26 (95% CI, −0.40–0.75), respectively. No evidence was found that the ICC increased in patients with longer onset-to-CT time using the automated ASPECTS method.

The median baseline ASPECTS generated by the automated method on test data ($n = 100$) was 8 (IQR, 7–9) versus a score of 7 (IQR, 6–9) on the ground truth DWI. Figure 3*A* shows a boxplot overlaid with a scatterplot showing the distribution of the automated CT ASPECTS at each individual ASPECTS on DWI. The intraclass correlation coefficient for total ASPECTS between the automated method and DWI was 0.76 (95% CI, 0.67–0.83). Figure 3*B* illustrates Bland-Altman agreement plots between the automated method and DWI for total ASPECTS. The mean difference in total ASPECTS between the automated method and DWI was minimal (0.3; limit of agreement, −3.3, 2.6).

Agreement on ASPECTS between the automated method and DWI using a dichotomized ASPECTS threshold of >4 versus ≤4 was good ($\kappa = 0.78$: 95% CI, 0.57–0.99). Sensitivity (97.8%: 95% CI, 92.2%–99.7%), specificity (80%: 95% CI, 34.8%–93.3%), F1 measure (0.98), and AUC (0.89: 95% CI, 0.81–0.94) were reasonably good. When ASPECTS was trichotomized (0–4, 5–7, 8–10), the agreement between the automated method and DWI was good as well (linear weighted $\kappa = 0.66$; 95% CI, 0.53–0.80). Agreement on ASPECTS between the automated method and DWI at the individual ASPECTS region level is reported in Table 1. κ for agreement between the 2 methods ranged from 0.36 to 0.64. The automated ASPECTS method demonstrated high specificity but modest sensitivity compared with DWI at the regional level. F1 measures are also shown in Table 1.

Sensitivity analysis was attempted by varying the threshold involvement of each ASPECTS region on expert-read DWI as >0% and >50% involvement in addition to the >20% involvement used for the primary analyses. Region-level agreement between the automated method and expert-rated DWI ASPECTS for all 3 thresholds is shown in Table 2. Agreement between the 2 methods was best when DWI ASPECTS was rated using the >50% threshold method.

The ICC for total ASPECTS between the expert-rated NCCT and DWI was 0.67 (95% CI, 0.55–0.77). The agreement between the expert-rated CT ASPECTS and DWI using a dichotomized ASPECTS threshold of >4 versus ≤4 was modest ($\kappa = 0.42$; 95% CI, 0.22–0.62). Sensitivity (81.1%: 95% CI, 71.5%–88.6%), specificity (90%: 95% CI, 55.5%–99.7%), F1 measure (0.89), and AUC (0.85: 95% CI, 0.77–0.92) were obtained.

The ICC for total ASPECTS between the expert-rated NCCT and the automated CT ASPECTS was 0.61 (95% CI, 0.47–0.72). The agreement between the expert-rated CT ASPECTS and the automated CT ASPECTS using a dichotomized ASPECTS threshold of >4 versus ≤4 was modest ($\kappa = 0.48$; 95% CI, 0.28–0.68). An example of expert-rated DWI ASPECTS, our automated CT ASPECTS, and expert-rated CT ASPECTS is shown in Fig 4.

Table 1: κ, accuracy, F1 measure, sensitivity, specificity, and AUC on each ASPECTS region

Region	κ (95% CI)	Accuracy (%) (95% CI)	F1 Measure	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	AUC (95% CI)
M1	0.59 (0.38–0.81)	90 (90/100) (84.1–95.9)	0.64	47.4 (9/19) (24.4–71.1)	100 (81/81) (95.5–100)	0.74 (0.64–0.82)
M2	0.52 (0.35–0.68)	76 (76/100) (67.6–84.4)	0.73	76.2 (32/42) (60.5–87.9)	75.9 (44/58) (62.8–86.1)	0.76 (0.67–0.84)
M3	0.47 (0.21–0.73)	88 (88/100) (81.6–94.4)	0.54	50 (7/14) (23–77)	94.2 (81/86) (87–98.1)	0.72 (0.62–0.81)
M4	0.36 (0.13–0.63)	85 (85/100) (78–92)	0.35	36.4 (4/11) (10.9–69.2)	91.1 (81/89) (83.1–96)	0.64 (0.54–0.73)
M5	0.54 (0.37–0.7)	77 (77/100) (68.8–85.3)	0.74	68.1 (32/47) (52.9–80.9)	84.9 (45/53) (72.4–93.3)	0.77 (0.67–0.84)
M6	0.39 (0.14–0.64)	86 (86/100) (79.2–92.8)	0.46	35.3 (6/17) (14.2–61.7)	96.4 (80/83) (89.8–99.2)	0.66 (0.56–0.75)
Lentiform	0.64 (0.47–0.81)	85 (85/100) (78–92)	0.75	71.0 (22/31) (52–85.8)	91.3 (63/69) (82–96.7)	0.81 (0.72–0.88)
Insula	0.62 (0.46–0.77)	81 (81/100) (73.3–88.7)	0.83	85.5 (47/55) (73.3–93.5)	75.6 (34/45) (60.5–87.1)	0.81 (0.71–0.88)
Caudate	0.63 (0.42–0.84)	90 (90/100) (84.1–95.9)	0.69	57.9 (11/19) (33.5–79.7)	97.5 (79/81) (91.4–99.7)	0.78 (0.68–0.85)
Internal capsule	0.59 (0.35–0.83)	91 (91/100) (85.4–96.6)	0.64	57.1 (8/14) (28.9–82.3)	96.5 (83/86) (90.1–99.3)	0.77 (0.67–0.85)
All regions	0.60 (0.54–0.66)	84.9 (849/1000) (82.7–87.1)	0.70	66.2 (178/269) (60.2–71.8)	91.8 (671/731) (89.6–93.7)	0.79 (0.75–0.83)

Table 2: Agreement on ASPECTS interpretation at a regional level and for dichotomized ASPECTS (>4 vs. ≤4) between the automated ASPECTS method and expert-read DWI ASPECTS using different DWI ASPECTS region-involvement thresholds

DWI ASPECTS Region-Involvement Thresholds	κ (95% CI)	Accuracy (%) (95% CI)	F1 Measure	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	AUC (95% CI)
20%						
All regions	0.6 (0.54–0.66)	84.9 (849/1000) (82.7–87.1)	0.70	66.2 (178/269) (60.2–71.8)	91.8 (671/731) (89.6–93.7)	0.79 (0.75–0.83)
>4 and ≤4	0.78 (0.57–0.99)	96 (96/100) (92.2–99.8)	0.98	97.8 (88/90) (92.2–99.7)	80 (8/10) (34.8–93.3)	0.89 (0.81–0.94)
50%						
All regions	0.64 (0.57–0.70)	88.8 (888/1000) (86.9–90.8)	0.71	68.2 (133/195) (61.2–74.7)	93.8 (755/805) (91.9–95.4)	0.81 (0.77–0.85)
>4 and ≤4	1 (1–1)	100 (100/100) (100–100)	1	100 (94/94) (96.2–100)	100 (6/6) (54.1–100)	1 (0.96–1)
0%						
All regions	0.56 (0.51–0.61)	79.5 (795/1000) (77–82)	0.72	69.7 (264/379) (64.8–74.2)	85.5 (531/621) (82.5–88.2)	0.78 (0.76–0.79)
>4 and ≤4	0.46 (0.26–0.66)	81 (81/100) (73.3–88.7)	0.88	93.2 (68/73) (84.7–97.7)	48.2 (13/27) (28.7–68.1)	0.71 (0.61–0.79)

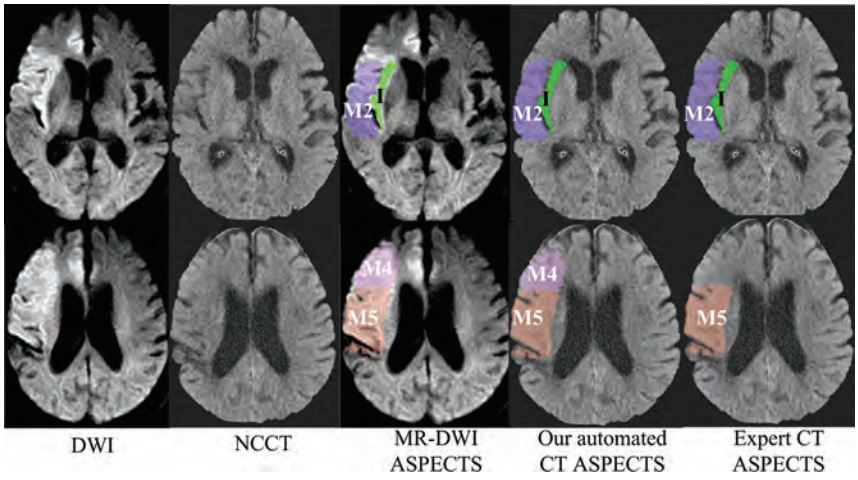


FIG 4. Examples of DWI ASPECTS, the automated CT ASPECTS derived in this study, and expert-read CT ASPECTS. ASPECTS regions with ischemic changes are shown in color.

DISCUSSION

Results from these analyses using 100 patient images show that the automated ASPECTS method proposed in this article agrees well with expert-read DWI ASPECTS at a regional level and for the total ASPECTS. Moreover, good agreement between the automated method and expert-read DWI ASPECTS for ASPECTS cut-points (>4 versus ≤4) may help evaluate patients for the presence or absence of large infarcts at baseline. These results also show that the automated ASPECTS method is not inferior to expert-read ASPECTS on NCCT.

A commercially available automated ASPECTS scoring system (e-ASPECTS; https://brainomix.com/e-aspects) based on a machine learning algorithm has shown an ability to detect early ischemic changes on NCCT at a level similar to that of junior stroke physicians while being noninferior to neuroradiologists.[19] That study used ASPECTS on baseline and follow-up CT scans as the ground truth for comparison. e-ASPECTS was further evaluated in a small study of 34 patients in which baseline CT and DWI scans were obtained <2 hours apart.[20] Another automated ASPECTS system combining filtering, bi-level and regional growth, feature selection, and a support vector machine was tested on 40 patients with AIS using DWI scans as the ground truth.[24] This method obtained a κ of 0.52 for dichotomized ASPECTS (>7 versus ≤7). However, the NCCT and diffusion-weighted imaging time was not reported, making it difficult to evaluate its clinical applicability. Other methods of scoring ASPECTS automatically have mostly been tested against expert-read ASPECTS on NCCT. A major strength of this study is the use of an ASPECTS read on acute DWI by an expert as the ground truth to validate the automated ASPECTS method. This assures that the validity of the automated method was tested to a very high standard.

The proposed automated ASPECTS scoring method is based on feature engineering and random forest learning. Random forest is considered one of the most recent and popular boosting methods and has proved classification performance for difficult problems in many medical image–analysis applications compared with other classifiers.[27,28] Random forest is an ensemble learning method that combines multiple weak classifiers (decision trees) and lets these decision trees vote for the most popular class. Each tree in the forest relies on a random vector sampled independently, and all trees in the forest have the same distribution. The growth of the tree is governed by

random vectors. A measure of randomness is introduced into the training, which can prevent the training classifier from getting stuck at a local minimum, thereby improving the accuracy and reducing the chances of overfitting.

Some previous automated methods have used first-order image features, such as Hounsfield unit (HU) or density and HU difference between the ischemic and contralateral side as features for their algorithms. These first-order image features have limitations in patients with subtle ischemic changes and when images have low signal-to-noise ratios and motion artifacts. The use of multiple higher order computational textural features as part of the machine learning algorithms in the automated ASPECTS method proposed here helps us improve the validity of our technique.

This study has some limitations. First, of the 157 training images randomly selected, only 26.1% (410/1570) of ASPECTS regions had ischemic changes versus 73.9% (1160/1570) normal ASPECTS regions. Improving the performance by tackling the imbalance in data distribution is a goal for our machine learning algorithms. Second, this analysis used imaging data from 1 site. NCCT image acquisition and quality vary across sites; we will, therefore, need to validate the automated ASPECTS method in other data from other sites. Third, only 10 patients had ASPECTS $\leq$4 of 100 test patients, thus raising some valid concerns about the stability of results in the dichotomized ASPECTS analysis. Validation on a larger dataset is required to demonstrate the robustness of these results. Fourth, only a single atlas was used to localize the ASPECTS regions, which might not be optimal for all patients compared with using a method based on multiple atlases. However, localization based on nonlinear registration using a single atlas can maintain a good trade-off between computational cost and accuracy because saving time is critical in the acute stroke setting. Improving registration accuracy using a single atlas remains an open problem for brain imaging despite some existing atlas-selection techniques.[29]

CONCLUSIONS

The automated ASPECTS method developed here could accurately and reliably assign ASPECTS on baseline NCCT scans in patients presenting with acute ischemic stroke. This work therefore further validates the utility of machine learning algorithms in developing software that can help and support physicians in interpreting brain scans of patients with acute ischemic stroke.

Disclosures: Mayank Goyal—*RELATED*: *Other*: GE Healthcare, *Comments*: licensing agreement concerning systems of stroke diagnosis; *UNRELATED: Consultancy*: Medtronic, Stryker, MicroVention, Mentice, *Comments*: consulting and advice concerning acute stroke management; *Grants/Grants Pending*: Stryker, *Comments*: funding for the UNMASK EVT study.* Michael D. Hill—*UNRELATED*: *Consultancy*: Merck, *Comments*: Adjudication Board for stroke outcomes for clinical trials in diabetes; *Grants/Grants Pending*: Medtronic, Stryker, Boehringer Ingelheim, Alberta Innovates, Canadian Institutes for Health Research, *Comments*: multiple grants for stroke clinical trials*; *Stock/Stock Options*: Calgary Scientific Inc, *Comments*: stock ownership in a private image-processing company. Bijoy K. Menon—*RELATED*: *Grant*: Canadian Institute of Health Research, *Comments*: operating grant from the Canadian Institute of Health Research.* *Money paid to the institution.

REFERENCES

1. Barber PA, Demchuk AM, Zhang J, et al. **Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy: ASPECTS Study Group—Alberta Stroke Programme Early CT Score.** *Lancet* 2000; 355:1670–74 CrossRef Medline
2. Pexman JW, Barber PA, Hill MD, et al. **Use of the Alberta Stroke Program Early CT Score (ASPECTS) for assessing CT scans in patients with acute stroke.** *AJNR Am J Neuroradiol* 2001;22:1534–42 Medline
3. Aoki J, Kimura K, Shibazaki K, et al. **DWI-ASPECTS as a predictor of dramatic recovery after intravenous recombinant tissue plasminogen activator administration in patients with middle cerebral artery occlusion.** *Stroke* 2013;44:534–37 CrossRef Medline
4. Nezu T, Koga M, Kimura K, et al. **Pretreatment ASPECTS on DWI predicts 3-month outcome following rt-PA: SAMURAI rt-PA Registry.** *Neurology* 2010;75:555–61 CrossRef Medline
5. 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
6. Hill MD, Rowley HA, Adler F, et al; PROACT-II Investigators. **Selection of acute ischemic stroke patients for intra-arterial thrombolysis with pro-urokinase by using ASPECTS.** *Stroke* 2003;34:1925–31 CrossRef Medline
7. Okazaki S, Moriwaki H, Minematsu K, et al. **Extremely early computed tomography signs in hyperacute ischemic stroke as a predictor of parenchymal hematoma.** *Cerebrovasc Dis* 2008;25:241–46 CrossRef Medline
8. Yoo AJ, Berkhemer OA, Fransen PS, et al; MR CLEAN investigators. **Effect of baseline Alberta Stroke Program Early CT Score on safety and efficacy of intra-arterial treatment: a subgroup analysis of a randomised phase 3 trial (MR CLEAN).** *Lancet Neurol* 2016;15: 685–94 CrossRef Medline
9. 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
10. Grotta JC, Chiu D, Lu M, et al. **Agreement and variability in the interpretation of early CT changes in stroke patients qualifying for intravenous rtPA therapy.** *Stroke* 1999;30:1528–33 CrossRef Medline
11. Menon B, Smith E, Modi J, et al. **Regional leptomeningeal score on CT angiography predicts clinical and imaging outcomes in patients with acute anterior circulation occlusions.** *AJNR Am J Neuroradiol* 2011;32:1640–45 CrossRef Medline
12. Menon BK, Puetz V, Kochar P, et al. **ASPECTS and other neuroimaging scores in the triage and prediction of outcome in acute stroke patients.** *Neuroimaging Clin N Am* 2011;21:407–23 CrossRef Medline
13. Bentley P, Ganesalingam J, Carlton Jones AL, et al. **Prediction of stroke thrombolysis outcome using CT brain machine learning.** *Neuroimaging Clin* 2014;4:635–40 CrossRef Medline
14. Przelaskowski A, Sklinda K, Bargieł P, et al. **Improved early stroke detection: wavelet-based perception enhancement of computerized tomography exams.** *Comput Biol Med* 2007;37:524–33 CrossRef Medline
15. Rajini NH, Bhavani R. **Computer aided detection of ischemic stroke using segmentation and texture features.** *Measurement* 2013;46: 1865–74 CrossRef
16. Tang FH, Ng DK, Chow DH. **An image feature approach for computer-aided detection of ischemic stroke.** *Comput Biol Med* 2011;41: 529–36 CrossRef Medline
17. Chawla M, Sharma S, Sivaswamy J, et al. **A method for automatic detection and classification of stroke from brain CT images.** *Conf Proc IEEE Eng Med Biol Soc* 2009;2009:3581–84 Medline
18. Hampton-Till J, Harrison M, Kühn AL, et al. **Automated quantification of stroke damage on brain computed tomography scans: e-ASPECTS.** *EMJ* 2015;3:69–74 https://www.emjreviews.com/neurology/article/automated-quantification-of-stroke-damage-on-brain-computed-tomography-scans-e-aspects/. Accessed August 6, 2015
19. Nagel S, Sinha D, Day D, et al. **e-ASPECTS software is non-inferior to neuroradiologists in applying the ASPECT score to computed**

tomography scans of acute ischemic stroke patients. *Int J Stroke* 2017;12:615–22 CrossRef Medline
20. Herweh C, Ringleb PA, Rauch G, et al. **Performance of e-ASPECTS software in comparison to that of stroke physicians on assessing CT scans of acute ischemic stroke patients.** *Int J Stroke* 2016;11:438–45 CrossRef Medline
21. Stoel BC, Marquering HA, Staring M, et al. **Automated brain computed tomographic densitometry of early ischemic changes in acute stroke.** *J Med Imaging (Bellingham)* 2015;2:014004 CrossRef Medline
22. Shieh Y, Chang CH. **Automated ASPECTS scoring system as a clinical support system for acute stroke care.** In: *Proceedings of 2012 IEEE-Engineering in Medicine and Biology Society International Conference on Biomedical and Health Informatics,* Hong Kong, China. January 5–7, 2012:691–94
23. Shieh Y, Chang CH, Shieh M, et al. **Computer-aided diagnosis of hyperacute stroke with thrombolysis decision support using a contralateral comparative method of CT image analysis.** *J Digit Imaging* 2014;27:392–406 CrossRef Medline
24. Su JL, Chan L, Huang SY. **Development of computer aids ASPECTS system for acute ischemic stroke patient: a preliminary study.** In: *Proceedings of the 2017 2nd International Conference for Innovation in Biomedical Engineering and Life Sciences,* Pantang, Malaysia. December 7, 2017:203–07
25. Inoue M, Olivot JM, Labreuche J, et al. **Impact of diffusion-weighted imaging Alberta stroke program early computed tomography score on the success of endovascular reperfusion therapy.** *Stroke* 2014;45:1992–98 CrossRef Medline
26. Obuchowski NA. **Nonparametric analysis of clustered ROC curve data.** *Biometrics* 1997;53:567–78 Medline
27. Criminisi A, Shotton J. *Decision Forests for Computer Vision and Medical Image Analysis.* London; Springer-Verlag; 2013
28. Maier O, Schröder C, Forkert ND, et al. **Classifiers for ischemic stroke lesion segmentation: a comparison study**. *PLoS One* 2015;10:e0145118 CrossRef Medline
29. Aljabar P, Heckemann RA, Hammers A, et al. **Multi-atlas-based segmentation of brain images: atlas selection and its effect on accuracy.** *Neuroimage* 2009;46:726–38 CrossRef Medline

Radiomics-Based Intracranial Thrombus Features on CT and CTA Predict Recanalization with Intravenous Alteplase in Patients with Acute Ischemic Stroke

W. Qiu, H. Kuang, J. Nair, Z. Assis, M. Najm, C. McDougall, B. McDougall, K. Chung, A.T. Wilson, M. Goyal, M.D. Hill, A.M. Demchuk, and B.K. Menon

ABSTRACT

BACKGROUND AND PURPOSE: Thrombus characteristics identified on non-contrast CT (NCCT) are potentially associated with recanalization with intravenous (IV) alteplase in patients with acute ischemic stroke (AIS). Our aim was to determine the best radiomics-based features of thrombus on NCCT and CT angiography associated with recanalization with IV alteplase in AIS patients and proximal intracranial thrombi.

MATERIALS AND METHODS: With a nested case-control design, 67 patients with ICA/M1 MCA segment thrombus treated with IV alteplase were included in this analysis. Three hundred twenty-six radiomics features were extracted from each thrombus on both NCCT and CTA images. Linear discriminative analysis was applied to select features most strongly associated with early recanalization with IV alteplase. These features were then used to train a linear support vector machine classifier. Ten times 5-fold cross-validation was used to evaluate the accuracy of the trained classifier and the stability of the selected features.

RESULTS: Receiver operating characteristic curves showed that thrombus radiomics features are predictive of early recanalization with IV alteplase. The combination of radiomics features from NCCT, CTA, and radiomics changes is best associated with early recanalization with IV alteplase (area under the curve = 0.85) and was significantly better than any single feature such as thrombus length ($P < .001$), volume ($P < .001$), and permeability as measured by mean attenuation increase ($P < .001$), maximum attenuation in CTA ($P < .001$), maximum attenuation increase ($P < .001$), and assessment of residual flow grade ($P < .001$).

CONCLUSIONS: Thrombus radiomics features derived from NCCT and CTA are more predictive of recanalization with IV alteplase in patients with acute ischemic stroke with proximal occlusion than previously known thrombus imaging features such as length, volume, and permeability.

ABBREVIATIONS: AUC = area under the curve; GLCM = gray-level co-occurrence matrix; LSW = level-spot waves; ROC = receiver operating characteristic

Acute ischemic stroke treatment using intravenous alteplase, a tissue-type plasminogen activator, is primarily focused on dissolving thrombus within the arterial tree. Thrombus characteristics such as length, burden, and permeability identified on multitechnique CT are potentially associated with recanalization with IV alteplase.[1] For example, a shorter length of middle cerebral artery M1 segment thrombus is associated with successful recanalization with IV alteplase treatment.[2,3] It is hypothesized that thrombus length along with collateral status and angioarchitecture within the cerebral arterial tree may be related to lysis with IV alteplase.[4] Patients with short-but-permeable thrombi are more likely to be recanalized with intravenous alteplase than those without.[5] Thrombus permeability, measured by attenuation characteristics on noncontrast CT and CT angiography, is associated with improved functional outcome, smaller final infarct volume, and higher recanalization rates.[6] Nonetheless, the ability of these imaging constructs to reliably predict recanalization with IV alteplase in patients with proximal intracranial thrombi is still limited.

Radiomics is an emerging field that converts imaging data into a high-dimensional feature space using a large number of auto-

Received August 3, 2018; accepted after revision October 21.

From the Departments of Clinical Neurosciences (W.Q., H.K., J.N., Z.A. M.N., C.M., A.T.W., B.M., M.G., M.D.H., A.M.D., B.K.M.), Calgary Stroke Program, Mechanical and Manufacturing Engineering (K.C.), Radiology (M.D.H., A.M.D., M.G., B.K.M.), and Community Health Sciences (M.D.H., B.K.M.), University of Calgary, Calgary, Alberta, Canada; Hotchkiss Brain Institute (M.D.H., A.M.D., M.G., B.K.M.), Calgary, Alberta, Canada; and Department of Radiology (J.N.), McMaster University, Hamilton, Ontario, Canada.

This study was funded through an operating grant from the Canadian Institute of Health Research.

Please address correspondence to Bijoy K. Menon, MD, Room 1079, 10th Floor, Foothills Medical Centre, 1403 29th St NW, Calgary, AB, Canada, T2N 2T9; e-mail: docbijoymenon@gmail.com

Indicates open access to non-subscribers at www.ajnr.org
Indicates article with supplemental on-line appendix and tables.
Indicates article with supplemental on-line photos.

http://dx.doi.org/10.3174/ajnr.A5918

Radiomics features used in the analysis

	Method	No. of Features
Set 1: First-order statistics	Voxel intensity distribution	34
Set 2: Shape and size	Thrombus length, surface, and volume	6
Set 3: Textural features	Gray-level co-occurrence matrix	78
	Gray-level run length matrix	66
	Neighborhood gray-level difference matrix	10
	Law texture	105
	Local binary pattern	27

matically extracted data-characterization algorithms.[7-9] Compared with measuring intracranial thrombus characteristics such as length, volume, and permeability, radiomics features distill multiple-but-subtle variations within the thrombus environment that cannot be appreciated by the human eye. We hypothesized that radiomics-based features of intracranial thrombus on NCCT and CTA are capable of measuring distinct characteristics of thrombus associated with recanalization with IV alteplase better than known thrombus characteristics measured on NCCT and CTA.

MATERIALS AND METHODS

Patient Selection

The primary objective of the study was to identify radiomics-based thrombus characteristics associated with recanalization of thrombus with IV alteplase in patients with intracranial ICA or M1 MCA segment thrombi. Because recanalization of thrombus with IV alteplase within these proximal intracranial arterial segments is rare, we used a case-control study design nested within the precise and rapid assessment of collaterals using multi-phase CTA in the triage of patients with acute ischemic stroke for IV or IA Therapy (PRoVe-IT).[10] We only included patients who had thin-slice NCCT (slice thickness ≤2.5 mm) and CTA (slice thickness ≤0.625 mm) performed on the same scanner. The study included all patients fulfilling study inclusion criteria with ICA/M1 MCA segment thrombi treated with IV alteplase who achieved recanalization as assessed on the first angiographic acquisition in patients undergoing endovascular therapy or on repeat CTA within 4 hours of baseline imaging in patients who did not undergo endovascular therapy. This time point for assessment of recanalization gives us a clinically relevant measure of recanalization with intravenous tPA that is relatively early. Patients were matched for age, sex, and stroke severity with controls (patients who also had ICA/M1 MCA segment thrombi treated with IV alteplase but without successful recanalization). After excluding 22 patients with excessive motion ($n = 3$), poor image quality ($n = 9$), and image coregistration failure ($n = 10$), we identified 30 patients and 37 controls for inclusion in the current study. The study was approved by the Conjoint Health Research Ethics Board at the University of Calgary.

Thrombus Segmentation

CTA images were automatically aligned with NCCT images using rigid registration with 3D Slicer (http://www.slicer.org), a publicly available software.[11] The quality of the registration was checked by visual inspection; manual correction was performed whenever the algorithm registration results were suboptimal. A senior neuroradiology fellow (J.N.) performed manual contouring of thrombi in a slice-by-slice manner from axial views of NCCT images using ITK-SNAP (www.itksnap.org),[12] while viewing the corresponding coregistered CTA images for guidance. This manual segmentation was used for the primary analysis. To analyze the variability introduced by manual segmentation, another neuroradiology fellow (Z.A.) manually segmented the same data. The results based on this segmentation were compared with the primary analysis. Pearson correlation analysis was also performed regarding the thrombus length and volume of 2 manual segmentations.

Radiomics Feature Extraction of Thrombus

Three hundred twenty-six predefined radiomics image features that describe thrombus characteristics were extracted in an automated manner from the above images.[7,13] These features can be divided into 3 sets: 1) first-order statistics, 2) shape and size, and 3) high-order statistical textural information (Table). The first set quantifies thrombus-intensity characteristics using first-order statistics calculated from all voxels within the thrombus region. Set 2 consists of features such as thrombus length, volume, and compactness.[2] Set 3 consists of textural features that can quantify textural heterogeneity within the thrombus.[13] More details on feature definition are in the On-line Appendix. These features were calculated in 3D, thereby taking spatial information compared with the neighboring voxels into account. To obtain a single-feature vector per thrombus, we calculated the mean value of each feature across all the voxels within the thrombus. Radiomic features were extracted from NCCT and CTA images independently. A third group, radiomics change features between NCCT and CTA, was generated additionally by subtracting NCCT features from CTA features. Besides the 3 groups described above, a combination of any 2 or 3 feature groups by concatenation was also used for analysis. All calculations were implemented in Matlab (R2016b; MathWorks, Natick, Massachusetts). For more implementation details about the support vector machine, please refer to the On-line Table 3.

Feature Selection and Classification

Linear discriminative analysis was applied to the 3 groups of features and all possible combinations of the 3 groups of features to rank them in order of importance. Receiver operating characteristic (ROC) analysis was used to compare features in each group and in all the various group combinations generated above; these features were then ranked from the highest-to-lowest importance in each group. The optimum number of features for each group was thereby determined by the feature combination yielding the highest area under the curve (AUC) value using a recursive feature-elimination strategy. To decrease the impact of overfitting, we performed experiments using selected features by 5-fold cross-validation with subjects randomly divided into 5 equally sized groups. Within cross-validation experiments, the support vector machine model for optimizing radiomics features and parameters was built on 4 of the 5 folds and evaluated on the subjects in the

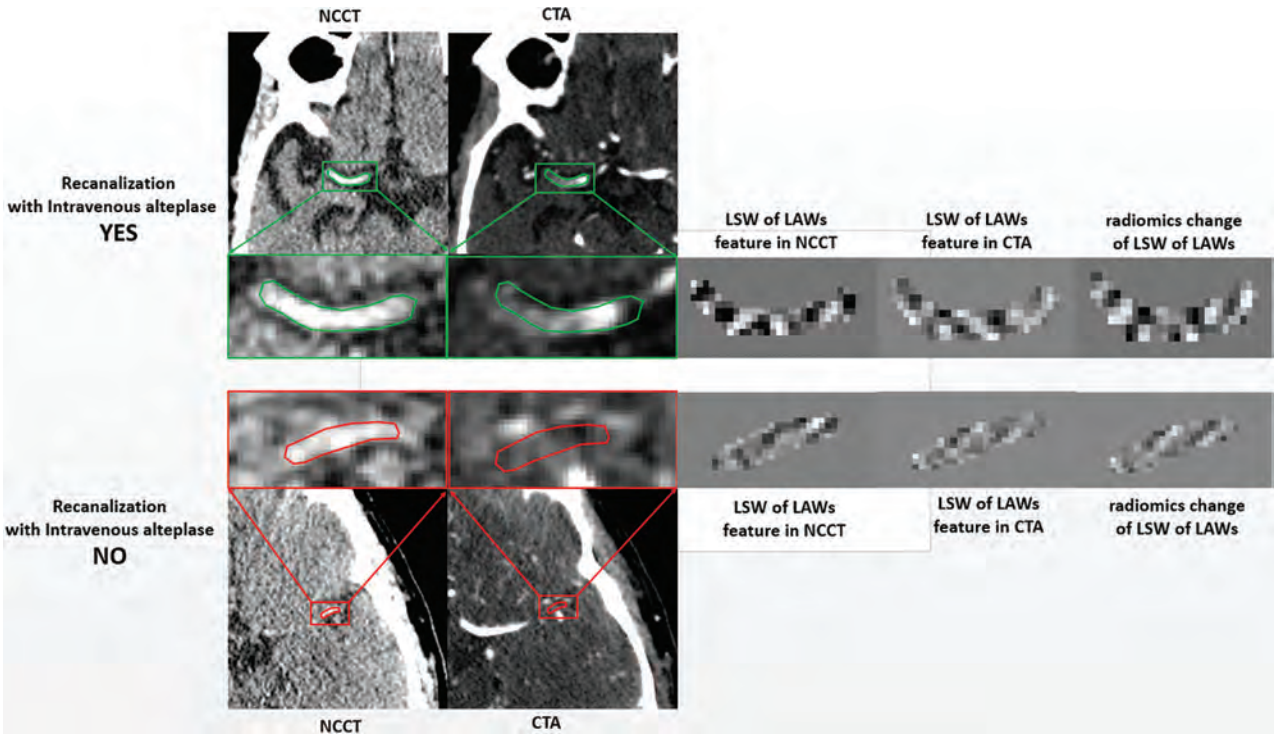


FIG 1. Visualized feature images of thrombus of 2 patients, one with early recanalization and another without recanalization. The selected feature, Laws SD of LSW, is demonstrated. The LSW value (gray-level) of each pixel in the feature map is measuring the heterogeneity in the local region of each pixel.

fifth fold. All cross-validation experiments were repeated 10 times to evaluate the stability of the selected features.

In addition, the selected features were compared with a clinically used feature, residual flow, subjectively graded using consensus by 2 experts from CTA images,[5] with ROC analysis. The selected features were also compared with thrombus permeability defined as a mean attenuation increase from NCCT to CTA, maximum attenuation in CTA, and a maximum attenuation increase from NCCT to CTA. The mean and maximum Hounsfield units of thrombus in NCCT and CTA were manually measured by placing 1–3 ROIs (area of 1–3 mm^2) on the thrombus (1 in each third of the thrombus) on NCCT and CTA.[14,15] However, in this study, the mean and maximum Hounsfield units of thrombus in NCCT and CTA were automatically calculated from the segmented whole thrombus in registered NCCT and CTA images. Finally, we created a multivariable support vector machine model for the clinically used features, such as thrombus length, volume, residual flow, and the 4 permeability measurements, for classification accuracy comparisons.

RESULTS

Patient Characteristics

Baseline demographics and other characteristics of the 67 patients included in this study are summarized in On-line Table 1. The difference regarding baseline clinical variables, etiology of ischemic stroke, and intracranial thrombus imaging characteristics was assessed using the Wilcoxon rank sum test for nonparametric data and the Fisher exact test of proportions for categoric data separately between patients who received intravenous alteplase with and without recanalization. There were no statistically significant differences between the 2 groups by the etiology of ischemic stroke ($P > 0.05$), clinical ($P > 0.05$), and imaging parameters ($P > 0.05$), excluding the occlusion site in internal carotid artery ($P < 0.01$) and proximal M1 segment MCA ($P < 0.01$), and CT imaging ($P < 0.01$) and IV alteplase to recanalization assessment time ($P < 0.01$).

Determining the Best Radiomics Features

The most discriminative features from each feature group associated with recanalization with IV alteplase are shown in On-line Table 2. In particular, the mean of Laws level-spot waves (LSW)[7] ranked high in NCCT (third), radiomics change (second), and the group combining 3 feature groups (third). Figure 1 demonstrates the radiomics feature map of LSW for 2 patients from the case and control groups, one from each group. The feature value (gray-level in this feature map) measures the heterogeneity in the local neighboring region of each pixel. The higher the feature values are, the less heterogeneous is the signal from the thrombus. The 2 thrombi in Fig 1 are different in appearance, as seen by the distribution of gray levels in the feature image. The thrombus that recanalized with intravenous alteplase appears more heterogeneous in terms of Laws LSW feature values compared with the thrombus that did not recanalize.

To determine the radiomics features that best discriminate patients from controls, we used the top 5 features from the NCCT feature group, 15 features from the CTA feature group, 6 features from the radiomics change group, and 12 features from the combination of 3 feature groups listed in On-line Table 2, individually, to generate ROC curves using 10 × 5-fold cross-validation. AUC values for different feature combinations were demon-

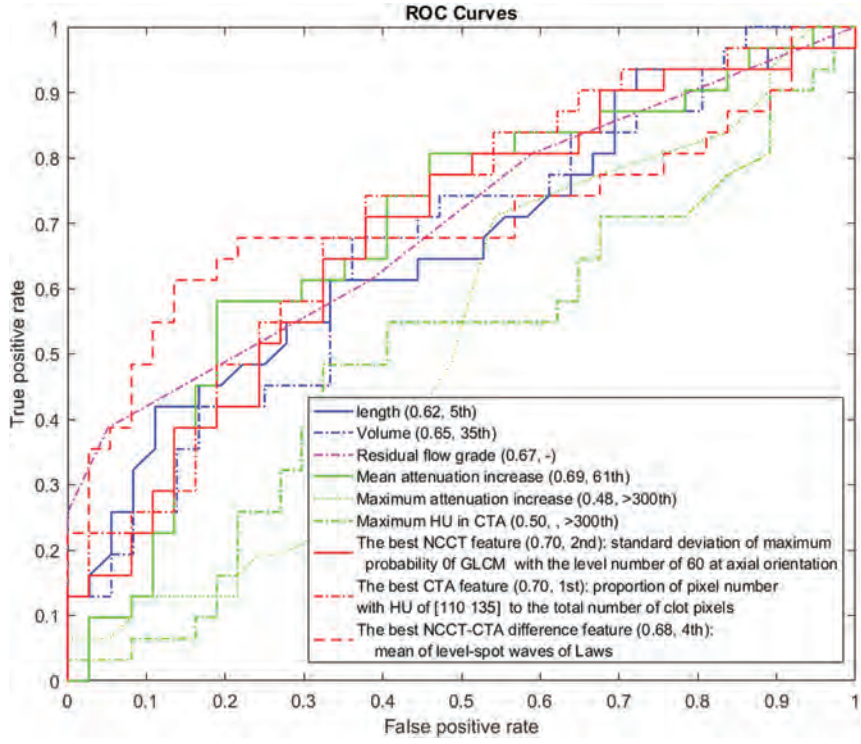


FIG 2. ROC results of the best radiomics features from different feature groups, compared with clinically relevant currently measurable features such as thrombus length, volume, and permeability measurements, as well as subjective assessment of the residual flow grade. The numbers in the parentheses following legend names denote AUC values and ranking after linear discriminative analysis.

strated in On-line Figure 1. The ROC curves of the best features from each group are plotted in Fig 2. The AUCs are 0.69 ± 0.01 for the best NCCT feature (ie, the SD of maximum probability of the gray-level co-occurrence matrix with the level number of 60 at sagittal orientation), 0.69 ± 0.01 for the best CTA feature (the proportion of pixel numbers with 110–135 HU to the total number of clot pixels), and 0.69 ± 0.03 for the best radiomics change (ie, the proportion of pixel numbers with 110–135 HU to the total number of thrombus pixels, respectively). The ROC curves of the top 5 features from each feature group are additionally demonstrated in the On-line Figure 2. An ANOVA test showed that there were no statistically significant differences among these features ($P = .15$). The prediction accuracy of these features is similar to that of biologically relevant features, such as thrombus length[4,16] (AUC = 0.66 ± 0.01, $P = .27$), thrombus volume (AUC = 0.66 ± 0.003, $P = .19$), thrombus permeability measured as[14] mean attenuation increase from NCCT to CTA (AUC = 0.68 ± 0.03, $P = .16$) (Fig 2), but significantly better than the thrombus permeability measurement that uses maximum attenuation in CTA[14] (AUC = 0.49 ± 0.01, $P < .001$) and a maximum attenuation increase from NCCT to CTA (AUC = 0.48 ± 0.07, $P < .001$). Additionally, a nonradiomics feature, residual flow grade, was compared with all the selected radiomics features and generated a similar prediction accuracy (AUC = 0.68 ± 0.03, $P = .21$).

Classification Accuracy

Figure 3 shows ROC curves using different feature combinations with 10 × 5-fold cross validation. In Fig 3*A*, the selected 5 NCCT features are most predictive with an AUC = 0.78 ± 0.02, followed by the 6 radiomics change features with an AUC = 0.77 ± 0.04, while the 15 CTA features are least predictive with AUC = 0.74 ± 0.03. However, combining these features further into different groups increases the prediction accuracy greatly (Fig 3*B*). In particular, 12 features selected from the combination group of NCCT, CTA, and radiomics change features provide the best prediction accuracy (AUC = 0.85 ± 0.03) and were significantly better than features such as thrombus length ($P < .001$), thrombus volume ($P < .001$), thrombus permeability as measured by a mean attenuation increase from NCCT to CTA ($P < .001$) or by maximum attenuation in CTA ($P < .001$) or by a maximum attenuation increase from NCCT to CTA ($P < .001$), and the nonradiomics feature of residual flow grade ($P < .001$). No evidence was found that a combination of the clinically used features improved prediction accuracy (AUC = 0.68 ± 0.03).

Because manual segmentation of thrombus could introduce variability in the analysis, a second manual segmentation of the same data was performed by another neuroradiology fellow (Z.A.). Pearson correlation coefficients regarding the thrombus length and volume of 2 manual segmentations were 0.95 and 0.94, respectively, between the 2 observers. When the analyses were replicated for the second manual segmentations as seen above, the same combination of the NCCT, CTA, and radiomics change features generated an AUC = 0.83 ± 0.02, not significantly different from the results of the first manual segmentation ($P = .27$).

DISCUSSION

Using a nested case-control study design, we show that intracranial thrombus heterogeneity on NCCT and CTA as captured by radiomics is associated with recanalization with intravenous alteplase. Radiomics-based feature analysis of intracranial large-vessel thrombi demonstrably increases our ability to predict recanalization with intravenous alteplase compared with current clinically measurable thrombus characteristics such as length, volume, and thrombus permeability.

Whether IV alteplase will likely result in recanalization in patients with proximal intracranial thrombus is, in our opinion, an important piece of information that can help physicians make appropriate decisions on triage, transport, and treatment of patients, especially when the risk of administering alteplase (intracranial and otherwise) is considered.[17] Patients in whom the probability of recanalization with intravenous alteplase is likely very low and who may have a higher risk of bleeding may benefit from direct transport to the angiosuite for endovascular treatment. Patients in whom the probability of recanalization with intravenous alteplase is very high and in whom endovascular ac-

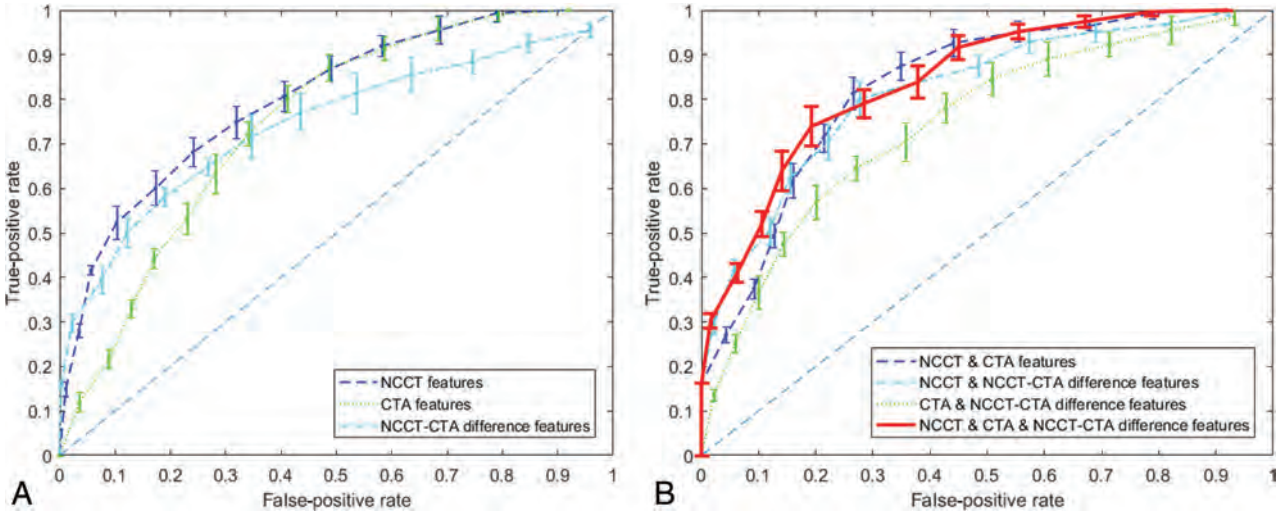


FIG 3. ROC curves for the prediction of early recanalization with IV alteplase using 67 patients (30 patients versus 37 controls). The *bars on the curves* represent the SDs of 10 repetitions. *A*, ROC curves for each feature group. *B*, ROC curves for different combinations of feature groups.

cess is difficult or risky could be given intravenous alteplase and observed for improvement. One may also weigh these considerations when deciding on transport in patients who are imaged at primary stroke centers and await transport to comprehensive stroke centers. We acknowledge that these decisions will not, however, be taken in isolation and that multiple clinical factors will also influence such decisions. Moreover, subgroup analysis of ongoing and future trials testing direct endovascular treatment versus intravenous alteplase followed by endovascular treatment in patients with proximal intracranial thrombus may also benefit to be informed by information that help determine likelihood of early recanalization with lytic agents. In addition, trials that test newer thrombolytic agents and augmented thrombolytic techniques may find this information useful in learning about optimal trial designs.

Radiomics features are computational in a high-dimensional feature space that is converted from imaging data using many data-characterization algorithms. In this study, 326 predefined radiomics features that describe thrombus characteristics, such as first-order statistics, shape and size, and high-order statistical textural feature, were automatically extracted from manually segmented thrombus in NCCT and CTA images. These radiomics features capture subtle variations within the thrombus environment not easily perceived by the human eye. Individually, our analysis shows that these features are as good as known thrombus imaging biomarkers such as thrombus length and permeability.[4,14] However, when these radiomics features are combined into a statistical model, the ability for predicting recanalization with intravenous alteplase is significantly improved ($P < .001$).

Variability in image acquisition across hospitals and scanners in clinical practice affects image analysis and therefore the expert reader's ability to identify thrombus characteristics such as density, length, and permeability. The imaging data used in our analysis were acquired from scanners at multiple centers, and the radiomics features were directly computed from the imaging data, without any preprocessing or normalization. Our analysis demonstrates that the extracted radiomics features are stable under various image-acquisition parameters. Our results therefore suggest the potential generalizability of our results. In our opinion, radiomics-based models predicting recanalization of thrombus with intravenous alteplase can be further improved with better standardization of imaging protocols.

Among 67 patients receiving IV alteplase, the patients with successful recanalization nearly doubled compared with the patients without recanalization, on the basis of a history of diabetes and smoking and anticoagulation therapy (On-line Table 1). Although no statistically significant differences were observed from the limited data in this study, it would be interesting to explore this trend further with a larger dataset. Another finding from On-line Table 1 is that longer times from alteplase start to recanalization and from CT to recanalization are associated with successful recanalization, and the patients with ICA and proximal M1 MCA occlusion have significantly low rates of recanalization (all $P <$ .01), even from these limited data. These results are consistent with the results in Menon et al.[17] With a plasma half-life of 6–7 minutes, alteplase is not likely to be biologically active 6 hours after administration. However, it is possible that the early thrombus debulking effects of alteplase translate to less overall thrombus, allowing endogenous tissue plasminogen activator to complete the remaining lysis required.[17]

This study has some limitations. Our sample size was small, primarily because recanalization of proximal intracranial thrombus with intravenous alteplase within a clinically relevant time is relatively rare. To address this limitation, we used a case-control study design. We do, however, acknowledge that such a design informs us about associations but not causality. To address overfitting of data, we used a series of steps, including the use of i) a feature-selection approach to first select the best features associated with recanalization, ii) a linear support vector machine with fewer parameters to be derived compared with more complicated nonlinear machine learning models such as random forests and deep neural networks, iii) L2 norm regularization to constrain the model fitting and decrease the effective *df* without reducing the actual number of parameters in the model, iv) and a 10× 5-fold cross-validation strategy to generate a more accurate indication of how well the model generalizes to unseen data.

To address variability in thrombus segmentation between experts, a second neuroradiology fellow manually contoured all thrombi using the same processing pipeline. Results were similar to those in the primary analysis. Feature representation influences classification accuracy. In this work, the mean value was used to represent the feature along all voxels within a thrombus,[7,13] without taking feature distribution into account. Feature ranking using linear discriminative analysis was applied on the whole dataset before cross-validation, which might introduce a selection bias,[18] thereby overestimating performance. More extensive validation on an independent test dataset is required to evaluate the stability of the selected features. We were unable to explore the underlying biologic meaning of the radiomics-based imaging. This type of correlation will need in vitro studies that look at radiomics and thrombus pathology together.

CONCLUSIONS

This proof-of-principle case-control study shows that statistical models using radiomics-based imaging features of intracranial large-vessel thrombi can help identify patients who are likely to achieve recanalization with intravenous alteplase. Future work on automating image thrombus segmentation and image analysis along with further validation of our results can potentially lead to the development of computer-aided image-analysis software that could help physicians determine whether patients with proximal occlusions would recanalize in a reasonable timeframe with intravenous alteplase. This knowledge may be useful in the design and testing of novel thrombolytic agents or enhanced thrombolysis strategies. Until then however, this method cannot be used for decision-making in acute stroke care, a time-sensitive situation in which intravenous thrombolysis has to be administered to eligible patients promptly and with minimal delay.

Disclosures: Wu Qiu—*RELATED*: *Grant*: Canadian Institute of Health Research.* Brooklyn McDougall—*UNRELATED*: *Employment*: Quality Improvement and Clinical Research Alberta Stroke Program Summer Studentship; *Grant*: Quality Improvement and Clinical Research https://www.ucalgary.ca/quicr/. Mayank Goyal—*UNRELATED*: *Patents (Planned, Pending or Issued)*: GE Healthcare, *Comments*: licensing agreement for Systems of Stroke diagnosis. Michael D. Hill—*UNRELATED*: *Grants/Grants Pending*: Medtronic, Stryker, Boehringer Ingelheim, Bayer Canada, Alberta Innovates, *Comments*: grants to the University of Calgary for clinical trials*; *Stock/Stock Options*: Calgary Scientific Inc, *Comments*: stock owner in private company. *Money paid to the institution.

REFERENCES

1. Al-Ajlan FS, Qazi E, Kim CK, et al. **Multimodality CT based imaging to determine clot characteristics and recanalization with intravenous tPA in patients with acute ischemic stroke.** *Neurovascular Imaging* 2017;3:2 CrossRef
2. Rohan V, Baxa J, Tupy R, et al. **Length of occlusion predicts recanalization and outcome after intravenous thrombolysis in middle cerebral artery stroke.** *Stroke* 2014;45:2010–17 CrossRef Medline
3. Riedel CH, Zimmermann P, Jensen-Kondering U, et al. **The importance of size: successful recanalization by intravenous thrombolysis in acute anterior stroke depends on thrombus length.** *Stroke* 2011;42:1775–77 CrossRef Medline
4. Qazi EM, Sohn SI, Mishra S, et al. **Thrombus characteristics are related to collaterals and angioarchitecture in acute stroke.** *Can J Neurol Sci* 2015;42:381–88 CrossRef Medline
5. Mishra S, Dykeman J, Sajobi T, et al. **Early reperfusion rates with IV tPA are determined by CTA clot characteristics.** *AJNR Am J Neuroradiol* 2014;35:2265–72 CrossRef Medline
6. Santos EM, Dankbaar JW, Treurniet KM, et al; DUST Investigators. **Permeable thrombi are associated with higher intravenous recombinant tissue-type plasminogen activator treatment success in patients with acute ischemic stroke.** *Stroke* 2016;47:2058–65 CrossRef Medline
7. Aerts HJ, Velazquez ER, Leijenaar RT, et al. **Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach.** *Nat Commun* 2014;5:4006 CrossRef Medline
8. Litjens GJ, Elliott R, Shih NN, et al. **Computer-extracted features can distinguish noncancerous confounding disease from prostatic adenocarcinoma at multiparametric MR imaging.** *Radiology* 2015;278:135–45 CrossRef Medline
9. Gillies RJ, Kinahan PE, Hricak H. **Radiomics: images are more than pictures, they are data.** *Radiology* 2015;278:563–77 CrossRef Medline
10. Menon BK, d'Esterre CD, Qazi EM, et al. **Multiphase CT angiography: a new tool for the imaging triage of patients with acute ischemic stroke.** *Radiology* 2015;275:510–20 CrossRef Medline
11. 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
12. 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
13. Awad J, Krasinski A, Parraga G, et al. **Texture analysis of carotid artery atherosclerosis from three-dimensional ultrasound images.** *Med Phys* 2010;37:1382–91 CrossRef Medline
14. Gensicke H, Pordelli P, Evans J, et al. **Comparison of different methods of thrombus perviousness measurement and impact on recanalization in the INTERRSeCT cohort study.** In: *Proceedings of the 3rd European Stroke Organisation Conference*, Prague, Czech Republic. May 16–18, 2017
15. Santos EM, Marquering HA, den Blanken MD, et al; MR CLEAN Investigators. **Thrombus permeability is associated with improved functional outcome and recanalization in patients with ischemic stroke.** *Stroke* 2016;47:732–41 CrossRef Medline
16. Riedel CH, Jensen U, Rohr A, et al. **Assessment of thrombus in acute middle cerebral artery occlusion using thin-slice nonenhanced computed tomography reconstructions.** *Stroke* 2010;41:1659–64 CrossRef Medline
17. Menon BK, Al-Ajlan FS, Najm M, et al; INTERRSeCT Study Investigators. **Association of clinical, imaging, and thrombus characteristics with recanalization of visible intracranial occlusion in patients with acute ischemic stroke**. *JAMA* 2018;320:1017–26 CrossRef Medline
18. Ambroise C, McLachlan GJ. **Selection bias in gene extraction on the basis of microarray gene-expression data.** *Proc Natl Acad Sci U S A* 2002;99:6562–66 CrossRef Medline

Improved White Matter Cerebrovascular Reactivity after Revascularization in Patients with Steno-Occlusive Disease

L. McKetton, L. Venkatraghavan, C. Rosen, D.M. Mandell, K. Sam, O. Sobczyk, J. Poublanc, E. Gray, A. Crawley, J. Duffin, J.A. Fisher, and D.J. Mikulis

ABSTRACT

BACKGROUND AND PURPOSE: One feature that patients with steno-occlusive cerebrovascular disease have in common is the presence of white matter (WM) lesions on MRI. The purpose of this study was to evaluate the effect of direct surgical revascularization on impaired WM cerebrovascular reactivity in patients with steno-occlusive disease.

MATERIALS AND METHODS: We recruited 35 patients with steno-occlusive disease, Moyamoya disease ($n = 24$), Moyamoya syndrome ($n = 3$), atherosclerosis ($n = 6$), vasculitis ($n = 1$), and idiopathic stenosis ($n = 1$), who underwent unilateral brain revascularization using a direct superficial temporal artery–to-MCA bypass (19 women; mean age, 45.8 ± 16.5 years). WM cerebrovascular reactivity was measured preoperatively and postoperatively using blood oxygen level–dependent (BOLD) MR imaging during iso-oxic hypercapnic changes in end-tidal carbon dioxide and was expressed as %Δ BOLD MR signal intensity per millimeter end-tidal partial pressure of CO_2.

RESULTS: WM cerebrovascular reactivity significantly improved after direct unilateral superficial temporal artery–to-middle cerebral artery (STA-MCA) bypass in the revascularized hemisphere in the MCA territory (mean ± SD, −0.0005 ± 0.053 to 0.053 ± 0.046 %BOLD/mm Hg; $P < .0001$) and in the anterior cerebral artery territory (mean, 0.0015 ± 0.059 to 0.021 ± 0.052 %BOLD/mm Hg; $P = .005$). There was no difference in WM cerebrovascular reactivity in the ipsilateral posterior cerebral artery territory nor in the vascular territories of the nonrevascularized hemisphere ($P < .05$).

CONCLUSIONS: Cerebral revascularization surgery is an effective treatment for reversing preoperative cerebrovascular reactivity deficits in WM. In addition, direct–STA-MCA bypass may prevent recurrence of preoperative symptoms.

ABBREVIATIONS: ACA = anterior cerebral artery; BOLD = blood oxygen level–dependent; CVR = cerebrovascular reactivity; MMD = Moyamoya disease; PCA = posterior cerebral artery; $P_{ET}CO_2$ = end-tidal partial pressure of CO_2; STA = superficial temporal artery

The superficial temporal artery (STA)-middle cerebral artery (MCA) bypass is a surgical revascularization technique often used to bypass upstream in symptomatic steno-occlusive disease. Cerebrovascular reactivity (CVR) is defined as the change in cerebral blood flow (CBF) in response to a vasoactive stimulus, and reflects the ability of the cerebral vasculature to augment CBF when cerebral perfusion pressure is reduced (pressure autoregulation)[1] and when neural activity is increased (neurovascular coupling).[2] Surgical revascularization can improve CVR and reverse the paradoxic reduction in regional blood known as "steal physiology" in the ipsilateral and often in contralateral gray matter (GM) hemispheres,[3,4] reverse cortical thinning,[5] improve neurocognitive function,[6] and reduce further ischemic events.[7]

White matter (WM) may also be adversely affected by steno-occlusive disease. Chronic hypoperfusion can result in ischemic injury, leading to axonal degeneration and glial proliferation.[8] This effect can result in WM lesions or WM hyperintensities on T2-weighted imaging or FLAIR imaging. WM lesions signify a progressive cerebral small-vessel disease that is often observed in patients with Moyamoya disease (MMD), patients with stroke, and in aging. Age-related reductions in CVR and perfusion have been found in WM lesions,[9] and WM abnormalities have been associated with reductions in cognitive function.[10] In addition, there has been a link among impaired WM CVR, cognitive decline

Received July 17, 2018; accepted after revision October 8.

From the Division of Neuroradiology, Joint Department of Medical Imaging (L.M., C.R., D.M.M., K.S., O.S., J.P., E.G., A.C., D.J.M.), Department of Anesthesia and Pain Management (L.V., J.A.F.), University Health Network, Toronto, Ontario, Canada; Russell H. Morgan Department of Radiology and Radiological Science (K.S.), John Hopkins School of Medicine, Baltimore, Maryland; and Department of Physiology (J.D., J.A.F.) and Institute of Medical Sciences (J.D., J.A.F., D.J.M.), University of Toronto, Toronto, Ontario, Canada.

Please address correspondence to David Mikulis, MD, Toronto Western Hospital, Joint Department of Medical Imaging, McLaughlin Pavilion, 3rd Floor, Room 431, 399 Bathurst St, Toronto, ON M5T 2S8, Canada; e-mail: david.mikulis@uhn.ca

Indicates article with supplemental on-line table.

Indicates article with supplemental on-line photo.

http://dx.doi.org/10.3174/ajnr.A5912

in Alzheimer disease,[11] and cognitive dysfunction and hemodynamic impairment due to steno-occlusive disease.[12] A previous study reported that WM lesions became reversible by revascularization surgery in MMD partly due to regional CBF increase.[13] Thus, revascularization surgery may help improve WM CVR that could lead to the conservation of or improvement in cognitive function.

The purpose of this retrospective study was to determine the effects of revascularization on WM CVR in patients with intracranial steno-occlusive disease. We identified 35 patients who fit our selection criteria. We measured WM CVR before and after surgical revascularization in the intervened and nonintervened hemispheres to investigate whether regaining CBF would improve WM CVR.

MATERIALS AND METHODS

Patients

The study was approved by the research ethics board of the University Health Network, and all patients signed informed consent. The patients included in this study were selected from a prospectively maintained data base of patients who underwent CVR blood oxygen level–dependent (BOLD) MR imaging testing as part of a series of research ethics board–approved studies at Toronto Western Hospital. The inclusion criteria for this study were the following: 1) patients with intracranial steno-occlusive disease who had a unilateral direct extracranial-intracranial bypass of the superficial temporal artery to MCA (STA-MCA); 2) the presence of steal physiology on the prerevascularization CVR map; 3) an available preoperative CVR study and at least 1 postoperative CVR study at 3 months or later postsurgery; 4) both preoperative and postoperative CVR studies performed with the same CO_2 stimulus protocol and acquisition parameters; and 5) all patient CVR data being sufficiently free of artifacts to be comparable, as determined by a neuroradiologist experienced in BOLD MR imaging CVR assessment (D.J.M.). Thirty-five patients were selected on the basis of inclusion criteria in this study (19 women; mean age at time of bypass, 45.8 ± 16.5 years). Pathology included MMD ($n = 24$), Moyamoya syndrome ($n = 3$), atherosclerosis ($n = 6$), vasculitis (right ICA stenosis) ($n = 1$), and idiopathic stenosis ($n = 1$). Further details are shown in the On-line Table.

Imaging

MR imaging was performed on a 3T system (Signa HDx platform; GE Healthcare, Milwaukee, Wisconsin) using an 8-channel phased array head coil. For each patient, an anatomic T1-weighted 3D spoiled gradient-echo sequence was acquired with the following parameters that varied from before to after 2009: voxel size = 0.781 × 0.781 × 2.2 mm/0.85 × 0.85 × 1 mm; TR = 7.88 ms; TE = 3 ms; matrix size = 256 × 256; 60/146 slices; FOV = 22 × 22 cm; flip angle = 12°. Each patient had the same pre- and postoperative BOLD MR imaging parameters acquired. As end-tidal partial pressure of CO_2 ($P_{ET}CO_2$) was manipulated, BOLD MR imaging data were acquired using a T2*-weighted single-shot spiral gradient-echo sequence with an echo-planar readout. Ten out of 35 patients were scanned before 2010 and had the same acquisition protocol for both pre- and postbypass CVRs, and 25 patients were scanned after 2010 and had a different acquisition protocol for their pre- and postbypass CVR BOLD scans. The only differences between protocols included the reduction of a 2-mm gap, lowering the flip angle by 5°, and increasing the slices from 20 to 30. Therefore, the BOLD parameters that varied from before 2010 to after 2010 were the following: voxel size = 3.75 × 3.75 × 5 mm, 2-mm gap/no gap, TR = 2 seconds, TE = 30 ms, matrix size = 64 × 64, 20/30 slices, FOV = 24 × 24 cm, flip angle = 90°/85°.

Vasodilatory Stimulus

The control of $P_{ET}CO_2$ and the end-tidal pressure of O_2 were achieved during MR imaging using an automated gas blender that regulates gas composition and flow to a sequential gas-delivery breathing circuit (RespirAct™; Thornhill Medical, Toronto, Ontario, Canada) according to the methods previously described.[14] The automated gas blender was connected to a soft plastic mask sealed to the face using transparent dressing film (Tegaderm Film, 1626W; 3M Health Care, St. Paul, Minnesota). Before and after 2013, the patients underwent either a 2-square wave (step) protocol or 1-square wave followed by a ramp (step and a ramp) protocol for manipulating $P_{ET}CO_2$ (from resting partial pressure of CO_2 to 10 mm Hg above resting) as previously detailed.[3,15] Both $P_{ET}CO_2$ and end-tidal pressure of O_2 values were selected from the raw partial pressure of carbon dioxide and partial pressure of O_2 tracings and were confirmed post hoc by visual inspection and corrected if needed by a custom-written program (LabVIEW; National Instruments, Austin, Texas). Of the 35 patients studied, 2 patients (P14 and P21) had their first CVR fail due to excessive motion and had their CVR scan restarted with no motion artifacts. This scan was used for analysis. One patient (P11) had the initial postoperative CVR scan fail due to technical issues with the scanner freezing, causing patient distress with the emergency button pressed. The patient returned for their follow-up the following month.

Data and Statistical Analyses

MR imaging and $P_{ET}CO_2$ data were imported to an independent workstation and preprocessed using AFNI software (http://afni.nimh.nih.gov/afni),[16] SPM8 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12), and Matlab R2015a (MathWorks, Natick, Massachusetts). $P_{ET}CO_2$ data were time-shifted to the point of maximum correlation between the rapid changes in $P_{ET}CO_2$ and BOLD signal. CVR was calculated as the slope of the line of best fit between the percentage change in BOLD signal versus $P_{ET}CO_2$. CVR was calculated on a voxel-by-voxel basis from the slope of a linear least-squares fit of the BOLD signal data series to the $P_{ET}CO_2$ values and was expressed as the percentage change in BOLD signal per change in $P_{ET}CO_2$ (%/mm Hg). The magnitude of the slope in each voxel was color-coded from a color spectrum in which yellow to red denoted a small-to-large positive slope relating to an increase in CBF, and light to dark blue denoted a small-to-large negative slope relating to reversal of CBF (ie, paradoxic vascular steal physiology) (Fig 1).

Anatomic T1-weighted images were segmented into cerebrospinal fluid (CSF), GM and WM. CVR masks were generated containing only WM, and were subsequently transformed into Montreal Neurological Institute space. Unihemispheric WM

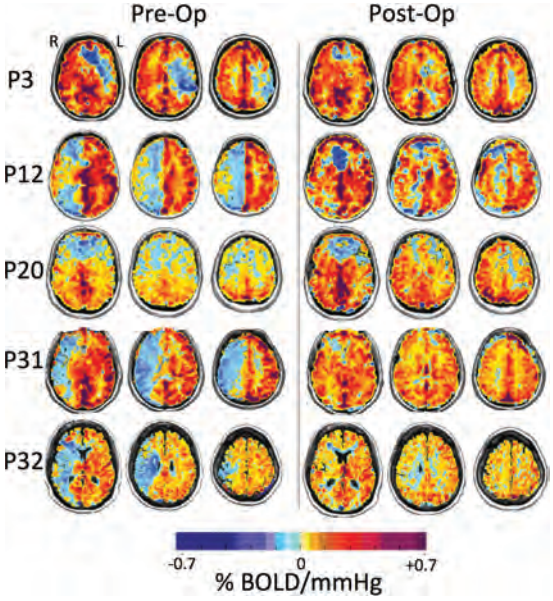


FIG 1. BOLD MR imaging CVR maps in standard space overlaid on an anatomic T1-weighted image of 5 representative patients (P3, P12, P20, P31, and P32) who underwent a left (P3) and right (P12, P20, P31, and P32) STA-MCA bypass. The first column shows CVR maps before bypass, and the second column shows CVR maps within a 1-year follow-up after revascularization surgery. CVR units are %ΔBOLD MR signal intensity per millimeter mercury $\Delta P_{ET}CO_2$. Areas shown in blue demonstrate decreased and paradoxic (negative) CVR as shown in P3 in the left hemisphere and P12, P20, P31, and P32 in the right hemisphere (left column). The resulting CVR maps visually demonstrate improvement (reduction of steal) postbypass (right column).

probability maps were thresholded at 70% probability and were used as a template for calculating CVR in each hemisphere. These were transformed into Montreal Neurological Institute space using SPM8. Unihemispheric statistical parametric mapping GM probability maps were thresholded at 70% in AFNI and served as a template for calculating hemispheric CVR for each participant. Each major arterial vascular territory comprising the MCA, posterior cerebral artery (PCA), and anterior cerebral artery (ACA) was identified by 2 neuroradiologists (D.J.M. and D.M.M.) and manually traced, creating ROIs. A Wilcoxon matched-pairs signed rank test (2-sided, $\alpha = .05$) was used for statistical analysis comparing the prerevascularization and postrevascularization WM CVR values (Fig 2). Results were considered significant after accounting for multiple comparisons by a Bonferroni correction if the per-comparison P value $<.05/(3$ comparisons) was .017. Additional analyses on GM CVR values pre- and postrevascularization are reported in the On-line Figure.

RESULTS

The mean duration between STA-MCA bypass surgery and the postoperative CVR study was 8.3 months (median, 5 months). Each patient's STA-MCA bypass was patent. There was no correlation between follow-up time and WM CVR changes in the revascularized MCA ($r = -0.14$, $P = .42$), PCA ($r = -0.12$, $P = .47$), and ACA ($r = 0.012$, $P = .94$) territory. The mean time between preoperative and postoperative CVR studies was 12.6 months (median, 8 months) (On-line Table).

WM CVR significantly improved after direct unilateral STA-MCA bypass in the revascularized hemisphere in the MCA territory (mean ± SD, −0.0005 ± 0.053 to 0.053 ± 0.046 %BOLD/mm Hg; $P < .0001$) and in the ACA territory (0.0015 ± 0.059 to 0.021 ± 0.052 %BOLD/mm Hg; $P = .0053$). There was no difference in WM CVR after bypass in the ipsilateral PCA territory (0.128 ± 0.059 to 0.13± 0.062 %BOLD/mm Hg; $P = .98$) and in the nonrevascularized MCA (mean ± SD, 0.091 ± 0.057 to 0.092 ± 0.055 %BOLD/mm Hg; $P = .74$), ACA (0.041 ± 0.052 to 0.043 ± 0.055 %BOLD/mm Hg; $P = .87$), or PCA (0.149 ± 0.051 to 0.145 ± 0.054 %BOLD/mm Hg; $P = .62$) territory (Fig 2).

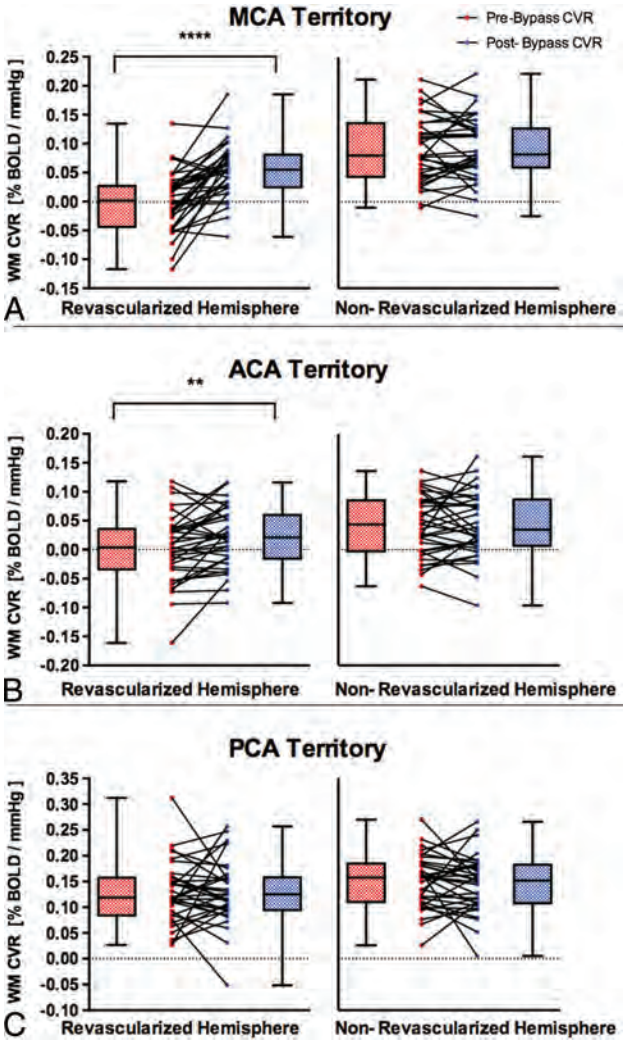


FIG 2. Comparisons of pre- and post-unilateral STA-MCA bypass on white matter CVR in the ipsilateral revascularized and contralateral nonrevascularized hemispheres in patients with MMD, Moyamoya syndrome, and steno-occlusive and atherosclerotic disease. The white matter CVR in revascularized and nonrevascularized hemispheres is shown in different vascular territories: MCA (*A*), ACA (*B*), and PCA (*C*). The *box* denotes the interquartile range (25%–75%), the *horizontal line* in the *box* denotes the median, and the *whiskers* denote the minimum and maximum values. *Double asterisks* indicate $P < .01$; *4 asterisks*, $P < .0001$.

GM CVR significantly improved after direct unilateral STA-MCA bypass in the revascularized hemisphere in the MCA territory (mean ± SD, 0.034 ± 0.07 to 0.11 ± 0.059 %BOLD/mm Hg; $P < .0001$). There was no difference in GM CVR after bypass in

the ipsilateral ACA territory (0.08 ± 0.10 to 0.09 ± 0.08 %BOLD/mm Hg; $P = .21$), PCA territory (0.25 ± 0.1 to 0.23 ± 0.1 %BOLD/mm Hg; $P = .23$), or in the nonrevascularized MCA (0.17 ± 0.1 to 0.16 ± 0.08 %BOLD/mm Hg; $P = .1$), ACA (0.26 ± 0.09 to 0.25 ± 0.09 %BOLD/mm Hg; $P = .12$), or PCA (0.25 ± 0.08 to 0.25 ± 0.1 %BOLD/mm Hg; $P = .09$) territory (On-line Figure).

DISCUSSION

This study shows that the ipsilateral WM CVR improved after unilateral revascularization of the affected hemisphere with a >10 time increase in the MCA and ACA territories. In addition, there were some negative changes postbypass in the nonrevascularized hemisphere. For example, of 35 patients, 18 patients had a negative change ranging from a 0.18%–10.1% decrease in CVR postbypass in the nonrevascularized MCA territory (Fig 2*A*). This could be due to the disease progression affecting the contralateral side and because the WM CVR signal is noisier (ie, the range of 0%–10% pre- and postoperative differences cannot be detected with confidence). Previous studies have shown that revascularization surgery improves cortical GM CVR and reduces the risk of acute ischemia.[3,4] The importance of our findings suggests that WM CVR impairment, that is related to WM lesions and associated with chronic ischemia, is also improved with revascularization.

Previous studies reported WM CVR characteristics in healthy controls. Sam et al,[3] reported the mean WM CVR to be 0.17% ± 0.06 %/mm Hg in 27 healthy control subjects (age range, 19–71 years). Thomas et al[17] found that WM CVR was about 4–8 times lower than GM CVR. There is a paucity of literature pertaining to WM CVR postrevascularization.

Our results are in agreement with those of a previous study that found significant improvement in WM CVR in the intervened MCA hemisphere but not in the nonintervened hemisphere after revascularization surgery.[3] Additionally, Sam et al,[3] showed that WM CVR improved bilaterally in both ACA and PCA territories in 20 patients with steno-occlusive disease, whereas our findings showed improvement only in the ipsilateral MCA and ACA territories after revascularization. We suggest that these discrepancies may have resulted from sample size differences, heterogeneity in the pathologies of patients, and variability in revascularization techniques in the form of extracranial-intracranial bypass, carotid endarterectomy, and encephaloduroarteriosynangiosis.

Patients with steno-occlusive disease have an increased risk of cerebral hypoperfusion that could consequently lead to brain atrophy, cognitive impairment, and dementia.[12] The most frequent presentation of intracranial steno-occlusive diseases such as MMD is transient ischemic attack or ischemic stroke.[18] In MMD, the risk of recurrent stroke is as high as 10% per year.[19] Steal physiology in MMD and other steno-occlusive diseases is associated with a high risk of stroke.[20] Furthermore, impaired CVR with steal physiology in steno-occlusive disease has been associated with cognitive dysfunction,[12] ipsilateral cortical thinning,[21] and increased diffusion in WM[22]; the latter is likely related to increased water diffusivity due to axonal degeneration.[23] In patients with atherosclerotic disease, CVR deficits are associated with increased risk of future ischemic attacks.[24] However the clinical efficacy of revascularization surgery in decreasing this risk remains controversial.[25]

As a potential benefit to neurocognition, extracranial-intracranial bypass has previously been reported to significantly improve cognitive function in steno-occlusive disease[26,27]; our study suggests that WM reperfusion may contribute to the improvement. Additionally, revascularization surgery has been shown to improve cognitive function in patients with symptomatic primary carotid stenosis undergoing carotid endarterectomy and carotid artery stent placement[28] and in patients with ICA stenosis undergoing carotid endarterectomy.[29] The caveat is that these reported results were variable and were part of nonrandomized studies. However, the Randomized Evaluation of Carotid Occlusion and Neurocognition trial reported no significant difference in 2-year cognitive changes between 16 patients who underwent current best medical therapy and 13 patients who underwent extracranial-intracranial bypass.[30] In both arms, patients had symptomatic ICA occlusion and increased oxygen extraction fraction on PET. Further studies are needed to assess CVR in patients who underwent extracranial-intracranial bypass compared with patients who underwent medical therapy (ie, consistent use of statins and antihypertensives).

The mechanism of restoration of WM CVR requires further investigation, but the phenomenon has been shown previously. Patients with steno-occlusive disease who had undergone extracranial-intracranial arterial bypass surgery had an increase in their total brain blood supply, particularly the restoration of perfusion in hemodynamically compromised brain tissue.[31] Revascularization surgery was also previously found to improve cerebral perfusion in MMD in the ipsilateral MCA territory in children[32] and adults.[33]

The assessment of cerebral hemodynamic insufficiency in steno-occlusive disease has remained a complex issue due to the lack of suitable methods and consistency between study centers.[34] Various imaging methods are used to evaluate cerebral hemodynamics, including CVR or oxygen extraction fraction using PET,[35] xenon-enhanced CT[36], perfusion-weighted MR imaging,[37] transcranial Doppler,[34] and SPECT.[38] BOLD CVR MR imaging and oxygen extraction fraction using PET are the most commonly used mapping methods. However, PET is costly, not widely available, more invasive, uses ionizing radiation,[39] and provides no indication of dynamic autoregulation.[40] Standardized BOLD MR imaging CVR using CO_2 as a vasodilatory stimulus is noninvasive, requires no exposure to ionizing radiation (important for use in the pediatric population) or contrast agents, and is a more available method for mapping whole-brain CVR. In addition, a recent study in patients with advanced steno-occlusive disease showed that BOLD MR imaging CVR performed very well against PET CVR using identical CO_2 stimuli.[41] BOLD CVR MR imaging is therefore an accessible and readily applied method for routine clinical use.

CVR can be measured using BOLD MR imaging as a surrogate for cerebral blood flow and hypercapnic changes in the $P_{ET}CO_2$ as the vasoactive stimulus.[42] In healthy individuals, there is a strong positive relationship between $P_{ET}CO_2$ and CBF over a large range of $P_{ET}CO_2$. In steno-occlusive disease, patients may exhibit

downstream regional decreases in CBF in response to hypercapnia. With mild hemodynamic impairment, reductions in CVR result from a diminished increase in blood flow, whereas in more severe cases, steal physiology occurs. Vascular steal arises from the redistribution of blood flow away from any region that has reduced vascular reserve to vascular beds with intact reserve, which results in a greater reduction in flow resistance.[15] In addition to vascular steal, other mechanisms including cerebrovascular autoregulation and/or the cerebral metabolic rate of O_2 up-regulation may contribute to negative BOLD CVR.[43] Here, an alternative explanation for a negative BOLD signal could be due to an increase in blood volume, particularly if the CBV changes happen in arteries or veins that have low oxygen saturation fractions as found in patients with hypoxia with no change in deoxyhemoglobin concentration in the blood. Additionally, Arteaga et al[43] reported negative BOLD in WM with low CBV (0.01–0.02 mL blood/mL parenchyma), which would need very sizeable changes in CBV to explain the magnitude of the CVR BOLD signal decrease. Although their stimulus included hyperoxia (and ours did not), it was proposed that the cerebral metabolic rate of O_2 may increase during the hypercapnic and hyperoxia stimulus because hypoxic tissue at baseline may metabolize the additional oxygen provided.[43] Partial voluming with CSF (because CSF has no CBF) in periventricular tissue may have contributed to paradoxic reduced flow; however, because the patients have larger regions of steal, we can rule this out.

There are a number of approaches used to implement a vasoactive stimulus (eg, breath-holding, administering a constant level of CO_2, injecting acetazolamide) that can result in large variations in the stimuli and thereby large variations in measured CVR.[44] Our study used BOLD MR imaging with precisely repeatable CO_2, which minimizes sources of variability and allows detailed mapping of the evolution of CVR changes in patients across time, accounting for voxelwise test-retest variability compared against a control CVR atlas.[9] This key feature is unique to this study because accurate assessments in CVR changes are challenging to separate from test-retest variability with nonrepeatable and unknown stimuli.

The present study did have limitations. The patient population under study included heterogeneous vascular pathology; nevertheless, each patient underwent standardized unilateral direct STA-MCA bypass surgery. We allowed liberal inclusion criteria with respect to age, medication, and disease process to maintain the generality of the results.

CONCLUSIONS

Unilateral extracranial-intracranial bypass improves WM CVR in the ipsilateral MCA and ACA territories, indicating improvement in underlying hemodynamic reserve. Prospective studies using quantitative, reproducible vasodilatory stimuli examining post-revascularization cognitive performance and other clinical parameters, including patients who have undergone common medical therapy, are warranted and strongly recommended to shed further light on the efficacy of bypass revascularization.

Disclosures: J.A.F. and J.D. are senior scientists at Thornhill Research, Inc., (TRI), a spin-off company affiliated with the University Health Network (UHN) that developed the RespirAct™, a non-commercial research tool made available by TMI to research institutions to enable CVR studies. D.J.M. is a shareholder in TRI. O.S. is a part time employee of TRI. The remaining authors have no disclosures and declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

1. Fantini S, Sassaroli A, Tgavalekos KT, et al. **Cerebral blood flow and autoregulation: current measurement techniques and prospects for noninvasive optical methods.** *Neurophotonics* 2016;3:31411 CrossRef Medline
2. Pillai JJ, Mikulis DJ. **Cerebrovascular reactivity mapping: an evolving standard for clinical functional imaging.** *AJNR Am J Neuroradiol* 2015;36:7–13 CrossRef Medline
3. Sam K, Poublanc J, Sobczyk O, et al. **Assessing the effect of unilateral cerebral revascularisation on the vascular reactivity of the non-intervened hemisphere: a retrospective observational study.** *BMJ Open* 2015;5:e006014 CrossRef Medline
4. Mandell DM, Han JS, Poublanc J, et al. **Quantitative measurement of cerebrovascular reactivity by blood oxygen level-dependent MR imaging in patients with intracranial stenosis: preoperative cerebrovascular reactivity predicts the effect of extracranial-intracranial bypass surgery.** *AJNR Am J Neuroradiol* 2011;32:721–27 CrossRef Medline
5. Fierstra J, Maclean DB, Fisher JA, et al. **Surgical revascularization reverses cerebral cortical thinning in patients with severe cerebrovascular steno-occlusive disease.** *Stroke* 2011;42:1631–37 CrossRef Medline
6. Baek HJ, Chung SY, Park MS, et al. **Preliminary study of neurocognitive dysfunction in adult moyamoya disease and improvement after superficial temporal artery-middle cerebral artery bypass.** *J Korean Neurosurg Soc* 2014;56:188–93 CrossRef Medline
7. Mesiwala AH, Sviri G, Fatemi N, et al. **Long-term outcome of superficial temporal artery–middle cerebral artery bypass for patients with Moyamoya disease in the US.** *Neurosurg Focus* 2008;24:E15 CrossRef Medline
8. Pantoni L, Garcia JH. **The significance of cerebral white matter abnormalities 100 years after Binswanger's report: a review.** *Stroke* 1995;26:1293–301 CrossRef Medline
9. Sobczyk O, Crawley AP, Poublanc J, et al. **Identifying significant changes in cerebrovascular reactivity to carbon dioxide.** *AJNR Am J Neuroradiol* 2016;37:818–24 CrossRef Medline
10. Deary IJ, Leaper SA, Murray AD, et al. **Cerebral white matter abnormalities and lifetime cognitive change: a 67-year follow-up of the Scottish Mental Survey of 1932.** *Psychol Aging* 2003;18:140–48 CrossRef Medline
11. Silvestrini M, Pasqualetti P, Baruffaldi R, et al. **Cerebrovascular reactivity and cognitive decline in patients with Alzheimer disease.** *Stroke* 2006;37:1010–15 CrossRef Medline
12. Balucani C, Viticchi G, Falsetti L, et al. **Cerebral hemodynamics and cognitive performance in bilateral asymptomatic carotid stenosis.** *Neurology* 2012;79:1788–95 CrossRef Medline
13. Komatsu K, Mikami T, Noshiro S, et al. **Reversibility of white matter hyperintensity by revascularization surgery in Moyamoya disease.** *J Stroke Cerebrovasc Dis* 2016;25:1495–502 CrossRef Medline
14. Slessarev M, Han J, Mardimae A, et al. **Prospective targeting and control of end-tidal CO2 and O2 concentrations.** *J Physiol* 2007;581:1207–19 CrossRef Medline
15. Sobczyk O, Battisti-Charbonney A, Fierstra J, et al. **A conceptual model for CO2-induced redistribution of cerebral blood flow with experimental confirmation using BOLD MRI.** *Neuroimage* 2014;92:56–68 CrossRef Medline
16. Cox RW. **AFNI: software for analysis and visualization of functional magnetic resonance neuroimages.** *Comput Biomed Res* 1996;29:162–73 CrossRef Medline
17. Thomas BP, Liu P, Park DC, et al. **Cerebrovascular reactivity in the brain white matter: magnitude, temporal characteristics, and age effects.** *J Cereb Blood Flow Metab* 2014;34:242–47 CrossRef Medline
18. Zhao WG, Luo Q, Jia JB, et al. **Cerebral hyperperfusion syndrome**

after revascularization surgery in patients with moyamoya disease. *Br J Neurosurg* 2013;27:321–25 CrossRef Medline
19. Chiu D, Shedden P, Bratina P, et al. **Clinical features of moyamoya disease in the United States.** *Stroke* 1998;29:1347–51 CrossRef Medline
20. Schoof J, Lubahn W, Baeumer M, et al. **Impaired cerebral autoregulation distal to carotid stenosis/occlusion is associated with increased risk of stroke at cardiac surgery with cardiopulmonary bypass.** *J Thorac Cardiovasc Surg* 2007;134:690–96 CrossRef Medline
21. Fierstra J, Poublanc J, Han JS, et al. **Steal physiology is spatially associated with cortical thinning.** *J Neurol Neurosurg Psychiatry* 2010; 81:290–93 CrossRef Medline
22. Conklin J, Fierstra J, Crawley AP, et al. **Impaired cerebrovascular reactivity with steal phenomenon is associated with increased diffusion in white matter of patients with Moyamoya disease.** *Stroke* 2010;41:1610–16 CrossRef Medline
23. Helenius J, Soinne L, Salonen O, et al. **Leukoaraiosis, ischemic stroke, and normal white matter on diffusion-weighted MRI.** *Stroke* 2002;33:45–50 CrossRef Medline
24. Silvestrini M, Vernieri F, Pasqualetti P, et al. **Impaired cerebral vasoreactivity and risk of stroke in patients with asymptomatic carotid artery stenosis.** *JAMA* 2000;283:2122–27 CrossRef Medline
25. Grubb RL, Powers WJ, Derdeyn CP, et al. **The Carotid Occlusion Surgery Study.** *Neurosurg Focus* 2003;14:e9
26. Dong Y, Teoh HL, Chan BP, et al. **Changes in cerebral hemodynamic and cognitive parameters after external carotid–internal carotid bypass surgery in patients with severe steno-occlusive disease: a pilot study.** *J Neurol Sci* 2012;322:112–16 CrossRef Medline
27. Nielsen H, Højer-Pedersen E, Gulliksen G, et al. **Reversible ischemic neurological deficit and minor strokes before and after EC/IC bypass surgery: a neuropsychological study.** *Acta Neurol Scand* 1986; 73:615–18 CrossRef Medline
28. Lal BK, Younes M, Cruz G, et al. **Cognitive changes after surgery vs stenting for carotid artery stenosis.** *J Vasc Surg* 2011;54:691–98 CrossRef Medline
29. Yoshida K, Ogasawara K, Kobayashi M, et al. **Improvement and impairment in cognitive function after carotid endarterectomy: comparison of objective and subjective assessments.** *Neurol Med Chir (Tokyo)* 2012;52:154–60 CrossRef Medline
30. Marshall RS, Festa JR, Cheung YK, et al; RECON Investigators. **Randomized Evaluation of Carotid Occlusion and Neurocognition (RECON) trial: main results.** *Neurology* 2014;82:744–51 CrossRef Medline
31. Neff KW, Horn P, Dinter D, et al. **Extracranial/intracranial arterial bypass surgery improves total brain blood supply in selected symptomatic patients with unilateral internal carotid artery occlusion and insufficient collateralization.** *Neuroradiology* 2004;46:730–37 CrossRef Medline
32. Blauwblomme T, Lemaitre H, Naggara O, et al. **Cerebral blood flow improvement after indirect revascularization for pediatric Moyamoya disease: a statistical analysis of arterial spin-labeling MRI.** *AJNR Am J Neuroradiol* 2016;37:706–12 CrossRef Medline
33. Kwon WK, Kwon TH, Park DH, et al. **Efficacy of superficial temporal artery-middle cerebral artery bypass in cerebrovascular steno-occlusive diseases: hemodynamics assessed by perfusion computed tomography.** *Asian J Neurosurg* 2017;12:519–24 CrossRef Medline
34. Reinhard M, Schwarzer G, Briel M, et al. **Cerebrovascular reactivity predicts stroke in high-grade carotid artery disease.** *Neurology* 2014; 83:1424–31 CrossRef Medline
35. Endo H, Inoue T, Ogasawara K, et al. **Quantitative assessment of cerebral hemodynamics using perfusion-weighted MRI in patients with major cerebral artery occlusive disease: comparison with positron emission tomography.** *Stroke* 2006;37:388–92 CrossRef Medline
36. Meyer JS, Shinohara T, Imai A, et al. **Imaging local cerebral blood flow by xenon-enhanced computed tomography? Technical optimization procedures.** *Neuroradiology* 1988;30:283–92 CrossRef Medline
37. Zakariaee S, Oghabian M, Firouznia K, et al. **Assessment of the agreement between cerebral hemodynamic indices quantified using dynamic susceptibility contrast and dynamic contrast-enhanced perfusion magnetic resonance imaging.** *J Clin Imaging Sci* 2018;8:2 CrossRef Medline
38. Cikrit DF, Dalsing MC, Harting PS, et al. **Cerebral vascular reactivity assessed with acetazolamide single photon emission computer tomography scans before and after carotid endarterectomy.** *Am J Surg* 1997;174:193–97 CrossRef Medline
39. Ragan DK, McKinstry R, Benzinger T, et al. **Depression of whole-brain oxygen extraction fraction is associated with poor outcome in pediatric traumatic brain injury.** *Pediatr Res* 2012;71: 199–204 CrossRef Medline
40. Derdeyn CP, Videen TO, Yundt KD, et al. **Variability of cerebral blood volume and oxygen extraction: stages of cerebral haemodynamic impairment revisited.** *Brain* 2002;125:595–607 CrossRef Medline
41. Fierstra J, van Niftrik C, Warnock G, et al. **Staging hemodynamic failure with blood oxygen-level-dependent functional magnetic resonance imaging cerebrovascular reactivity: a comparison versus gold standard (15O-)H2O-positron emission tomography.** *Stroke* 2018;49:621–29 CrossRef Medline
42. Mandell DM, Han JS, Poublanc J, et al. **Mapping cerebrovascular reactivity using blood oxygen level-dependent MRI in patients with arterial steno-occlusive disease: comparison with arterial spin labeling MRI.** *Stroke* 2008;39:2021–28 CrossRef Medline
43. Arteaga DF, Strother MK, Faraco CC, et al. **The vascular steal phenomenon is an incomplete contributor to negative cerebrovascular reactivity in patients with symptomatic intracranial stenosis.** *J Cereb Blood Flow Metab* 2014;34:1453–62 CrossRef Medline
44. Powers WJ, Clarke WR, Grubb RL, et al. **Extracranial-intracranial bypass surgery for stroke prevention in hemodynamic cerebral ischemia: the Carotid Occlusion Surgery Study randomized trial.** *JAMA* 2011;306:1983 CrossRef Medline

Infarct Volume Predicts Hospitalization Costs in Anterior Circulation Large-Vessel Occlusion Stroke

C.D. Streib, S. Rangaraju, D.T. Campbell, D.G. Winger, S.L. Paolini, A.J. Zhang, B.T. Jankowitz, A.P. Jadhav, and T.G. Jovin

ABSTRACT

BACKGROUND AND PURPOSE: Anterior circulation large-vessel occlusion stroke, one of the most devastating stroke subtypes, is associated with substantial economic burden. We aimed to identify predictors of increased acute care hospitalization costs associated with anterior circulation large-vessel occlusion stroke.

MATERIALS AND METHODS: Comprehensive cost-tracking software was used to calculate acute care hospitalization costs for patients with anterior circulation large-vessel occlusion stroke admitted July 2012 to October 2014. Patient demographics and stroke characteristics were analyzed, including final infarct volume on follow-up neuroimaging. Predictors of hospitalization costs were determined using multivariable linear regression including subgroup cost analyses by treatment technique (endovascular, IV tPA-only, and no reperfusion therapy) and sensitivity analyses incorporating patients initially excluded due to early withdrawal of care.

RESULTS: Three hundred forty-one patients (median age, 69 years; interquartile range, 57–80 years; median NIHSS score, 16; interquartile range, 13–21) were included in our primary analysis. Final infarct volume, parenchymal hematoma, baseline NIHSS score, ipsilateral carotid stenosis, age, and obstructive sleep apnea were significant predictors of acute care hospitalization costs. Final infarct volume alone accounted for 20.87% of the total cost variance. Additionally, final infarct volume was consistently the strongest predictor of increased cost in primary, subgroup, and sensitivity analyses.

CONCLUSIONS: Final infarct volume was the strongest predictor of increased hospitalization costs in anterior circulation large-vessel occlusion stroke. Acute stroke therapies that reduce final infarct volume may not only improve clinical outcomes but may also prove cost-effective.

ABBREVIATIONS: ACLVO = anterior circulation large-vessel occlusion; EWOC = early withdrawal of care; FIV = final infarct volume; IQR = interquartile range; mTICI = modified TICI score; PH1 = parenchymal hematoma type 1; PH2 = parenchymal hematoma type 2

It is estimated that direct medical costs for stroke treatment in the United States in 2015 reached $38 billion and will rise to $51.3 billion in 2020.[1] Large-vessel occlusion stroke, one of the most clinically devastating stroke subtypes, is believed to drive acute hospitalization costs in a disproportionate manner relative to other ischemic stroke etiologies.[2,3] Recent advances in endovascular treatment for anterior circulation large-vessel occlusion (ACLVO) stroke have dramatically improved functional outcomes,[4-10] with the unintended consequence of concentrating acute ACLVO stroke care and the corresponding health care expenses in select tertiary referral hospitals. The economic burden to these hospitals is an especially timely consideration.

Prior studies addressing the cost of hospitalization in ischemic stroke have found that endovascular treatment, intubation, baseline NIHSS, atrial fibrillation, ischemic heart disease, stroke sub-

Received April 26, 2018; accepted after revision October 21.

From the Department of Neurology (C.D.S., A.J.Z.), University of Minnesota, Minneapolis, Minnesota; Department of Neurology, Stroke Institute (C.D.S., S.L.P., B.T.J., A.P.J., T.G.J., S.R., D.T.C.), University of Pittsburgh Medical Center, Pittsburgh. Pennsylvania; Department of Neurology (S.R.), Emory University, Atlanta, Georgia; Department of Neurology (D.T.C.), WellStar Kennestone Hospital, Marietta, Georgia; and Clinical Translational Science Institute (D.G.W.), University of Pittsburgh, Pittsburgh, Pennsylvania.

Dr Streib was a NIH StrokeNET fellow supported by 1U01NS086489–02 (Principal Investigator Tudor G. Jovin and Lawrence Wechsler). Dr Rangaraju was a National Institutes of Health StrokeNET fellow supported by 1U10NS086607–01 (Principal Investigator, Michael Frankel) and is also a recipient of a clinical research training fellowship from the American Brain Foundation.

This project was supported by the National Institutes of Health through grant No. UL1-TR-000005.

Data previously presented as a platform presentation at: International Stroke Conference, February 22–24, 2017; Houston, Texas.

Please address correspondence to Tudor G. Jovin, MD, UPMC Stroke Institute, 200 Lothrop St, Suite C-400, Pittsburgh, PA 15213; e-mail: jovitg@upmc.edu; @PittStroke

Indicates open access to non-subscribers at www.ajnr.org
Indicates article with supplemental on-line tables.
Indicates article with supplemental on-line photos.
http://dx.doi.org/10.3174/ajnr.A5917

type, diabetes mellitus, age, sex, and dehydration are significant predictors of stroke hospitalization costs.[11-16] Hospitalization costs are typically calculated from billing charges, diagnostic codes, and insurer or Medicare payments. However, such indirect accounting methods do not accurately reflect the true cost of delivering medical care from the perspective of the hospital.[15-18] In July 2012, the University of Pittsburgh Medical Center (UPMC) implemented proprietary, comprehensive cost-tracking software, which captures hospitalization costs with considerably greater patient-level detail than traditional methods.[19] Our aim was to use this data capture paradigm to determine critical predictors of hospitalization costs in ACLVO stroke.

MATERIALS AND METHODS

Data Sources and Subjects

Institutional review board approval was obtained. Our patient cohort was derived from 2 prospectively collected databases: our Get With The Guidelines stroke data base and our endovascular stroke data base. Patients with acute stroke presenting to our hospital with a primary diagnosis of ACLVO stroke between July 2012 and September 2014 were eligible for analysis. To avoid confounding, we restricted the analysis to patients receiving intensive medical treatment. Withdrawal of care leading to discharge or death within the first week of hospitalization constituted an exclusion criterion. Patients with lack of follow-up imaging or extended hospitalization due to factors unrelated to stroke were also excluded.

Patient demographics, medical history, admission laboratory values, medications, stroke characteristics, and neuroimaging findings were incorporated into the analysis. Stroke characteristics and neuroimaging findings included the following: affected hemisphere, baseline NIHSS score, level of arterial occlusion, final infarct volume (FIV), parenchymal hematoma type 1 (PH1), and parenchymal hematoma type 2 (PH2). Etiologic factors such as underlying atrial fibrillation, ipsilateral carotid stenosis, and carotid dissection were also studied. The primary reperfusion technique was classified as endovascular treatment, IV tPA, or no reperfusion therapy; patients who received both endovascular treatment and IV tPA were analyzed as having endovascular treatment.

Measurements

The level of occlusion was defined as the extracranial internal carotid artery, intracranial internal carotid artery, middle cerebral artery M1 division, middle cerebral artery M2 division, or tandem lesions, as determined by catheter-based angiography when available or the initial vessel imaging study (CTA or MRA). The Alberta Stroke Program Early CT Scores were interpreted by the attending vascular neurologist and recorded at the time of admission for patients undergoing endovascular therapy. Revascularization status, the modified Thrombolysis in Cerebral Infarction score (mTICI), was assessed by the interventionalist post-endovascular treatment and prospectively recorded. FIV was calculated on MR imaging or CT scans obtained 6–48 hours postadmission by measuring the infarct volume on each slice and then summating the infarct volumes of the individual slices according to previously published methodology.[20] We have demonstrated high correlation between infarct volumes calculated by CT and MR imaging within our database in prior studies.[20]

Cost Analysis

Our institution developed novel, automated comprehensive cost-tracking software, which was implemented July 1, 2012. The cost algorithms of the software incorporate patient-level information to generate individualized patient-level cost data for each admission. For example, rather than assigning a uniform cost to all "stroke protocol" brain MR imaging scans, costs are calculated by MR imaging acquisition time. Accordingly, MR imaging scans with longer acquisition times are assigned a higher cost value, which accounts for the increased nursing and MR imaging technician resources required, as well as the depreciation of the MR imaging scanner (simplified equation: $Cost_{MRI} = (MRI\ Tech\ Salary \times Time_{MRI}) + (Nurse\ Salary \times Time_{MRI}) + (MRI\ Depreciation \times Time_{MRI})$. This calculation differentiates the cost of MR imaging for an uncooperative, aphasic patient with a high NIHSS score versus a cooperative patient with a low NIHSS score. This level of detail is captured across all hospital cost domains, including the following: supplies, drugs, blood products, clinical ancillary services, diagnostic testing, imaging, laboratory, dietary, intensive care unit nursing, non-intensive care unit nursing, pharmacy, surgical services, housekeeping, and miscellaneous expenses, which were included in our analysis. Costs were calculated from the hospital's perspective and encompass the duration of the patient's acute care hospitalization. Hospitalization costs were analyzed objectively, independent of reimbursement considerations.

Physicians' costs as calculated by our software were highly variable because of differences in physicians' salaries and the source of physicians' salary support. These costs did not accurately correlate with the levels of service provided. To avoid introducing imprecision into the analysis, we excluded physicians' costs. Similarly, transfer costs, which occur randomly, and indirect costs, such as administrative salaries of nonhealth care providers, were also excluded. We did not adjust for inflation because the study duration was limited to a 28-month period.

Statistical Analysis

Distributions of continuous variables were assessed for normality. We performed a log-transformation of the hospital cost data to correct for rightward skewing. All baseline demographics and stroke characteristics were included in univariate analyses. Variables with a P value $< .25$ in univariate linear regression analyses were entered into multivariable linear regression to determine base cost models. Variables with a P value $< .05$ in multivariable linear regression were considered statistically significant. The final model was evaluated for heteroscedasticity graphically by plotting the fitted-versus-residual values for each subject and formally using the Breusch-Pagan test. If the variance inflation factor was >2, it would be addressed by removing the less biologically plausible variable from the model. The goodness-of-fit of the models was assessed by R^2. For ease of interpretation, β coefficients of each statistically significant variable were exponentiated. Following this reverse transformation, the exponentiated β coefficients represent the percentage change in hospitalization costs

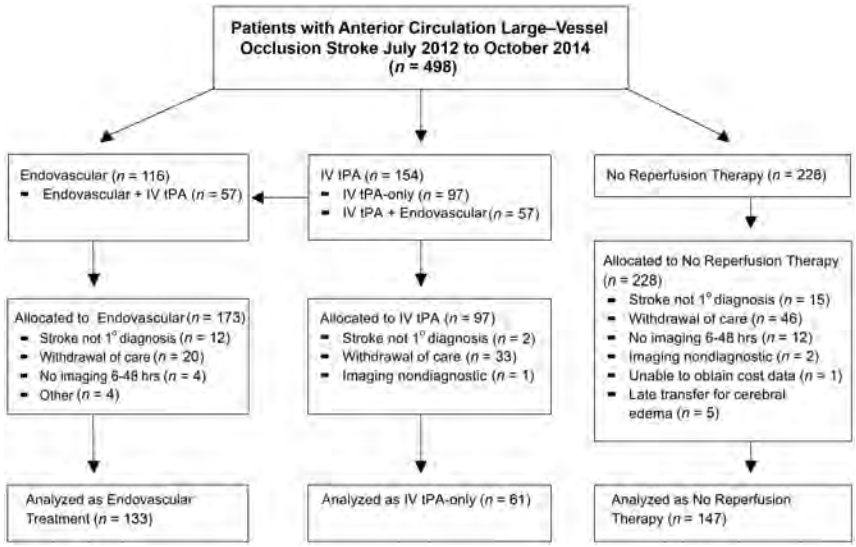


FIG 1. Study flow diagram. Endovascular stroke data base July 2014 to October 2014: three hundred twenty-four total stroke interventions (66 posterior circulation, 88 at an affiliated hospital without cost-tracking) with 173 patients eligible for anterior circulation large-vessel occlusion.

attributable to a 1-unit increase of each continuous variable or the percentage change associated with the presence of a categorical variable. Within each model, the magnitude of the effect size of each variable was quantified by partial eta^2.

Because endovascular treatment and IV tPA administration are proven predictors of increased hospitalization costs,[11,15,16] the primary analysis was performed on all eligible patients with ACLVO while controlling for treatment technique (endovascular, IV tPA-only, and no reperfusion therapy) by including treatment technique in the multivariable linear regression model. We then performed subgroup analyses on each treatment arm. Finally, we performed sensitivity analyses incorporating previously excluded patients to assess the robustness of our findings. The statistical analysis was conducted with STATA software, Version 13.0 (StataCorp, College Station, Texas).

RESULTS

Of the 498 patients who presented to our institution with an ACLVO stroke during the study period, 345 patients met the general inclusion criteria. Of those patients, 4 were excluded because their hospitalization was prolonged >21 days for reasons unrelated to stroke (myotonic dystrophy = 1, alcohol withdrawal = 1, delayed discharge placement = 2). Three hundred forty-one patients were included in the primary analysis and comprised 133 patients with endovascular treatment, 61 patients with IV tPA-only, and 147 patients who received no reperfusion therapy (Fig 1). Median hospitalization costs were $21,871 (interquartile range [IQR], $15,672–$31,363) for the endovascular therapy group, $14,456 (IQR, $7626–$19,701) for the IV tPA-only group, and $13,401 (IQR, $8308–$23,589) for the no reperfusion therapy group. The median age of our cohort was 69 years (IQR, 57–80 years) with a median baseline NIHSS score of 16 (IQR, 13–21). Complete patient demographics and stroke characteristics are summarized in Tables 1 and 2, respectively.

Primary Analysis

Results from our univariate analysis (On-line Table 1) guided construction of our best-fit model using multivariable linear regression. In the primary analysis of all patients with ACLVO while controlling for the primary reperfusion technique, significant predictors of log-transformed hospitalization costs included the following: FIV ($P < 0.001$); the composite variable PH1 or PH2 ($P < .001$); baseline NIHSS score ($P = .002$); obstructive sleep apnea ($P = .0040$); age ($P = .009$); and ipsilateral carotid stenosis ($P = .020$). Age was the only variable that was inversely associated with hospitalization costs. The final model explained 42.07% of the variance in hospitalization costs ($R^2 = 42.07\%$). The FIV was identified as the strongest predictor of hospitalization costs, accounting for 20.87% of the total hospitalization cost variance as calculated by partial eta^2. The variables did not demonstrate significant collinearity, nor did the model have heteroscedasticity in the primary or subgroup analyses.

Subgroup Analysis

The FIV was the only variable that remained statistically significant across all subgroup models. Irrespective of the treatment technique, FIV was the strongest predictor of hospitalization costs as determined by partial eta^2. PH1 or the related composite variable, PH1 or PH2, were strongly correlated with cost in all except the IV tPA-only subgroup, as was obstructive sleep apnea. The baseline NIHSS score was a significant predictor in all models with the exception of the subgroup with no reperfusion therapy. Ipsilateral carotid stenosis and the mTICI score were statistically significant predictors of cost for the endovascular treatment subgroup only, while age was inversely associated with cost in the subgroup with no reperfusion therapy. The P values and exponentiated β coefficients of the statistically significant variables for each model are summarized in Table 3.

Sensitivity Analysis

Of the 99 patients excluded from the primary and subgroup analyses due to early withdrawal of care (EWOC), complete data were available for 88 patients. EWOC led to significantly lower hospitalization cost (median, $6664; IQR, $4868–$13,765) compared with aggressive medical and interventional treatment ($P < .0001$). Patients with EWOC were older, with larger infarct volumes, higher baseline NIHSS scores, and increased rates of PH1 and PH2 (On-line Table 2). When patients with EWOC were included, age, FIV, and PH1 remained significant predictors of cost, but the model had a considerably poorer fit ($R^2 = 30.10\%$) and the association among cost, FIV, and PH1 was attenuated due to confounding. Accounting for EWOC status generated a considerably more robust model ($R^2 = 42.53$) in which FIV, followed by PH1, remained consistently strong predictors of increased hospitalization cost (On-line Table 3).

Table 1: Baseline patient demographics[a]

	All Patients (*n* = 341)	Endovascular Therapy (*n* = 133)	IV tPA-Only (*n* = 61)	No Reperfusion Therapy (*n* = 147)
Cost (median) (IQR)	$16,446 ($9823–$27,165)	$21,871 ($15,672–$31,363)	$14,456 ($7626–$19,701)	$13,401 ($8308–$23,589)
Age (median) (IQR)	69 (57–80)	68 (58–80)	75 (59–83)	66 (55–79)
Female	165 (48.4%)	65 (48.9%)	33 (54.1%)	67 (45.6%)
A. Fib	162 (47.5%)	65 (48.9%)	37 (60.7%)	60 (40.9%)
Alcohol abuse	27 (7.9%)	7 (5.3%)	4 (6.6%)	16 (10.9%)
Anticoagulation	29 (8.5%)	11 (8.3%)	5 (8.2%)	13 (8.8%)
CAD	78 (23.0%)	26 (19.6%)	19 (31.7%)	33 (22.6%)
CAS	72 (21.1%)	23 (17.3%)	10 (16.4%)	39 (26.5%)
CHF	59 (17.5%)	15 (11.4%)	17 (28.3%)	27 (18.5%)
CKD	19 (5.6%)	3 (2.3%)	4 (6.7%)	12 (8.2%)
Dementia	15 (4.4%)	2 (1.5%)	7 (11.7%)	6 (4.1%)
Dissection	17 (5.0%)	9 (6.8%)	2 (3.3%)	6 (4.1%)
DMII	103 (30.2%)	51 (38.3%)	25 (41.0%)	27 (18.4%)
Hyperlipidemia	149 (43.7%)	64 (48.1%)	29 (47.5%)	56 (38.1%)
Hypertension	291 (85.3%)	111 (83.5%)	55 (90.2%)	125 (85.0%)
Illicit drug use	8 (2.3%)	0 (0.0%)	0 (0.0%)	8 (5.5%)
Obesity	44 (12.9%)	12 (9.0%)	19 (31.1%)	13 (8.8%)
OSA	31 (9.2%)	12 (9.0%)	8 (13.1%)	11 (7.5%)
Psychiatric	33 (9.7%)	9 (6.8%)	7 (11.5%)	17 (11.6%)
Tobacco	124 (36.6%)	40 (30.1%)	21 (34.4%)	63 (42.9%)

Note:—A. Fib indicates atrial fibrillation; CAD, coronary artery disease; Dissection, ipsilateral carotid dissection; CAS, ipsilateral carotid artery stenosis; CHF, congestive heart failure; CKD, chronic kidney disease; DMII, diabetes mellitus II; OSA, obstructive sleep apnea; Psychiatric, any psychiatric comorbidity.
[a] Values represent total number and percentage unless otherwise indicated.

Table 2: Stroke characteristics[a]

	All Patients (*n* = 341)	Endovascular Therapy (*n* = 133)	IV tPA-Only (*n* = 61)	No Reperfusion Therapy (*n* = 147)
Baseline NIHSS (median) (IQR)	16 (13–21)	16 (13–19)	16 (12–22)	17 (13–21)
FIV (mL) (median) (IQR)	59.9 (17.1–129)	27 (9.4–86.0)	62.5 (8.7–118.7)	93 (37.8–170.0)
Left hemisphere	175 (51.3%)	65 (48.9%)	38 (62.3%)	72 (48.9%)
Level of occlusion				
Extracranial ICA	7 (2.0%)	2 (1.5%)	4 (6.6%)	1 (0.6%)
Tandem occlusion	39 (11.4%)	31 (23.3%)	2 (3.3%)	6 (4.1%)
Intracranial ICA	79 (23.2%)	16 (12.0%)	9 (14.8%)	54 (36.7%)
M1 division of MCA	166 (48.7%)	74 (55.6%)	33 (54.1%)	59 (40.1%)
M2 division of MCA	50 (14.7%)	10 (7.5%)	13 (21.3%)	27 (18.4%)
PH1	14 (4.1%)	3 (2.3%)	8 (13.1%)	4 (2.7%)
PH2	9 (2.6%)	5 (3.8%)	3 (4.9%)	1 (0.7%)
ASPECTS (median) (IQR)	Data unavailable	8 (7–9)	Data unavailable	Data unavailable
Wake-up stroke	Data unavailable	19 (14.3%)	Data unavailable	Data unavailable
Onset to treatment (min) (median) (IQR)		310 (236–465)	138 (115–155)	
IA intubation		5 (3.7%)		
Recanalization (mTICI)				
mTICI 0		1 (0.8%)		
mTICI 1		1 (0.8%)		
mTICI 2a		4 (3.0%)		
mTICI 2b		91 (68.7%)		
mTICI 3		36 (26.9%)		

Note:—IA indicates intra-arterial (ie, endovascular treatment).
[a] Values represent total number and percentage unless otherwise indicated.

A second sensitivity analysis included 11 of 12 patients whose follow-up imaging was performed outside our prespecified 6- to 48-hour time window. Statistically significant predictors of hospitalization costs were identical to those in the primary analysis (On-line Table 3).

DISCUSSION

We found that in patients with ACLVO stroke, FIV, a well-known predictor of clinical outcome,[8,9,21,22] is also the most robust determinant of hospitalization costs. FIV had the strongest association with hospitalization costs in univariate analysis (Fig 2), a relationship that persisted in all multivariable analyses, including our primary analysis of all patients with ACLVO and subgroup analyses of patients with endovascular therapy, IV tPA, or no reperfusion therapy. Although it is intuitive that larger stroke volumes lead to more resource-intensive hospital admissions, we do not believe that this fundamental relationship between infarct volume and hospitalization costs has been previously reported.

We constructed our cost-prediction models from baseline patient demographics and stroke characteristics alone. Given that we limited our models to variables with P values $< .05$ and inten-

Table 3: Significant predictors of cost

	P Value	Partial Eta^2	Exponentiated β-Coefficient (95% CI)[b]
Primary analysis: all ACLVO[a] (n = 341)			
FIV (mL)	<.0001	20.87%	1.0037 (1.0029–1.0048)
PH types 1 and 2	<.0001	5.22%	1.6399 (1.3050–2.0610)
NIHSS	.0024	2.76%	1.0161 (1.0057–1.0266)
Obstructive sleep apnea	.0040	2.48%	1.3391 (1.0981–1.6329)
Age (yr)	.0089	2.06%	0.9946 (0.9906–0.9986)
Ipsilateral carotid stenosis	.0195	1.65%	1.1902 (1.0285–1.3774)
Subgroup analysis by treatment group			
Endovascular therapy (n = 133)			
FIV (mL)	.0003	10.03%	1.0024 (1.0011–1.0037)
PH type 1	.0006	9.24%	2.4942 (1.4926–4.1680)
Obstructive sleep apnea	.0018	7.71%	1.5211 (1.1727–1.9730)
NIHSS	.0120	5.08%	1.0206 (1.0046–1.0368)
Ipsilateral carotid stenosis	.0290	3.85%	1.2481 (1.0234–1.5221)
mTICI score[c]	.0364	1.79%	Multilevel variable
IV tPA-only (n = 61)			
FIV (mL)	<.0001	29.11%	1.0043 (1.0025–1.0060)
NIHSS	.0300	7.99%	1.0273 (1.0273–1.0524)
No reperfusion therapy (n = 147)			
FIV (mL)	<.0001	27.92%	1.0042 (1.0031–1.0053)
PH types 1 and 2	.0005	8.37%	2.6371 (1.5458–4.4988)
Obstructive sleep apnea	.0300	3.29%	1.5057 (1.0407–2.1784)
Age (yr)	.0330	3.20%	0.9925 (0.9857–0.9937)

[a] Controlled for treatment type: endovascular, IV-tPA, and no reperfusion therapy.
[b] The exponentiated β-coefficient represents the multiplicative change in cost associated with an increase in 1 unit of a continuous variable (or the presence of a categoric variable), while holding all other variables constant.
[c] Multilevel variable with *F*-distribution.

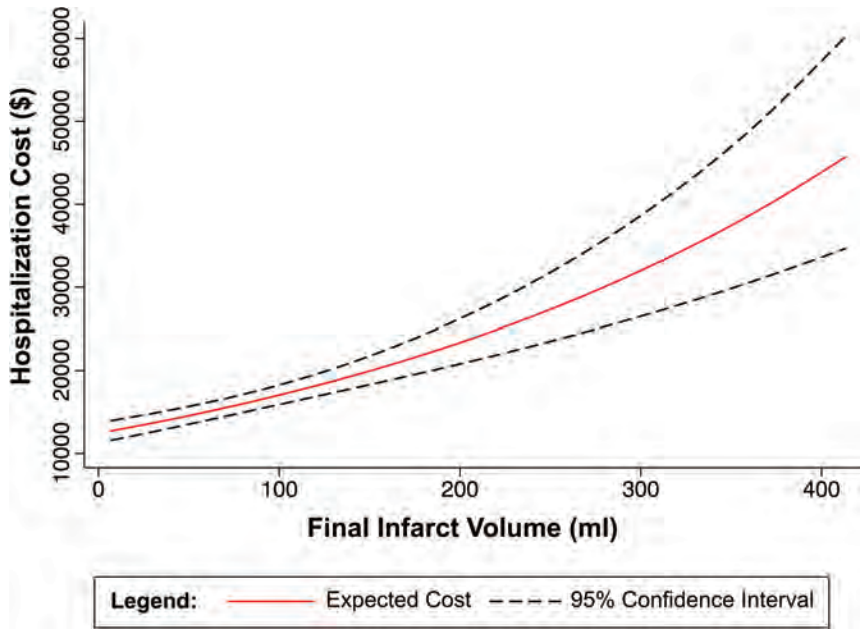


FIG 2. Univariate analysis: cost versus final infarct volume.

tionally excluded postadmission variables such as length of stay, intubation, and decompressive craniectomy, we find it remarkable that our primary model still accounted for 42.07% of the total variance of (log-transformed) hospitalization costs. FIV alone accounted for 20.87% of the variance in hospitalization costs. When interpreted across the range of observed FIVs, a patient with a 249.2-mL infarct (95th percentile) would have a 148.61% higher hospitalization cost than an otherwise identical patient with a 2.0-mL infarct (fifth percentile). Under this hypothetic scenario, with the mean hospitalization cost of $20,351.23 as a reference point, such a change in FIV would increase hospitalization costs by $30,244.13, or $122.35 for each additional milliliter of stroke burden (Table 4). Notably, in our analyses, the relationship between cost and FIV is exponential rather than linear, meaning that larger increases in FIV have a more pronounced impact (Fig 2). Another consideration is that our primary objective was to accurately identify critical predictors of increased ACLVO acute care hospitalization costs. By excluding costs that were highly variable (ie, physicians' costs) or occurring at random (ie, transfer costs) from the model, we improved our ability to correctly identify consistent predictors of hospitalization costs but under-represented the total cost to the hospital.

In addition to FIV, we also found that PH1, baseline NIHSS, obstructive sleep apnea, age, and ipsilateral carotid stenosis were significant predictors of hospitalization costs in our primary analysis. With the exception of ipsilateral carotid stenosis, our findings are consistent with prior studies of cost or clinical outcome.[11,13,23,24] Increased hospitalization costs in patients with ACLVO with ipsilateral carotid stenosis are presumably a consequence of additional expenses accrued from endarterectomy or carotid stent placement.

Differences in the cost models generated by our subgroup analysis may be explained by inherent differences among patient populations. For example, carotid stenosis was a significant predictor of hospitalization cost only in the endovascular therapy subgroup. Rates of carotid revascularization in patients with moderate or severe symptomatic carotid stenosis ranged from 100% in the endovascular group to 30% and 28% in the IV tPA-only and no reperfusion therapy groups, respectively. This discrepancy is a result of our institutional practice of revascularizing all symptomatic carotid arteries during the initial hospitalization unless contraindicated by the risk of hemorrhagic conversion attributable to high FIV. Therefore, patients with larger strokes, such as those seen in the IV tPA-only and no reperfusion therapy subgroups, undergo carotid revascularization less frequently, decreasing the impact of carotid stenosis on acute care hospitalization costs. Conversely,

Table 4: Expected cost burden per additional infarct volume[a]

Increased Stroke Burden (mL)	Expected Increase in Hospitalization Cost	Net Expected Cost Increase	Cost per Additional 1 mL of Stroke Burden
1	0.37%	$75.11	$75.11
10	3.75%	$763.75	$76.38
50	20.23%	$4116.33	$82.33
100	44.54%	$9065.25	$90.65
150	73.78%	$15,015.15	$100.10
200	108.93%	$22,168.52	$110.84
247.2[b]	148.61%	$30,244.13	$122.35
300	202.00%	$41,108.50	$137.03
350	263.09%	$53,539.62	$152.97
400	336.52%	$68,485.11	$171.21
413.3[c]	358.44%	$72,946.41	$176.50

[a] With reference to mean cost = $20,351.23 and primary multivariable analysis (Table 3), holding all variables constant while increasing the FIV.
[b] Range of FIVs observed (fifth–95th percentile).
[c] Full range of observed FIVs (0–100th percentile).

age was not a significant predictor of cost in the endovascular subgroup. A potential explanation for this finding is that only elderly patients with excellent baseline functional status were selected for endovascular stroke treatment.

Our subgroup of patients undergoing endovascular thrombectomy was treated before the publication of randomized controlled trials demonstrating the effectiveness of endovascular therapy. However, these randomized controlled trials did not meaningfully change our institutional practice, and our findings remain generalizable. The NIHSS scores and Alberta Stroke Program Early CT Scores of our endovascular therapy cohort are also comparable with those reported in recent endovascular stroke randomized controlled trials, and our patients were similarly screened for pretreatment functional independence. Additionally, the primary endovascular treatment technique used at our institution was the stent retriever, resulting in >90% mTICI 2b or 3 reperfusion rates. We also studied a wide time window of stroke onset to treatment (interquartile range, 236–465 minutes) with ~15% of the cohort having wake-up stroke (Tables 1 and 2). Finally, there have not been substantial changes in post thrombectomy medical care or substantial inflation that would impact the validity of our key findings.

One of the limitations of our study is that the findings may be specific to high-volume, tertiary referral, academic hospitals. However, while absolute costs may differ across hospitals, predictors of cost are likely to be similar because they are ultimately determinants of resource use. FIV is known to be associated with inpatient interventions such as tracheostomy, feeding tube placement, and decompressive craniectomy, which both add to hospitalization cost and prolong hospital admission.[25,26] Further supporting this hypothesis, a post hoc exploratory analysis of our dataset demonstrated a clear, significant relationship ($P < .001$) between increasing FIV and prolonged hospitalization (On-line Fig 1). Thus, we believe our findings should be applicable to a broad range of hospitals. Additionally, intermediate and long-term care, key components of stroke health care costs, were not taken into account in our analysis. In the Endovascular Therapy for Ischemic Stroke with Perfusion-Imaging Selection (EXTEND-IA) trial, Campbell et al[9] reported that patients who received endovascular treatment had significantly lower FIVs and returned home earlier than the control group. Accordingly, it can be extrapolated that lower FIVs are likely to result in decreased intermediate and long-term health care costs, though this hypothesis requires further investigation.[27]

Another limitation is the 6- to 48-hour time window used to calculate infarct volumes, which is a reflection of our clinical practice. We obtain early MR imaging scans on patients undergoing endovascular recanalization treatment to assess for hemorrhage and guide postintervention antithrombotic therapy, whereas patients without endovascular recanalization are scanned later for prognostication and management of cerebral edema. Patients receiving IV tPA undergo follow-up imaging 24 hours post-treatment per protocol, while the timing of imaging in patients who do not receive reperfusion therapy is determined on a case-by-case basis. Early imaging may result in an underappreciation of infarct volume. However, most scans obtained <24 hours from admission were either for patients having undergone endovascular recanalization or those not receiving reperfusion therapy who presented with large completed infarcts (On-line Fig 2); these patients are unlikely to experience substantial infarct growth.[9] By standardizing our infarct volume calculations to scans obtained between 6 and 48 hours from admission, we limited potential inconsistency in infarct volume calculations attributable to more variable imaging timing and progression and/or resolution of cerebral edema. We acknowledge that the non-uniform timing of follow-up imaging may introduce imprecision into our calculations, but we do not believe it alters our fundamental findings.

A further limitation is that our findings only apply to ACLVO stroke. We intentionally excluded posterior circulation large-vessel occlusions because treatment paradigms in this patient population are not driven uniformly by high-level randomized clinical trial data. Additionally, it appears that in ACLVO stroke lesion location may not be tightly correlated with outcome.[28] Outcomes and hospital admission complexity in posterior circulation stroke, however, may be more dependent on lesion location due to the ramifications of brain stem infarction on consciousness and respiratory drive.[29]

Finally, 99 patients were excluded from our primary analysis because of EWOC. Patients with EWOC represent a clinically distinct patient population, and their inclusion in our sensitivity analysis confounds the model. However, controlling for EWOC status produced the same fundamental finding with FIV the most robust predictor of increased hospitalization costs. Furthermore, although EWOC significantly decreases hospitalization costs, it remains a clinically undesirable outcome. Pursuing EWOC may be appropriate in a subset of patients with ACLVO, but ideally cost-effectiveness strategies target cost savings and improved clinical outcomes in parallel.

CONCLUSIONS

We found that FIV, a well-known predictor of stroke-related disability,[7,10,21,22] was the strongest predictor of increased hospital-

ization costs in ACLVO stroke at our institution. Furthermore, this relationship was exponential—that is, greater absolute increases in FIV have a considerably greater effect on cost. A notable strength of our cost analysis is the use of detailed patient-level cost data rather than traditional indirect accounting methods. Additionally, our study has a relatively large sample size and broad inclusion criteria and demonstrated consistent findings across primary, subgroup, and sensitivity analyses. To our knowledge, this is the first time that the fundamental relationship between FIV and hospitalization costs has been reported. This finding has considerable economic implications for the treatment of ACLVO stroke. Therapies that reduce FIV not only improve clinical outcomes, but may also be critical to providing cost-effective treatment.

ACKNOWLEDGMENTS

The authors would like to thank Jon Petrie for assistance in obtaining and interpreting institutional cost data.

Disclosures: Christopher D. Streib—*RELATED*: *Grant*: National Institutes of Health StrokeNet, *Comments*: National Institutes of Health StrokeNet Fellowship grant University of Pittsburgh (1U01NS086489–02).* Srikant Rangaraju—*UNRELATED*: *Grants/Grants Pending*: National Institutes of Health/National Institute of Neurological Disorders and Stroke K08.* Daniel G. Winger—*RELATED*: *Grant*: National Institutes of Health, *Comments*: The project described was supported by the National Institutes of Health through grant number UL1-TR-000005.* Tudor G. Jovin—*UNRELATED*: *Consultancy*: Cerenovus, Biogen; *Stock/Stock Options*: Anaconda, Silk Road, Route 92, Blockade Medical, FreeOx Biotech. Brian T. Jankowitz—*UNRELATED*: *Consultancy*: Stryker, Medtronic; *Money paid to the institution.

REFERENCES

1. Heidenreich PA, Trogdon JG, Khavjou OA, et al; American Heart Association Advocacy Coordinating Committee, Stroke Council, Council on Cardiovascular Radiology and Intervention, Council on Clinical Cardiology, Council on Epidemiology and Prevention, Council on Arteriosclerosis, Thrombosis and Vascular Biology, Council on Cardiopulmonary, Critical Care, Perioperative and Resuscitation, Council on Cardiovascular Nursing, Council on the Kidney in Cardiovascular Disease, Council on Cardiovascular Surgery and Anesthesia, and Interdisciplinary Council on Quality of Care and Outcomes Research. **Forecasting the future of cardiovascular disease in the United States: a policy statement from the American Heart Association.** *Circulation* 2011;123:933–44 CrossRef Medline
2. Smith WS, Lev MH, English JD, et al. **Significance of large vessel intracranial occlusion causing acute ischemic stroke and TIA.** *Stroke* 2009;40:3834–40 CrossRef Medline
3. Rai AT, Evans K, Riggs JE, et al. **Intravenous thrombolysis of large vessel occlusions is associated with higher hospital costs than small vessel strokes: a rationale for developing stroke severity-based financial models.** *J Neurointerv Surg* 2016;8:423–28 CrossRef Medline
4. Nogueira RG, Jadhav AP, Haussen DC, et al; DAWN Trial Investigators. **Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct.** *N Engl J Med* 2018;378:11–21 CrossRef Medline
5. Albers GW, Marks MP, Kemp S, et al; DEFUSE 3 Investigators. **Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging.** *N Engl J Med* 2018;378:708–18 CrossRef Medline
6. Saver JL, Goyal M, Bonafe A, et al; SWIFT PRIME Investigators. **Stent-retriever thrombectomy after intravenous t-PA vs. t-PA alone in stroke.** *N Engl J Med* 2015;372:2285–95 CrossRef Medline
7. Jovin TG, Chamorro A, Cobo E, et al; REVASCAT Trial Investigators. **Thrombectomy within 8 hours after symptom onset in ischemic stroke.** *N Engl J Med* 2015;372:2296–306 CrossRef Medline
8. 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
9. Campbell BC, Mitchell PJ, Kleinig TJ, et al; EXTEND-IA Investigators. **Endovascular therapy for ischemic stroke with perfusion-imaging selection.** *N Engl J Med* 2015;372:1009–18 CrossRef Medline
10. Berkhemer OA, Fransen PS, Beumer D, et al; MR CLEAN Investigators. **A randomized trial of intraarterial treatment for acute ischemic stroke.** *N Engl J Med* 2015;372:11–20 CrossRef Medline
11. Simpson KN, Simpson AN, Mauldin PD, et al; IMS III Investigators. **Drivers of costs associated with reperfusion therapy in acute stroke: the Interventional Management of Stroke III Trial.** *Stroke* 2014;45:1791–98 CrossRef Medline
12. Wang G, Joo H, Tong X, et al. **Hospital costs associated with atrial fibrillation for patients with ischemic stroke aged 18–64 years in the United States.** *Stroke* 2015;46:1314–20 CrossRef Medline
13. Wang G, Zhang Z, Ayala C, et al. **Costs of hospitalization for stroke patients aged 18–64 years in the United States.** *J Stroke Cerebrovasc Dis* 2014;23:861–68 CrossRef Medline
14. Huang YC, Hu CJ, Lee TH, et al. **The impact factors on the cost and length of stay among acute ischemic stroke.** *J Stroke Cerebrovasc Dis* 2013;22:e152–158 CrossRef Medline
15. Brinjikji W, Rabinstein AA, Cloft HJ. **Hospitalization costs for acute ischemic stroke patients treated with intravenous thrombolysis in the United States are substantially higher than Medicare payments.** *Stroke* 2012;43:1131–33 CrossRef Medline
16. Brinjikji W, Kallmes DF, Rabinstein AA, et al. **Hospitalization costs for patients with acute ischemic stroke treated with endovascular embolectomy in the United States.** *Stroke* 2011;42:3271–73 CrossRef Medline
17. Jones SA, Gottesman RF, Shahar E, et al. **Validity of hospital discharge diagnosis codes for stroke: the Atherosclerosis Risk in Communities Study.** *Stroke* 2014;45:3219–25 CrossRef Medline
18. Chen M. **Cost-effectiveness of endovascular therapy for acute ischemic stroke.** *Neurology* 2012;79:S16–21 CrossRef Medline
19. Boulton C. UPMC Selling Analytics to Curb Health Care Costs. *Wall Street Journal*. April 28, 2014. http://blogs.wsj.com/cio/2014/04/28/upmc-builds-analytics-to-curb-health-care-costs/. Accessed May 20, 2015
20. Rangaraju S, Liggins JT, Aghaebrahim A, et al. **Pittsburgh outcomes after stroke thrombectomy score predicts outcomes after endovascular therapy for anterior circulation large vessel occlusions.** *Stroke* 2014;45:2298–304 CrossRef Medline
21. Ribo M, Flores A, Mansilla E, et al. **Age-adjusted infarct volume threshold for good outcome after endovascular treatment.** *J Neurointerv Surg* 2014;6:418–22 CrossRef Medline
22. Zaidi SF, Aghaebrahim A, Urra X, et al. **Final infarct volume is a stronger predictor of outcome than recanalization in patients with proximal middle cerebral artery occlusion treated with endovascular therapy.** *Stroke* 2012;43:3238–44 CrossRef Medline
23. Mansukhani MP, Bellolio MF, Kolla BP, et al. **Worse outcome after stroke in patients with obstructive sleep apnea: an observational cohort study.** *J Stroke Cerebrovasc Dis* 2011;20:401–05 CrossRef Medline
24. Rao NM, Levine SR, Gornbein JA, et al. **Defining clinically relevant cerebral hemorrhage after thrombolytic therapy for stroke: analysis of the National Institute of Neurological Disorders and Stroke tissue-type plasminogen activator trials.** *Stroke* 2014;45:2728–33 CrossRef Medline
25. Walcott BP, Miller JC, Kwon CS, et al. **Outcomes in severe middle cerebral artery ischemic stroke.** *Neurocrit Care* 2014;21:20–26 CrossRef Medline
26. Thomalla G, Hartmann F, Juettler E, et al; Clinical Trial Net of the German Competence Network Stroke. **Prediction of malignant middle cerebral artery infarction by magnetic resonance imaging within 6 hours of symptom onset: a prospective multicenter**

observational study. *Ann Neurol* 2010;68:435–45 CrossRef Medline
27. Leppert MH, Campbell JD, Simpson JR, et al. **Cost-effectiveness of intra-arterial treatment as an adjunct to intravenous tissue-type plasminogen activator for acute ischemic stroke.** *Stroke* 2015;46: 1870–76 CrossRef Medline
28. Rangaraju S, Streib C, Aghaebrahim A, et al. **Relationship between lesion topology and clinical outcome in anterior circulation large vessel occlusions.** *Stroke* 2015;46:1787–92 CrossRef Medline
29. Campbell D, Rangaraju S, Streib C, et al. **Pittsburgh Outcomes After Stroke Thrombectomy Vertebrobasilar (POST VB) score predicts outcomes after endovascular therapy for basilar artery occlusions**. In: *Proceedings of the International Stroke Conference*, San Diego, California; February 12–14, 2014

MRI Quantitative T2* Mapping to Predict Dominant Composition of In Vitro Thrombus

R. Bourcier, R. Pautre, M. Mirza, C. Castets, J. Darcourt, J. Labreuche, L. Detraz, H. Desal, J.-M. Serfaty, and C. Toquet

ABSTRACT

BACKGROUND AND PURPOSE: MR imaging quantitative T2* mapping, which provides information about thrombus composition and specifically the red blood cell content, may be obtained in the setting of acute ischemic stroke before treatment. This could be useful to adapt the endovascular strategy. We aimed to analyze the red blood cell content of in vitro thrombi in relation to the thrombus-T2* relaxation time.

MATERIALS AND METHODS: Thirty-five thrombus analogs of different compositions were scanned with an MR imaging quantitative T2* mapping sequence. Two radiologists, blinded to thrombus composition, measured the thrombus-T2* relaxation time twice at an interval of 2 weeks. Quantitative histologic evaluations of red blood cell content were performed. Inter- and intraobserver reproducibility of the thrombus-T2* relaxation time was assessed by calculating intraclass correlation coefficients. Finally, a Spearman product moment correlation between the thrombus-T2* relaxation time and red blood cell content was performed.

RESULTS: The median thrombus-T2* relaxation time was 78.5 ms (range, 16–268 ms; interquartile range, 60.5 ms). The median red blood cell content was 55% (range, 0%–100%; interquartile range, 75%). Inter- and intraobserver reproducibility of the thrombus-T2* relaxation time was excellent (>0.9). The Spearman rank correlation test found a significant inverse correlation between thrombus-T2* relaxation time and red blood cell content ($\rho = -0.834$, $P < .001$).

CONCLUSIONS: MR imaging quantitative T2* mapping can reliably identify the thrombus red blood cell content in vitro. This fast, easy-to-use sequence could be implemented in routine practice to predict stroke etiology and adapt devices or techniques for endovascular treatment of acute ischemic stroke.

ABBREVIATIONS: RBC = red blood cell; SVS = susceptibility vessel sign; TT2*RT = thrombus-T2* relaxation time

Sudden occlusion of an intracranial artery by a thrombus represents the pivotal event in anterior acute ischemic stroke. Consequently, the primary goal of acute stroke treatment is to restore perfusion by lysing the thrombus with intravenous thrombolysis and/or retrieving the thrombus with mechanical thrombectomy.[1] The red blood cell (RBC) content may be especially critical to tailor the endovascular treatment. As demonstrated by Hashimoto et al,[2] the interaction between the stent-retriever struts and the thrombus is likely dependent on the number of RBCs in the thrombus.[3] Hence a first-line imaging-screening protocol that can provide the proportion of RBCs in the whole thrombus would be helpful when planning a treatment strategy. In this field, the T2* sequence offers acquisition times suited to the context of emergency acute ischemic stroke. The susceptibility vessel sign (SVS) on T2* MR imaging is defined by a hypointense signal, exceeding the diameter of the contralateral artery, located at the site of the thrombus.[4] A positive SVS (SVS+) correlates with an RBC-dominant thrombus, whereas a lack of SVS (SVS−) indicates a fibrin-dominant thrombus.[3]

However, a binary assessment of thrombus composition (either SVS+ or SVS−) may not reflect the complex nature of thrombi in which different regions may have different content of RBCs.[5-7] Moreover, SVS+ can be difficult to determine in

Received August 8, 2018; accepted after revision October 21.

From the Departments of Neuroradiology (R.B., R.P., L.D., H.D.) and Cardiac and Vascular Imaging (J.-M.S.), Hôpital René et Guillaume Laennec, Centre Hospitalier Universitaire de Nantes, Nantes, France; Neuravi Thromboembolic Initiative (M.M.), Galway, Ireland; Siemens Healthineers France (C.C.), Saint-Denis, France; Department of Neuroradiology (J.D.), Hôpital Purpan, Centre Hospitalier Universitaire de Toulouse, Toulouse, France; Department of Biostatistics (J.L.), Université de Lille, Centre Hospitalier Universitaire Lille, Lille, France; and Department of Pathology (C.T.), Hôtel Dieu, Centre Hospitalier Universitaire de Nantes, Nantes, France.

Jean-Michel Serfaty and Claire Toquet contributed equally to this work.

Please address correspondence to Romain Bourcier, MD, PhD, Service de Neuroradiologie, Centre Hospitalier Universitaire Guillaume et René Laennec, Boulevard du Professeur Jacques Monodm, 44 093 Nantes, France; e-mail: romain.bourcier2@gmail.com, romain.bourcier@chu-nantes

http://dx.doi.org/10.3174/ajnr.A5938

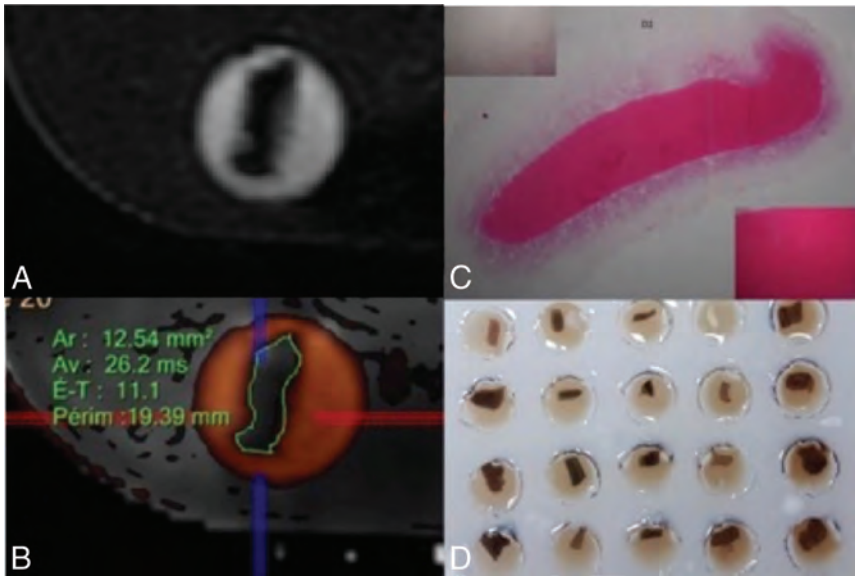


FIG 1. *A*, MR images of 3D-CISS analog thrombi. *B*, MR images of coregistered T2 mapping and 3D-CISS analog thrombus. Thrombi were manually contoured and analyzed in their largest surface, with the support of multiplanar reconstruction. Thrombi were identified by their localization in the matrix to allow comparison with histologic data. *C*, Macroscopic view of floating thrombi. *D*, Matrix of agarose gel with holes filled with physiologic serum. Ar indicates surface; Av, mean T2* relaxation time; E-T, SD.

clinical practice because some thrombi may show a hypointense signal with a diameter equal to or slightly above that of the contralateral artery. Last, the diagnostic accuracy of the SVS to determine RBC content varies dramatically among MR imaging machines.[8] Because T2* maps are commonly used to assess the cardiac and liver iron load in patients with thalassemia,[9,10] measuring the thrombus T2* relaxation time (TT2*RT) could provide quantitative insight into thrombus composition rather than a qualitative evaluation using the SVS.[11] In this study, we aimed to analyze the RBC content of in vitro thrombi in relation to the thrombus-T2* relaxation time and to examine the intra- and interobserver agreement of the TT2*RT measurements.

MATERIALS AND METHODS

Blood Collection and Thrombus Formation

Detailed methods for blood collection and thrombus formation were previously published.[12] In brief, ovine blood was obtained from a licensed facility (Ash Stream, Hollymount, Ireland). Seven different thrombi analog types were created, with 5 replicate thrombi for each type. To separate the whole-blood constituents, we performed centrifugation, and coagulation was generally initiated following the addition of a calcium chloride solution to the blood components. Erythrocyte, plasma, and white cell content were controlled by volume in the precoagulated solution. Finally, thrombi were prepared with different ratios of RBCs and fibrin. Clotted material was allowed to mature for approximately 30 minutes to 1 hour at 37°C. Then, thrombi were formed in either a static or dynamic environment and mixed with citrated plasma, thrombin, or platelets to promote thrombus retraction. Of note, the thrombus-creation results show that the environment influences thrombus composition. Specifically, even with the same initial volume and constituents of blood, a thrombus created in static conditions had a vastly different final histologic composition than the dynamic thrombus. To simulate realistic blood conditions and represent the wide range of human thrombus compositions, we used both types of environmental conditions to create thrombi. They were suspended in a physiologic serum within an MR imaging–compatible agarose matrix (Fig 1*D*).

MR Imaging Analysis

Each thrombus was placed in a vial 4 cm deep and 8 mm wide, filled with physiologic serum. The thrombi were fully submerged at the bottom of the vials but were not embedded in the agarose gel. Thirty-five different thrombi were scanned simultaneously in an MR imaging machine. Experiments were performed on a 1.5T MR imaging unit (Magnetom Aera; Siemens, Erlangen, Germany) equipped with a system capable of 50 mT/m-1 maximum strength and 200 mT-1.s-1 maximum slew rate. A volume coil was used for excitation, while a head/neck coil composed of 20 receiver channels was used for signal reception.

A quantitative T2* mapping sequence based on multi-gradient-echo was used with the following imaging parameters: TR, 442 ms; 12 different TEs: 3.30, 4.72, 6.19, 9.12, 13.08, 17.04, 20.98, 24.92, 28.86, 32.80, 36.74, 40.68 ms; slice thickness, 3 mm; gap between slices, 0.3 mm; FOV, 260 × 260 mm; flip angle, 20°; matrix, 205 × 256; voxel size, 1.3 × 1.3 × 3 mm^3; pass band, 930 Hz; acquisition time, 2 minutes and 32 seconds. We performed a 3D-CISS sequence with the following parameters: TR, 6.09 ms; TE, 2.81; section thickness, 0.5 mm; intersection gap, 0 mm; FOV, 190 × 154 mm; flip angle, 62°; matrix, 320 × 240; voxel size, 0.3 × 0.3 × 0.3; pass band, 300 Hz; acquisition time, 2.49 minutes.

The TT2*RT values were measured after coregistration with 3D-CISS images to visualize the boundaries of the thrombi and manually select an ROI to include the entire thrombus on an axial slice. The software automatically provided a mean TT2*RT value in milliseconds (in the ROI), the SD, and the size of the area in square millimeters.

After the MR imaging acquisition, the thrombus analogs were immediately embedded in paraffin for histologic analysis. The measurements by 2 radiologists (a junior radiologist and an experienced neuroradiologist) blinded to the histologic composition of the thrombi were performed in 2 sessions 2 weeks apart to examine inter- and intraobserver reliability.

Histologic Analysis

Immediately after the MR imaging acquisition, the thrombi were fixed in a 10% formalin solution. Formalin-fixed specimens were then embedded in paraffin, cut into 7-μm-thick sections, and stained with hematoxylin-eosin. Two experienced pathologists without knowledge of the laboratory-announced thrombus composition and the imaging findings per-

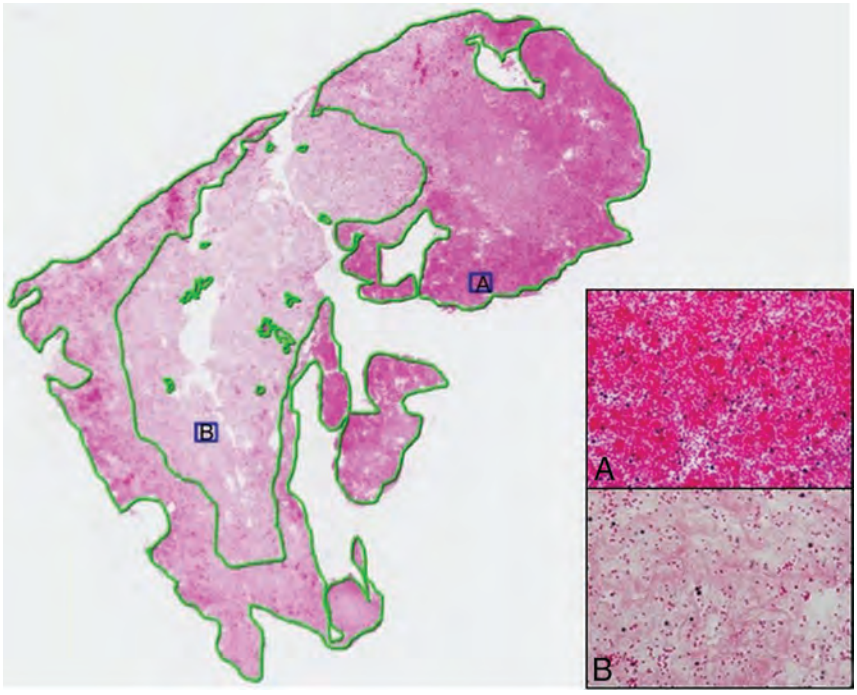


FIG 2. This thrombus consists of RBC (*A*), fibrin/platelet aggregations (*B*), and very rare white bloods cells. The RBC-dominant areas were identified on microscopic examination; then, a *green line* was drawn to select each of them. The proportion of the RBC component was calculated as the ratio of selected dominant-RBC areas to the entire thrombus area (hematoxylin-eosin; original magnification, G×20; G×400).

formed the histologic evaluation independently. Discordant results were re-evaluated and were resolved by consensus. We identified the different components of the thrombi: fibrin/platelet aggregations, RBCs, and rarely found white blood cells. Slides were scanned and analyzed using NDP.view software (https://ndp-view.software.informer.com/). First, a low-power-field examination was used to identify pure fibrin clots (>95% fibrin), pure RBC clots (>95% RBCs), and heterogeneous clots. Preliminary evidence suggested that 2 pathologists working independently could reproducibly evaluate a 10% change in RBC content. Hence, we devised a simple scaling system (Fig 2) based on a 10% threshold that had a stepwise incremental increase in RBC content for this semiquantitative analysis. When the thrombus was heterogeneous, a 2-step procedure was devised. First, the RBC-dominant areas were evaluated by a semiquantitative analysis using a lattice composed of 10 × 10 squares at a high-power field. Second, the proportion of the RBC component was calculated as the ratio (percentage) of selected dominant RBC areas to the entire thrombus area. A 10% incremental scale was again applied.

Statistical Analysis

MR imaging quantitative T2* mapping results are expressed as median, range, and interquartile range due to the non-normal distribution. Normality of distributions was assessed using histograms and the Shapiro-Wilk test. Inter- and intraobserver reproducibility of the TT2*RT measurements was assessed by calculating intraclass correlation coefficients. The intraclass correlation coefficients were estimated using a 2-way random effects model (absolute agreement). Intraclass correlation coefficient values of <0.5, between 0.5 and 0.75, between 0.75 and 0.9, and >0.90 were interpreted as poor, moderate, good, and excellent reproducibility, respectively.[13] The mean of the TT2*RT values (4 readings) was used in descriptive and correlation analyses. The correlation between TT2*RT values and histologic RBC content was assessed by calculating the Spearman rank correlation (ρ) coefficient. Ninety-five percent confidence intervals for intraclass correlation coefficients and ρ values were obtained using bootstrap methods (2000 bootstrap samples). All analyses were computed using R statistical and computing software (Version 3.5.1; http://www.r-project.org/).

RESULTS

A total of 35 thrombi were histologically processed and analyzed. Inter- and intraobserver reproducibility of the TT2*RT measurements was excellent, with an intraclass correlation coefficient of 0.915 (95% CI, 0.783–0.981) for interobserver reproducibility at the first session, 0.944 (95% CI, 0.814–0.991) for interobserver reproducibility at the second session, 0.977 (95% CI, 0.961–0.988) for intraobserver reproducibility of the junior radiologist, and 0.982 (95% CI, 0.967–0.990) for intraobserver reproducibility of the senior neuroradiologist.

The median TT2*RT value was 78.5 ms (range, 16–268 ms; interquartile range, 60.5 ms). The median TT2*RT SD was 28.5 ms (range, 6–84 ms; interquartile range, 25.5 ms). The median TT2*RT area that reflects the cross-sectional size of the thrombi was 26 mm^2 (range, 17–67 mm^2; interquartile range, 23.3 mm^2). The median RBC content was 55% (range, 0%–100%; interquartile range, 75%).

A strong negative correlation between TT2*RT and RBC content was found, with a correlation coefficient of −0.834 (95% CI, −0.659 to −0.915) (Fig 3). For fibrin-dominant thrombi, the median TT2*RT value was 104.25 ms (range, 63–230 ms; interquartile range, 61.9 ms), whereas for the RBC thrombi, the median TT2*RT value was 54 ms (range, 20–150 ms; interquartile range, 45.9 ms) (Fig 4).

As shown in Fig 3, the scatterplot of RBC content against TT2*RT revealed an outlier (90% RBC content, TT2*RT of 150 ms) with a substantial interobserver discordance (reader 1: TT2*RT of 80 ± 25 ms; reader 2: TT2*RT of 153 ± 163.5 ms), representing a 63% variation between observers. Histologic analysis indicated a 90% RBC content with a central area mostly composed of RBCs. However, a large fibrin-rich halo surrounded the thrombus. This unusual "2-layered" thrombus created discordance and misinterpretation due to inconsistently measured areas in either the RBC or fibrin-rich layers (Fig 1*C*).

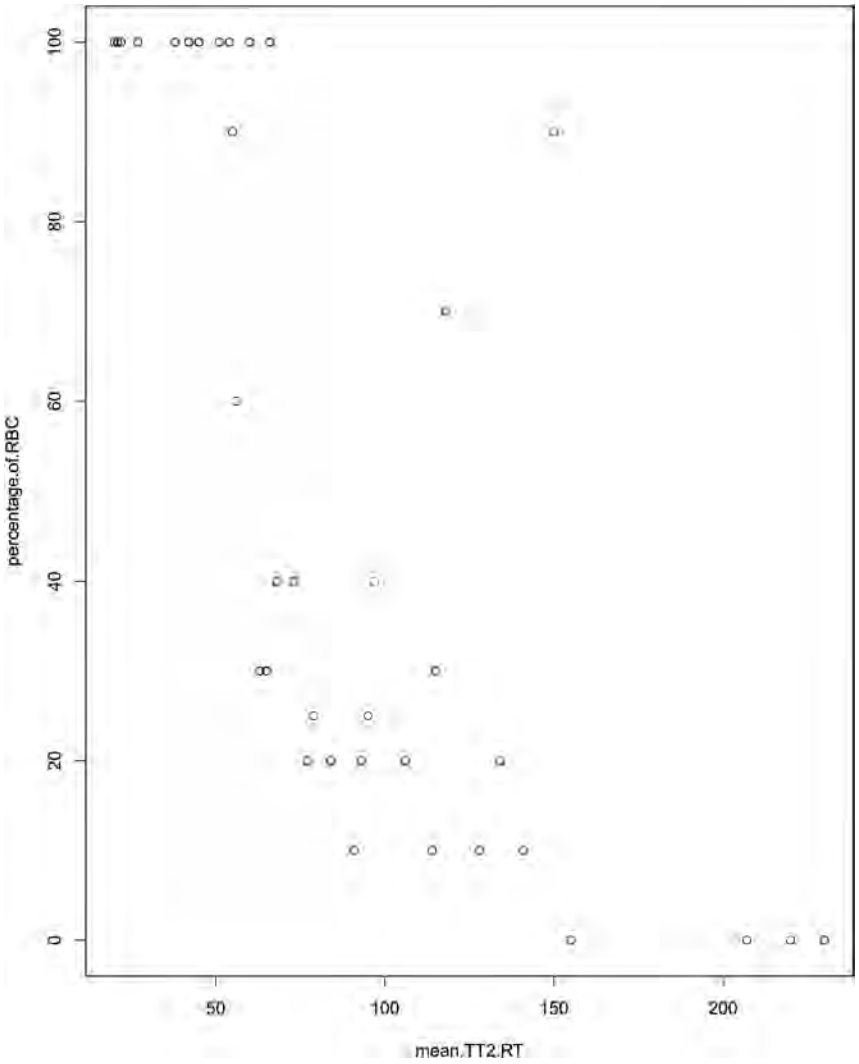


FIG 3. Semiquantitative evaluation (10% steps) of the red blood cell content in the thrombi and corresponding values of the mean thrombus-T2* relaxation time in milliseconds.

DISCUSSION

We found an inversely proportional correlation between the thrombus RBC content measured by histology and the TT2*RT value using MR imaging with in vitro thrombi. This new and short MR imaging sequence requires <3 minutes for acquisition and has excellent intra- and interobserver agreement, providing a reliable quantitative biomarker for dominant thrombus composition detectible in routine clinical practice.

Randomized controlled trials have demonstrated the superiority of bridging therapy over best medical care for patients with confirmed large-vessel occlusion but without providing any information on thrombus composition.[14-18] However, studies on human thrombi retrieved from patients with anterior acute ischemic stroke have revealed varying compositions.[5,6] Because the 2 major components of thrombi are RBC and fibrin, the fibrin proportion is roughly the inverse of the RBC content[5,6,19] (eg, if the RBC content is 20%, the fibrin content will be around 80%, and conversely). When one engages a thrombus with a stent retriever or aspiration catheter, optimal outcomes are likely best achieved when the entire thrombus can be retrieved in 1 pass without losing any fragments. A fibrin-rich thrombus can be difficult to engage and may adhere to the vessel wall; on the other hand, a weakly organized RBC-rich thrombus can be prone to fragmentation.[20,21]

Thus, MR imaging can offer a quantitative approach to measuring thrombus composition. The SVS has been studied in the literature on ischemic stroke treated by intravenous thrombolysis or endovascular means with end points defined either in terms of recanalization or functional outcome.[11,22-25] However, the qualitative evaluation of the SVS (either + or −) limits its ability to assess the complex and variable composition of thrombi. As previously demonstrated, a linear variation of the RBC/fibrin ratio is observed in samples of retrieved thrombi.[3,5,6] Furthermore, the diagnostic accuracy of the SVS to predict thrombus composition varies according to the MR imaging machine.[8] This variation seems to be a crucial pitfall for any extensive use of the SVS in large multicenter studies. The appearance of the SVS is related to technical parameters such as TE, voxel size, and magnetic field.[26] Given these limitations, the SVS should be read in relation to a precise T2* map calculated using several TRs with highly controlled imaging parameters. Indeed, a gradient-echo sequence is commonly used to identify brain hemorrhage, but the parameters of this sequence are not optimized for the detection and characterization of a thrombus. In this study, we acquired 12 different echo sequences, from 3.30 to 40.68 ms, allowing us to precisely measure T2* on tissue with both high and low T2*.

Hence, a quantitative imaging method might be a useful way of analyzing thrombus composition more precisely. A study found that acute intramural hematoma due to cerebral artery dissection could be determined more accurately on T2* quantification mapping.[27] One recent study analyzed the performance of T2* mapping (compared with SVS) to detect thrombi.[28] The TT2*RT value has also been found to be related to the number of attempts to reach successful recanalization after endovascular treatment, but the correlation between TT2*RT and the histologic composition of the thrombus has never been analyzed.[11] In our study, we used distinct thrombus analogs varying in composition, ranging from RBC-dominant to fibrin-dominant.[12] This model of thrombus creation holds promise for translational research on thrombus composition and permeability.[29] Using in vitro thrombi provides the opportunity to validate the capacity of the sequence to predict the proportion of RBCs. Thus, in vitro thrombi are more relevant than ex vivo thrombi for 2 main reasons: First, our study required thrombi with a composition ranging from extremely RBC-rich to fibrin-only, corresponding to the

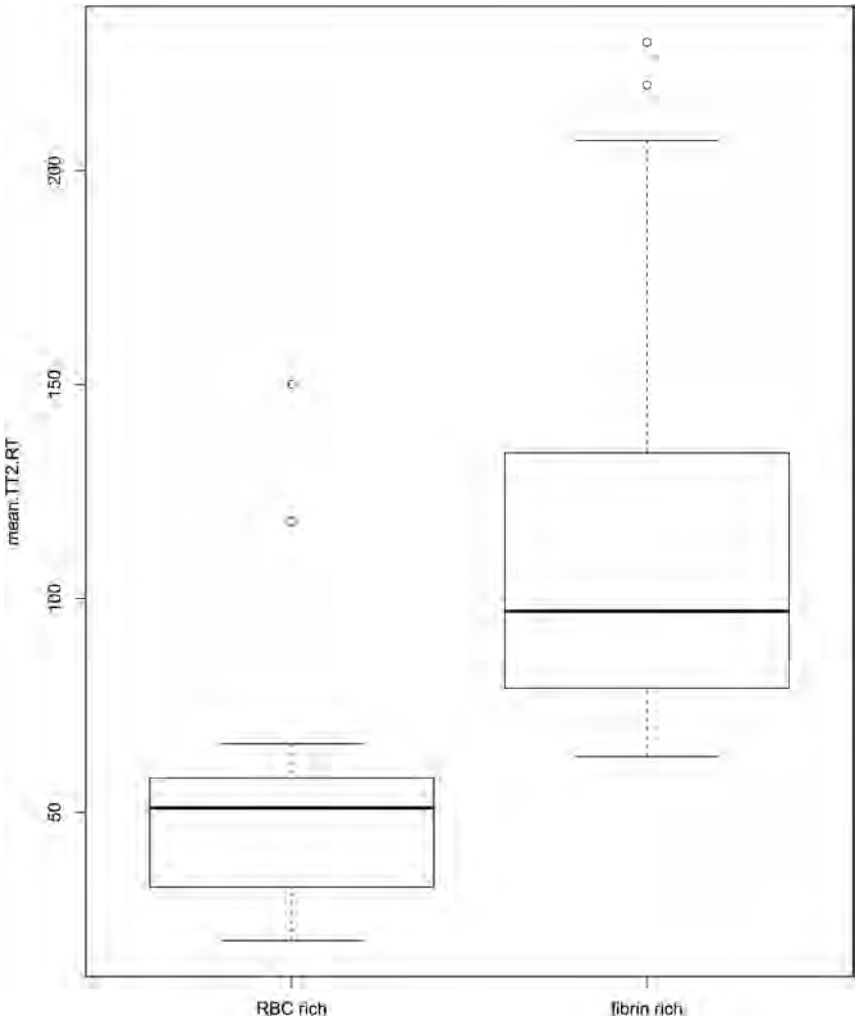


FIG 4. Boxplot of the red blood cell–dominant thrombi and of the fibrin-dominant thrombi with corresponding values of the mean thrombus-T2* relaxation time in milliseconds.

wide range of thrombi encountered in clinical practice, some of which can be very difficult to retrieve. In particular, fibrin-rich thrombi are difficult to remove with traditional stent-retriever methods; hence, there is a need to predict these types of thrombi. Second, the use of intravenous thrombolysis before thrombus extraction can modify the thrombus characteristics. In particular, thrombi retrieved after substantial chemical and mechanical manipulations are likely to be vastly dissimilar to the original occlusion.

Our study has several limitations. First, the results of TT2*RT might depend on the magnetic field homogeneity and susceptibility differences between the thrombus and the medium surrounding it. To test the influence of these parameters, one would need to repeat the measurements with thrombi positioned in different orientations. This was not done in our study; the thrombi were randomly placed within their vials. As a consequence, during the MR imaging acquisition, the long axis of the thrombi had multiple directions either in the direction of the phases or the frequency, just like the blood vessels in vivo. Moreover, to limit the effect of heterogeneities on images, we performed a 3D shim before each acquisition, allowing a narrow water-frequency peak. This adjustment only takes 30 seconds, is fully compatible with use in routine clinical practice, and allows the robustness and precision of the T2* value to be substantially improved.

Then, a critical point is that the 1.5T Siemens unit offers the possibility of shortening or lengthening the interecho intervals. Thus, in the sequence presented here, we concentrated the echoes around 30 ms for optimal accuracy. However, the T2* value is also dependent on magnetic field strength, and the next step is to analyze the inter-MR imaging machine agreement for this sequence. Last, one can argue that susceptibility-weighted imaging sequences could be useful for detecting intracranial thrombi. This is most likely the case because these sequences are more sensitive to the RBC content, but they may not efficiently predict the proportion of RBCs in the thrombus. Furthermore, because of the longer TEs, the susceptibility-weighted imaging requires a longer acquisition time, which can be detrimental in the context of acute ischemic stroke.

CONCLUSIONS

An MR imaging quantitative T2* mapping sequence that records the TT2*RT can reliably identify the thrombus RBC content in vitro. TT2*RT values are inversely related to the RBC content of thrombi. This new short MR imaging sequence provides a quantitative biomarker of thrombus composition accessible in clinical routine. Normalization of the MR imaging quantitative T2* mapping sequence between MR imaging vendors should be considered in the future if large multicenter studies aim to include information on thrombus composition before mechanical thrombectomy.

ACKNOWLEDGMENTS

We thank Siemens and Neuravi for their nonfunded, scientific support and Nick Barton for language editing.

REFERENCES

1. 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
2. Hashimoto T, Hayakawa M, Funatsu N, et al. **Histopathologic analysis of retrieved thrombi associated with successful reperfusion after acute stroke thrombectomy.** *Stroke* 2016;47:3035–37 CrossRef Medline
3. Brinjikji W, Duffy S, Burrows A, et al. **Correlation of imaging and histopathology of thrombi in acute ischemic stroke with etiology and outcome: a systematic review.** *J Neurointerv Surg* 2017;9:529–34 CrossRef Medline
4. Rovira A, Orellana P, Alvarez-Sabín J, et al. **Hyperacute ischemic stroke: middle cerebral artery susceptibility sign at echo-planar gradient-echo MR imaging.** *Radiology* 2004;232:466–73 CrossRef Medline

5. Liebeskind DS, Sanossian N, Yong WH, et al. **CT and MRI early vessel signs reflect clot composition in acute stroke.** *Stroke* 2011; 42:1237–43 CrossRef Medline
6. Marder VJ, Chute DJ, Starkman S, et al. **Analysis of thrombi retrieved from cerebral arteries of patients with acute ischemic stroke.** *Stroke* 2006;37:2086–93 CrossRef Medline
7. Boeckh-Behrens T, Schubert M, Förschler A, et al. **The impact of histological clot composition in embolic stroke.** *Clin Neuroradiol* 2016;26:189–97 CrossRef Medline
8. Bourcier R, Détraz L, Serfaty JM, et al. **MRI interscanner agreement of the association between the susceptibility vessel sign and histologic composition of thrombi.** *J Neuroimaging* 2017;27:577–82 CrossRef Medline
9. Hanneman K, Nguyen ET, Thavendiranathan P, et al. **Quantification of myocardial extracellular volume fraction with cardiac MR imaging in thalassemia major.** *Radiology* 2016;279:720–30 CrossRef Medline
10. Gandon Y, Olivié D, Guyader D, et al. **Non-invasive assessment of hepatic iron stores by MRI.** *Lancet* 2004;363:357–62 CrossRef Medline
11. Bourcier R, Brecheteau N, Costalat V, et al. **MRI quantitative T2* mapping on thrombus to predict recanalization after endovascular treatment for acute anterior ischemic stroke.** *J Neuroradiol* 2017;44: 241–46 CrossRef Medline
12. Duffy S, Farrell M, McArdle K, et al. **Novel methodology to replicate clot analogs with diverse composition in acute ischemic stroke.** *J Neurointerv Surg* 2017;9:486–91 CrossRef Medline
13. 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
14. 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
15. Campbell BC, Mitchell PJ, Kleinig TJ, et al; EXTEND-IA Investigators. **Endovascular therapy for ischemic stroke with perfusion-imaging selection.** *N Engl J Med* 2015;372:1009–18 CrossRef Medline
16. Berkhemer OA, Fransen PS, Beumer D, et al. **A randomized trial of intraarterial treatment for acute ischemic stroke.** *N Engl J Med* 2015; 372:11–20 CrossRef Medline
17. Jovin TG, Chamorro A, Cobo E, et al; REVASCAT Trial Investigators. **Thrombectomy within 8 hours after symptom onset in ischemic stroke.** *N Engl J Med* 2015;372:2296–306 CrossRef Medline
18. Saver JL, Goyal M, Bonafe A, et al; SWIFT PRIME Investigators. **Stent-retriever thrombectomy after intravenous t-PA vs. t-PA alone in stroke.** *N Engl J Med* 2015;372:2285–95 CrossRef Medline
19. Boeckh-Behrens T, Kleine JF, Zimmer C, et al. **Thrombus histology suggests cardioembolic cause in cryptogenic stroke.** *Stroke* 2016;47: 1864–71 CrossRef Medline
20. Fennell VS, Setlur Nagesh SV, Meess KM, et al. **What to do about fibrin rich "tough clots"? Comparing the Solitaire stent retriever with a novel geometric clot extractor in an in vitro stroke model.** *J Neurointerv Surg* 2018;10:907–10 CrossRef Medline
21. Machi P, Jourdan F, Ambard D, et al. **Experimental evaluation of stent retrievers' mechanical properties and effectiveness.** *J Neurointerv Surg* 2017;9:257–63 CrossRef Medline
22. Soize S, Batista AL, Rodriguez Regent C, et al. **Susceptibility vessel sign on T2* magnetic resonance imaging and recanalization results of mechanical thrombectomy with stent retrievers: a multicentre cohort study.** *Eur J Neurol* 2015;22:967–72 CrossRef Medline
23. Naggara O, Raymond J, Domingo Ayllon M, et al. **T2* "susceptibility vessel sign" demonstrates clot location and length in acute ischemic stroke.** *PLoS One* 2013;8:e76727 CrossRef Medline
24. Kimura K, Sakamoto Y, Iguchi Y, et al. **Clinical and MRI scale to predict very poor outcome in tissue plasminogen activator patients.** *Eur Neurol* 2011;65:291–95 CrossRef Medline
25. Aoki J, Kimura K, Shibazaki K, et al. **Location of the susceptibility vessel sign on T2*-weighted MRI and early recanalization within 1 hour after tissue plasminogen activator administration.** *Cerebrovasc Dis Extra* 2013;3:111–20 CrossRef Medline
26. Hodel J, Rodallec M, Gerber S, et al. **Susceptibility weighted magnetic resonance sequences "SWAN, SWI and VenoBOLD": technical aspects and clinical applications** [in French]. *J Neuroradiol* 2012; 39:71–86 CrossRef Medline
27. Kato A, Shinohara Y, Yamashita E, et al. **Usefulness of R2* maps generated by iterative decomposition of water and fat with echo asymmetry and least-squares estimation quantitation sequence for cerebral artery dissection.** *Neuroradiology* 2015;57:909–15 CrossRef Medline
28. Shinohara Y, Kato A, Yamashita E, et al. **R2* map by IDEAL IQ for acute cerebral infarction: compared with susceptibility vessel sign on T2*-weighted imaging.** *Yonago Acta Med* 2016;59:204–09 Medline
29. Santos EM, Marquering HA, den Blanken MD, et al; MR CLEAN Investigators. **Thrombus permeability is associated with improved functional outcome and recanalization in patients with ischemic stroke.** *Stroke* 2016;47:732–41 CrossRef Medline

Two-Layered Susceptibility Vessel Sign and High Overestimation Ratio on MRI Are Predictive of Cardioembolic Stroke

R. Bourcier, I. Derraz, S. Bracard, C. Oppenheim, and O. Naggara, on behalf of the THRACE Investigators

ABSTRACT

SUMMARY: In a prospective study among patients in the THRombectomie des Artères Cerebrales trial, we analyzed the diagnostic accuracy of 2 imaging biomarkers, the 2-layered susceptibility vessel sign and a high overestimation ratio, obtained on pretreatment brain T2* sequences, to identify cardioembolic stroke etiology (107/260 patients). In combination, these 2 biomarkers, on 1.5T or 3T systems (159 and 101 patients, respectively), demonstrated high specificity (0.77 at 1.5T and 1 at 3T) and their simultaneous presence is highly associated with cardioembolism.

ABBREVIATIONS: CES = cardioembolic stroke; HOR = high overestimation ratio; NPV = negative predictive value; overR = overestimation ratio; PPV = positive predictive value; SVS = susceptibility vessel sign; THRACE = THRombectomie des Artères Cerebrales; TL-SVS = 2-layered susceptibility vessel sign

Unraveling the etiology in acute ischemic stroke caused by large-vessel occlusion is important for secondary stroke prevention strategies. However, in many patients, risk factors for both large-artery atherosclerotic and cardioembolic stroke (CES) etiology are present. The composition of thrombus in obstructed arteries varies depending on whether the embolic source is cardioembolism or large-artery atherosclerosis.[1,2] On admission brain MR imaging, the susceptibility vessel sign (SVS), defined as a hypointense signal exceeding the diameter of the contralateral artery at the thrombus site,[3] may be seen on T2*-weighted images. The SVS was reported to be associated with cardioembolism, but with conflicting results and low specificity.[4-8] Recently, 2 characteristics of the SVS, overestimation ratio (overR), a quantitative evaluation of the SVS, and the 2-layered susceptibility vessel sign (TL-SVS), were reported to predict CES.[7,8] Indeed, both the TL-SVS and high overR (HOR) were described as exhibiting almost perfect specificity for CES.[7,8] However, these studies were retrospective, monocentric, and used time-consuming susceptibility-weighted imaging or 3T systems, and the TL-SVS and HOR were not both evaluated in individual patients. With >200 patients included with a gradient-echo sequence at either 1.5T or 3T, the multicenter THRombectomie des Artères Cerebrales (THRACE) trial[9] offers an opportunity to analyze the diagnostic accuracy of these 2 imaging biomarkers for the prediction of CES.

MATERIALS AND METHODS

THRACE was a randomized controlled trial conducted in 26 centers in France. Study design and protocol are detailed elsewhere.[10] Patients with acute ischemic stroke were eligible for inclusion if they were 18–80 years of age, had an NIHSS score of 10–25, had an occlusion of the intracranial ICA or the M1 segment of the MCA on MRA, could be administered intravenous thrombolysis within 4 hours of symptom onset, and could undergo thrombectomy within 5 hours of symptom onset. Patients with cervical ICA occlusion or subocclusive stenosis were excluded. Written informed consent was obtained from all patients or their legal representatives. The study protocol was approved by the Comité de Protection des Personnes III Nord Est Ethics Committee and the research boards of the participating centers.

For the present study, we included patients screened with an MR imaging gradient-echo sequence for whom analysis of the SVS was available. We excluded patients with severe motion artifacts on gradient-echo sequences when the presence or absence of SVS could not be reliably determined. We also excluded 5 patients screened with susceptibility-weighted imaging instead of a gradient-echo sequence. MR imaging acquisition parameters of the sequence were left to the discretion of the recruitment centers

Received June 7, 2018; accepted after revision September 5.

From the Department of Neuroradiology (R.B.), University Hospital Guillaume et René Laennec, Nantes, France; Department of Neuroradiology (R.B.), University Hospital of Nancy, Nancy, France; Department of Neuroradiology and Institut National de la Santé et de la Recherche Médicale CIC 1433 Innovation Technologique (I.D., S.B.), Université Paris-Descartes, Paris, France; Institut National de la Santé et de la Recherche Médicale U894 (C.O., O.N.), Sainte-Anne Hospital, Paris, France; and Pediatric Radiology Department (O.N.), Necker Enfants Malades, Paris, France.

Catherine Oppenheim and Olivier Naggara contributed equally to this work.

Please address correspondence to Romain Bourcier, MD, PhD, Service de Neuroradiologie, Centre hospitalier universitaire Guillaume et René Laennec, Boulevard Jacques Monod, 44 093 Nantes, France; e-mail: romain.bourcier2@gmail.com

Indicates article with supplemental on-line appendix.

Indicates article with supplemental on-line photo.

http://dx.doi.org/10.3174/ajnr.A5865

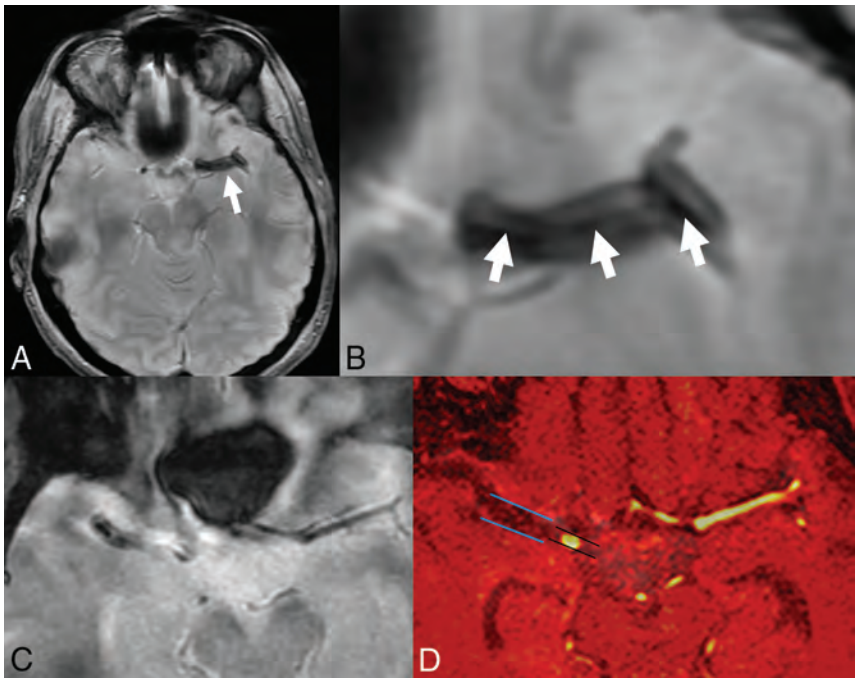


FIGURE. *A* and *B*, Two-layered susceptibility vessel sign. A 62-year-old man with left middle cerebral artery occlusion because of cardioembolism. 1.5T T2*-weighted gradient recalled-echo imaging (*white arrow*) shows a 2-layered SVS with a low-intensity signal core (*arrowheads*) surrounded by a higher intensity signal. *C* and *D*, Susceptibility vessel sign and overestimation ratio of SVS. A 69-year-old man with right middle cerebral artery occlusion. Overestimation ratio of SVS was 2.21 and the etiological work-up concluded to a cardiogenic embolism. The sequences of MR imaging from *C*, T2*-weighted gradient-echo images. *D*, The coregistered images of TOF-MRA and T2*-gradient-echo.

according to their routine practice and were therefore not standardized.

Readers searched for the TL-SVS, defined as an inhomogeneous SVS with a linear low-intensity signal core surrounded by a peripheral higher intensity signal.[8] The overR was manually measured according to the method previously described (ie, overR = Width of SVS / Width of Large Artery).[7] The ipsilateral large-artery width proximal to the occlusion was used to calculate the overR. Note that absence of the SVS was considered as an overR = 1 (Figure).

Stroke etiology was determined at the end of a complete etiologic work-up by a stroke neurologist, according to the Trial of Org 10172 in Acute Stroke Treatment classification.[11] In the THRACE trial, the recommended general work-up, according to the European Stroke Organisation Guidelines,[10] included physiologic parameters and routine blood tests, a 12-lead electrocardiogram, and continuous electrocardiography. In addition, 24-hour Holter electrocardiography monitoring was performed when arrhythmias were suspected and no other causes of stroke were found. The echocardiography was recommended when electrocardiographic abnormalities or evidence of cardiac disease on history was reported or when no other identifiable causes of stroke were found.

For details of the statistical analysis, see the On-line Appendix.

RESULTS

Two hundred sixty patients, including 107 (41%) patients with CES (31/101 patients with 3T MR imaging and 76/159 patients with 1.5T MR imaging), were identified. Furthermore, the etiologic work-up found large-artery atherosclerosis in 40 (15%) patients, "other determined cause of stroke" in 16 (6%) patients, and "undetermined cause of stroke" in 97 (37%) patients.

Overall, the median overR was 1.59 (range, 1.0–2.9; interquartile range. 0.29). The median overR was 1.54 (range, 1.0–2.9; interquartile range, 0.36) and 1.60 (range, 1.0–2.7; interquartile range, 0.54) at 3T and 1.5T, respectively.

The TL-SVS was present in 94 (36%) patients (overall 43/101 patients screened at 3T and 43/80 patients with the SVS screened at 3T; overall 51/159 patients screened at 1.5T and 51/136 patients with the SVS screened at 1.5T).

At 1.5T, the association of the TL-SVS and HOR yielded a specificity of 0.77 and a positive predictive value of 0.60.

At 3T, the association of the TL-SVS and HOR yielded a specificity of 1, a negative predictive value of 0.74, and a positive predictive value of 1 (Tables 1–3).

DISCUSSION

In the present study, using 3T systems, the simultaneous presence of the TL-SVS and HOR offered perfect specificity and a positive predictive value for CES. Using 1.5T systems, we also showed that the simultaneous presence of the TL-SVS and HOR had 77% specificity for CES.

Because exact stroke etiology cannot be determined with a high degree of reliability in >30% of strokes,[12] a direct relationship between imaging characteristics and stroke etiology would be useful. We demonstrated that in real world conditions (ie, brain MR imaging performed in 26 different centers on either a 1.5 or a 3T MR imaging unit, using a gradient-echo sequence of <1 minute), the TL-SVS and HOR can have practical implications for determining stroke etiology at the acute phase.

Recently, from the THRACE trial, qualitative visual binary grading of the SVS offered a high sensitivity of the SVS to predict a stroke of cardioembolic origin (0.89; 95% CI, 0.83–0.95), but specificity was very low (0.21; 95% CI, 0.14–0.27).[13] In the present study, we demonstrate that adding a more precise analysis of SVS characteristics (ie, TL-SVS and overR), increased specificity. This was achieved at the expense of lowering the sensitivity, as usually occurs when using a combination of biomarkers rather than a single sign. However, the combined analysis of the TL-SVS and HOR with cutoff values specific to the MR imaging system field strength increased specificity values.

Two recent retrospective studies using 3T systems analyzed separately the diagnostic accuracy of the TL-SVS and HOR for

Table 1: Diagnostic performance of the high overestimation ratio for the prediction of a cardioembolic etiology of stroke

3T (n = 101) overR >2.25		1.5T (n = 159) overR >1.33	
Sensitivity	0.10	Sensitivity	0.86
Specificity	0.94	Specificity	0.37
PPV	0.43	PPV	0.56
NPV	0.70	NPV	0.74
Accuracy	0.68	Accuracy	0.60

Table 2: Diagnostic performance of the 2-layered susceptibility vessel sign for the prediction of a cardioembolic etiology of stroke

3T (n = 101) TL-SVS		1.5T (n = 159) TL-SVS	
Sensitivity	0.45	Sensitivity	0.39
Specificity	0.59	Specificity	0.75
PPV	0.33	PPV	0.59
NPV	0.71	NPV	0.57
Accuracy	0.54	Accuracy	0.58

Table 3: Diagnostic performance of the combined 2-layered susceptibility vessel sign and the high overestimation ratio for the prediction of a cardioembolic etiology of stroke

3T (n = 101) overR >2.25 and TL-SVS		1.5T (n = 159) overR >1.33 and TL-SVS	
Sensitivity	0.23	Sensitivity	0.38
Specificity	1.00	Specificity	0.77
PPV	1.00	PPV	0.60
NPV	0.74	NPV	0.58
Accuracy	0.76	Accuracy	0.58

CES prediction. The authors found specificities of 97%[8] and 91%,[7] respectively, for the TL-SVS and overR of >2.01 to predict a CES.[7] In our study, by means of a gradient-echo sequence rather than susceptibility-weighted imaging, the specificity of the TL-SVS and HOR both appeared lower on 1.5T and 3T systems.

The present prospective study is the largest to date that has evaluated the TL-SVS and overR in the context of stroke etiology. Our study has several limitations. First, THRACE was not a trial designed to specifically investigate stroke etiology. However, work-ups were performed in referral comprehensive stroke centers that ensure high quality and extensive etiologic work-ups in accordance with the current guidelines.[10] Second, no information was registered on the anticoagulation regimen, the use of antiplatelet agents, or the platelet count before imaging, all factors that may have influenced the SVS. Third, the diagnostic accuracy of the SVS to predict thrombus composition varies according the MR imaging machines,[14] and there was no standardization of gradient-echo sequence parameters among the 26 participating centers in the THRACE trial. In our opinion, this limitation enhances the generalizability of our findings.

CONCLUSIONS

The presence of both the TL-SVS and HOR offers a high specificity for CES etiology on both 1.5 and 3T systems. When a patient with stroke has both of these imaging biomarkers, a cardioembolic source is highly likely.

ACKNOWLEDGMENTS

We thank Marc Soudant from the Institut National de la Santé CIC 1433 Epidémiologie Clinique for conducting the statistical analysis and Nick Barton for language editing.

REFERENCES

1. Marder VJ, Chute DJ, Starkman S, et al. **Analysis of thrombi retrieved from cerebral arteries of patients with acute ischemic stroke.** *Stroke* 2006;37:2086–93 CrossRef Medline
2. Boeckh-Behrens T, Kleine JF, Zimmer C, et al. **Thrombus histology suggests cardioembolic cause in cryptogenic stroke.** *Stroke* 2016;47:1864–71 CrossRef Medline
3. Rovira A, Orellana P, Alvarez-Sabín J, et al. **Hyperacute ischemic stroke: middle cerebral artery susceptibility sign at echo-planar gradient-echo MR imaging.** *Radiology* 2004;232:466–73 CrossRef Medline
4. Brinjikji W, Duffy S, Burrows A, et al. **Correlation of imaging and histopathology of thrombi in acute ischemic stroke with etiology and outcome: a systematic review.** *J Neurointerv Surg* 2017;9:529–34 CrossRef Medline
5. Cho KH, Kim JS, Kwon SU, et al. **Significance of susceptibility vessel sign on T2*-weighted gradient echo imaging for identification of stroke subtypes.** *Stroke* 2005;36:2379–83 CrossRef Medline
6. Kang DW, Jeong HG, Kim DY, et al. **Prediction of stroke subtype and recanalization using susceptibility vessel sign on susceptibility-weighted magnetic resonance imaging.** *Stroke* 2017;48:1554–59 CrossRef Medline
7. Zhang R, Zhou Y, Liu C, et al. **Overestimation of susceptibility vessel sign: a predictive marker of stroke cause.** *Stroke* 2017;48:1993–96 CrossRef Medline
8. Yamamoto N, Satomi J, Tada Y, et al. **Two-layered susceptibility vessel sign on 3-Tesla T2*-weighted imaging is a predictive biomarker of stroke subtype.** *Stroke* 2015;46:269–71 CrossRef Medline
9. Bracard S, Ducrocq X, Mas JL, et al; THRACE investigators. **Mechanical thrombectomy after intravenous alteplase versus alteplase alone after stroke (THRACE): a randomised controlled trial.** *Lancet Neurol* 2016;15:1138–47 CrossRef Medline
10. European Stroke Organisation (ESO) Executive Committee, ESO Writing Committee. **Guidelines for management of ischaemic stroke and transient ischaemic attack 2008.** *Cerebrovasc Dis* 2008;25:457–507 CrossRef Medline
11. Adams HP Jr, Bendixen BH, Kappelle LJ, et al. **Classification of subtype of acute ischemic stroke: definitions for use in a multicenter clinical trial—TOAST. Trial of Org 10172 in Acute Stroke Treatment.** *Stroke* 1993;24:35–41 CrossRef Medline
12. Goldstein LB, Jones MR, Matchar DB, et al. **Improving the reliability of stroke subgroup classification using the Trial of ORG 10172 in Acute Stroke Treatment (TOAST) criteria.** *Stroke* 2001;32:1091–98 CrossRef Medline
13. Bourcier R, Derraz I, Delasalle B, et al; THRACE investigators. **Susceptibility vessel sign and cardioembolic etiology in the THRACE trial.** *Clin Neuroradiol* 2018 Jun 12. [Epub ahead of print] CrossRef
14. Bourcier R, Détraz L, Serfaty JM, et al. **MRI interscanner agreement of the association between the susceptibility vessel sign and histologic composition of thrombi.** *J Neuroimaging* 2017;27:577–82 CrossRef Medline

Differences in the Calculated Transvenous Pressure Drop between Chronic Hydrocephalus and Idiopathic Intracranial Hypertension

G.A. Bateman and A.R. Bateman

ABSTRACT

BACKGROUND AND PURPOSE: Chronic hydrocephalus is associated with dilated ventricles despite a normal intracranial pressure. In idiopathic intracranial hypertension, the ventricles are normal despite an elevated intracranial pressure. This apparent paradox has largely remained unexplained. It is suggested that a pressure difference between the superficial and deep venous territories of the brain could account for the variation between the 2 diseases. The purpose of this paper is to investigate the cause of this pressure difference.

MATERIALS AND METHODS: Using MR phase-contrast imaging, we calculated the hydraulic diameters of the sagittal and straight sinuses in 21 patients with hydrocephalus, 20 patients with idiopathic intracranial hypertension, and 20 age-matched controls. The outflow resistance of each sinus was estimated using the Poiseuille equation. The outflow pressure was estimated using the flow data. A smaller subset of the patients with hydrocephalus had these studies repeated after successful shunt insertion.

RESULTS: In hydrocephalus, the sagittal sinuses were 21% smaller than those in controls ($P < .001$); the straight sinuses were not significantly different. In idiopathic intracranial hypertension, both sinuses were not significantly different from those of controls. The pressure drop from the sagittal sinus to the end of the straight sinus was elevated by 1.2 mm Hg in hydrocephalus ($P = .001$) but not significantly different from that in controls in idiopathic intracranial hypertension. Shunt insertion dilated the sagittal sinuses in hydrocephalus, leaving them 18% larger than normal and eliminating the transvenous pressure change.

CONCLUSIONS: There is a transvenous pressure difference in hydrocephalus that is absent in idiopathic intracranial hypertension. This difference is eliminated by shunt insertion. The findings may have a bearing on ventricular dilation.

ABBREVIATION: IIH = idiopathic intracranial hypertension

It has been estimated that >30,000 shunt procedures are performed in the United States every year to treat hydrocephalus.[1] This is despite a lack of understanding of what causes the disease or how a shunt exactly treats the condition. In chronic hydrocephalus, there is deformation of the brain parenchyma, which requires an expenditure of energy. In a hydraulic system, energy expenditure requires a pressure difference. It is assumed that a shunt treats this pressure difference. However, previous searches for a pressure drop between the ventricles and the subarachnoid space[2] and the brain parenchyma[3] or a pulse pressure difference[4] have been unsuccessful. Nevertheless, previous work by one of the current authors suggests that a pressure difference may exist between the superficial and deep venous territories in normal pressure hydrocephalus,[5] and this could supply the energy required for ventricular dilation. In idiopathic intracranial hypertension (IIH), the ventricles do not dilate; therefore, a pressure difference should not exist in this disease.

Recently, a 25% reduction in the sagittal sinus cross-sectional area in hydrocephalus but no change in the IIH sagittal sinus cross-sectional area has been found.[6] This finding would imply that an increase in pressure difference between the torcular and midsagittal sinus could exist in hydrocephalus but not in IIH. The purpose of the current study was to measure the sagittal and straight sinus cross-sectional areas and circumferences to calcu-

Received July 5, 2018; accepted after revision October 2.

From the Department of Medical Imaging (G.A.B.), John Hunter Hospital, Newcastle, New South Wales, Australia; Newcastle University Faculty of Health (G.A.B.), Callaghan Campus Newcastle, New South Wales, Australia; and Biomedical Engineering (A.R.B.), University of New South Wales, Sydney, New South Wales, Australia.

Paper previously presented, in part, in abstract form at: Hydrocephalus 2017, Ninth Meeting of the International Society for Hydrocephalus and CSF Disorders, September 23–25, 2017; Kobe, Japan.

Please address correspondence to Grant A. Bateman, MD, Department of Medical Imaging, John Hunter Hospital, Locked Bag 1, Newcastle Region Mail Center, 2310 Australia; e-mail: grant.bateman@hnehealth.nsw.gov.au

Indicates open access to non-subscribers at www.ajnr.org
Indicates article with supplemental on-line tables.
Indicates article with supplemental on-line photo.

http://dx.doi.org/10.3174/ajnr.A5883

late the hydraulic diameter and resistance of each. These data, together with blood flow data, will allow an estimate of the pressure difference from the midsagittal sinus to the end of the straight sinus (ie, the transvenous pressure difference). Review of a subset of the patients with hydrocephalus post-shunt insertion may shed light on the therapeutic effect of this procedure.

MATERIALS AND METHODS

Subjects

In a previous study undertaken by one of the authors,[7] there were 21 patients with chronic idiopathic hydrocephalus (mean age, 45 ± 10 years; 7 women and 14 men). There were 20 controls (mean age, 44 ± 10 years, with 8 women and 12 men).[7] These patients were entered into the current study, and the clinical data for these patients can be reviewed in the prior publication.[7] In 8 of these patients, the study protocol was repeated following shunt insertion. The average valve pressure set by the neurosurgeon was 105 ± 23 mm H_2O. The age of these patients was 30 ± 10 years, with 3 males and 5 females. The follow-up was 1.7 ± 2.7 years later. Twenty patients with idiopathic intracranial hypertension were entered from a prior publication undertaken by one of the current authors.[8] There were 18 women and 2 men of average age 40 ± 11 years.[8] In all 20, there was an increase in CSF opening pressure above 25 cm H_2O, with the mean pressure being 31 ± 4 cm H_2O. There was a normal ventricular size and no apparent cause for the pressure rise in these patients. Informed consent was obtained for all patients entering the original study,[7] which was given the authorization HNEHREC 07/03/21/5.03. Informed consent was also obtained for all patients entering this study. The study was reviewed by the Hunter New England Human Research Ethics Committee with authorization HNEHREC 16/06/15/5.06.

MR Imaging Technique

All patients were scanned on a 1.5T superconducting magnet (Verio; Siemens, Erlangen, Germany). In all patients, a standard brain MR imaging consisting of T1 sagittal, T2 axial, FLAIR axial, and diffusion-weighted axial images was performed. A time-of-flight MR venogram was obtained in the off sagittal plane with a slice thickness of 3 mm. An MR phase-contrast flow-quantification sequence was acquired with retrospective cardiac gating. The TR was 26.5 ms; TE, 6.9 ms; flip angle, 15°; slice thickness, 5 mm; matrix, 192 × 512; FOV, 150 mm; with a single excitation. The velocity-encoding value was 40 cm/s. The plane was selected to pass through the sagittal sinus 3 cm above the torcular and through the midpart of the straight sinus. The MR imaging was sourced from the hospital PACS; therefore, all measurements were performed on the original data. In all patients and controls, the cross-sectional area and wetted circumference of the sinuses were measured manually from the magnitude images of the flow-quantification series by one of the authors (G.A.B.) using the workstation measurement tool. The straight sinus length was measured manually from the MRV images from the junction with the vein of Galen to the torcular using the workstation curved length tool. The sagittal sinus length was measured from a point in the sagittal sinus that was in the same horizontal plane (the patient being supine) as the end of the straight sinus.

Analysis

The hydraulic diameter of each venous sinus segment was calculated using the formula:

$$1) \qquad Hd = 4A \,/\, Circ,$$

where Hd is the hydraulic diameter, A is the cross-sectional area of the sinus, and $Circ$ is the wetted circumference of the sinus.[9] The resistance of each sinus was calculated using the Poiseuille equation:

$$2) \qquad R = 8\mu L \,/\, \pi r^4,$$

where R is the resistance of each sinus, μ is the viscosity of blood (assumed to be 0.0035 Pa × s[10]), L is the measured length of each sinus, and r is the sinus radius, which is taken to be the hydraulic diameter divided by 2. The pressure drop through the sagittal and straight sinuses (ΔP) for each subject was calculated by multiplying the resistance value by the sinus flow (Q) in each instance:

$$3) \qquad \Delta P = RQ.$$

This was converted from pascals to millimeters of mercury for ease of discussion.

Statistical Analysis

Normality for all data was tested using a Shapiro-Wilk test with significance set at .05. Group means and SDs were obtained for each of the measurements. A nonpaired t test, with a P value $< .05$, was used to indicate statistical significance. A paired t test was used for the patients with hydrocephalus pre- and post-shunt insertion. All statistical analysis was performed using Matlab software (MathWorks, Natick, Massachusetts).

RESULTS

Measurements

The sinus hydraulic diameter, resistance, and pressure results are summarized in On-line Table 1. The pre- and postshunt results are summarized in On-line Table 2.

The mean sagittal sinus lengths for the controls and patients with hydrocephalus and IIH were 13.1 ± 1.2 cm, 13.5 ± 2.1 cm, and 13.1 ± 1.5 cm, respectively, and were not significantly different. In hydrocephalus, the hydraulic diameter of the sagittal sinus was reduced by 21% ($P < .001$) and the resistance was increased by 166% ($P = .002$) compared with controls. None of the other hydraulic diameters or resistances were significantly different. In controls, the sagittal and straight sinus blood flow averaged 5.7 ± 1.5 mL/s and 1.6 ± 0.5 mL/s, respectively; 4.6 ± 1.1 mL/s and 1.4 ± 0.4 mL/s in hydrocephalus, respectively; and 5.9 ± 1.3 mL/s and 1.8 ± 0.5 mL/s in IIH, respectively. In hydrocephalus, there was a 100% increase in the pressure drop across the sagittal sinus and a 38% decrease in the pressure in the straight sinus ($P = .01$ and .04). This result led to an increase in a sagittal sinus-to-straight sinus pressure difference of 1.2 mm Hg ($P = .001$) compared with controls. In IIH, the pressure differences were not significantly different from those of controls.

Following shunt insertion, in the smaller cohort of patients with hydrocephalus, the sagittal sinuses increased in size by 49% ($P = .007$). This change reduced the sagittal sinus resistance by 85% ($P = .05$) and eliminated the transvenous pressure difference.

DISCUSSION

The current study builds on previous work.[7,8] The sagittal sinus and straight sinus blood flow data have been previously published for the patients with hydrocephalus and controls.[7] The sagittal sinus blood flow has been previously published for patients with IIH.[8] All other data are new to this article.

The pressure drop through a vessel can be calculated using the Poiseuille equation.[11] Fall et al[12] used this equation, finding the normal resistance of the sagittal sinus to be 20.1 Pa/mL/s and the straight sinus to be 40.2 Pa/mL/s. Our method varies in 2 important ways from this article. First, we measured a shorter length of the sagittal sinus because we measured the same vertical distance from the MR imaging table as the end of the straight sinus to negate the necessity of allowing for the hydrostatic pressure difference; and second, we used the hydraulic diameter rather than assuming the sinuses to be circular in cross-section (personal correspondence with Dr Baledent PhD email March 2017). The flow through a triangular tube, such as a venous sinus, is much less efficient than in a cylindric tube.[13] The hydraulic diameter takes this reduced efficiency into account.[9] This combination of factors gave our controls a lower sagittal sinus resistance than that of Fall et al[12] of 15.9 Pa/mL/s and a higher straight sinus resistance of 108 Pa/mL/s. The sagittal sinus area and circumference have been measured by others using intravascular sonography in IIH, and the mean hydraulic diameter of 6.5 mm[14] is identical to that in our control cohort, suggesting an acceptable precision in our method.

Transvenous Pressure Difference

In controls, the pressure drop across the sagittal sinus averaged 0.7 mm Hg, and the straight sinus, 1.3 mm Hg, giving a transvenous pressure difference of −0.6 mm Hg (the negative value indicating that the deep territory sinus is of higher pressure). In hydrocephalus, the average sagittal sinus pressure drop was doubled at 1.4 mm Hg. The straight sinus resistance was not significantly different from that in the controls, but the blood flow was less, giving a lower pressure drop of 0.8 mm Hg. The average transvenous pressure difference in hydrocephalus was thus +0.6 mm Hg or 1.2 mm Hg higher than that in controls. In IIH, the sinus pressure difference was not significantly different from that in controls. Therefore, an increased transvenous pressure difference correlated with dilated ventricles and a normal difference with small ventricles.

Paradox of the Change in Sinus Size in Hydrocephalus

In a smaller cohort of 8 patients, the MR imaging protocol was repeated after shunt insertion. Pre-shunt insertion, all 21 patients showed small convex sagittal sinuses, similar to that in the Figure, *A*. Post-shunt insertion, all patients showed the free walls of the sinuses to be bowing outward (Figure, *B*). The change in appearance was noted to be a helpful sign that the shunt was working. The dilated sinuses looked like a filled spinnaker on a yacht (ie, the billowing sail sign; see the On-line Figure for further examples). The change in size highlights a paradox. The free walls of the sinuses sit between the CSF and the venous blood. The walls are concave when the pressure is higher in the CSF than in the lumen (positive transmural pressure difference) and convex when the pressure difference reverses.[6] Thus, in the 8 patients, the transmural pressure difference reversed after the shunt insertion. The degree that the sinus wall deflects depends on the magnitude of the transmural pressure difference and the stiffness of the sinus wall (ie, the elastic modulus).[6]

In healthy controls, the transmural pressure difference is known to be 4 mm Hg, and the sinus wall bulges slightly inward, giving a cross-sectional area of 42.1 mm^2.[6] The transmural pressure difference in patients with hydrocephalus was previously estimated to be 2.8 mm Hg,[7] with this smaller difference producing a much larger wall deflection than in the controls (see Figure, *C* for an example of the pressure differences).[6] The inference would be that the elastic modulus of the sagittal sinus in hydrocephalus is much lower (ie, the walls are floppy), but this cannot be true. The speed of pulse wave propagation along the sagittal sinus is increased by 2.4 times in normal pressure hydrocephalus.[15] The speed of a pulse wave is given by the equation $V = \sqrt{Eh / \rho r}$, where E is the elastic modulus, h is the vessel wall thickness, ρ the fluid density, and r is the vessel internal radius.[16] Given that the wall thickness and fluid density are constants, the ratio of the elastic modulus to the internal radius of the sinus in hydrocephalus must increase 5.8 times. The radius is reduced by 20% compared with that in controls. Therefore, the elastic modulus must be 4.6 times stiffer in hydrocephalus to account for the pulse wave velocity. The only way the sinus wall can be stiffer and yet still deflect further with a smaller pressure difference would be if the walls were irreversibly stretched. Note, post-shunt insertion, the sinuses greatly dilate with the CSF pressure set by the shunt valves in this cohort averaging 7.7 mm Hg. A normal sagittal sinus pressure is 7.5 mm Hg.[17] Thus, the sinus pressure post-shunt insertion must be minimally above the CSF pressure, but the sinuses still overdilate (Figure, *D–F*).

Venous Territories and Pressure Differences

The venous territories drained by the sagittal and straight sinuses are known to be separate. Intracerebral venous anastomoses through the centrum semiovale toward the convexity are nonexistent or negligible in humans.[18] Instead, a venous watershed exists, separating paraventricular white matter from a layer of subcortical white matter.[18] Thus, a pressure difference could be maintained between the superficial and deep venous territories. The capillary and venular pressure in the brain is thought to be maintained by a Starling resistor due to the compression of the connecting veins as they pass through the subarachnoid space.[19] Thus, a pressure difference in the sinuses would be irrelevant to the brain parenchyma if both brain territories had their pressures set by the intracranial pressure. However, Portnoy et al[2] maintained that the Starling mechanism is absent in the deep system and the pressure in the deep parenchyma is similar to that in the dural sinuses. The current study confirms these findings. The pressure difference between the CSF and the straight sinus is high in hydrocephalus, so vein of Galen compression should be obvious. No venous compression was evident in any of the MRVs. The patient in the Figure, *C* had the largest CSF-to-vein of Galen pressure difference of the cohort, but no compression was seen. Thus, it appears that the reduced venous pressure in the deep sinuses is free to propagate throughout the entire deep system.

The modeling of hydrocephalus in dogs shows an increase

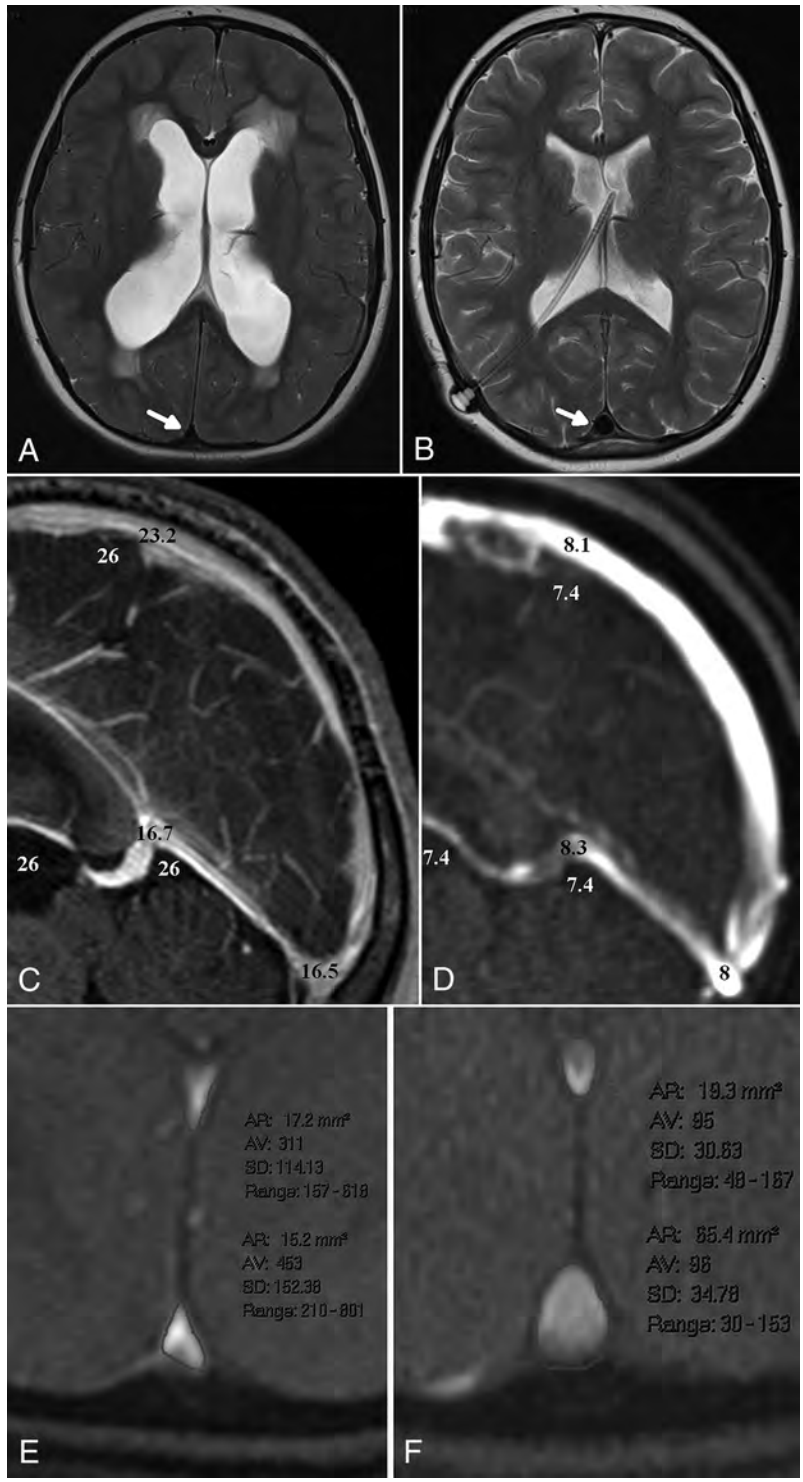


FIGURE. *A*, A T2 image of a 16-year-old girl with idiopathic hydrocephalus. Note the transependymal CSF spread. The *arrow* shows the small concave sagittal sinus. *B*, A T2 image following shunt insertion. The valve setting was 100 mm H_20. The sinus has increased in size and has convex margins (ie, the billowing sail sign). *C*, A postcontrast T1 reconstruction pre-shunt insertion. The *white numbers* are the CSF pressure from a lumbar puncture. The *black numbers* represent the calculated sinus pressures. The 2.8-mm Hg pressure drop from the CSF to the sagittal sinus is from a previous work.[7] Note the 9.3-mm Hg pressure drop from the CSF to the vein of Galen with no sign of venous collapse. The thalamostriate vein was not compressed but had moved out of plane. *D*, A reconstruction from the MRV post-shunt insertion, with CSF and venous pressures appended. The CSF pressure of 7.4 mm Hg is the shunt tube setting. The sagittal sinus pressure of 8.1 mm Hg is an estimate, given the convex sinus. *E*, A flow-quantification magnitude image pre-shunt insertion. The straight sinus area is 17.2 mm^2, and the sagittal sinus area is 15.2 mm^2. *F*, A flow-quantification magnitude image post-shunt insertion shows the straight sinus unchanged at 19.3 mm^2, but the sagittal sinus is much larger at 65.4 mm^2.

in sagittal sinus pressure, reduced sinus compliance, and abnormal collateral veins.[20] The entire sinus length was constricted in the hydrocephalic dogs,[20] identical to findings in the current study.

MR elastography measures the stiffness of the parenchyma of an organ. The stiffness of the brain increases with venous compression in the neck.[21] MR elastography in normal pressure hydrocephalus shows increased stiffness in the parietal and occipital lobes (drained by the sagittal sinus[22]) and decreased stiffness in the periventricular white matter[23] (drained by the deep system[22]). Thus, the MR elastography suggests an increased pressure in the superficial venous system and decreased pressure in the deep system.

Is a 1.2-mm Hg Pressure Difference Enough?

The current study has a mixture of patients with active and balanced hydrocephalus, so the calculated pressure difference may be lower than if only active cases were used. Therefore, the patients with shunts with active hydrocephalus have higher transvenous pressures than the remainder of the cohort (1.5 versus 0.6 mm Hg).

Hydrocephalus is associated with CSF ventricular reflux[24] and evidence of reversed aqueduct flow.[25] Transependymal CSF absorption may exist in hydrocephalus, but is a 1.2-mm Hg pressure difference enough to dilate the ventricles? It has been suggested that a transmural pressure difference need not be large to dilate the ventricles.[25] Kim et al[26] found that the ventricular enlargement could be explained by a transmantle pressure gradient of 1.0 mm Hg, similar to the current estimate.

Studies using sagittal sinus ligation in animal models uniformly produce an elevation in CSF pressure but no evidence of hydrocephalus in the acute setting.[27] Similarly, acute thrombosis of the sagittal sinus does not produce hydrocephalus despite an obvious pressure difference between the venous territories. Therefore, there must be another variable involved over and above the pressure drop.

CSF Capillary Absorption

The absorption of CSF into the capillaries requires passage across the blood-brain barrier. Hladky and Barrand[28] argued that net absorption of CSF across an intact blood-brain barrier is not sustainable regardless of the hydrostatic pressure because the salt would be left behind and a rapid increase in osmotic pressure would negate the hydrostatic pressure difference within minutes. This argument suggests why ligating the sagittal sinus does not produce ventricular dilation (ie, the blood-brain barrier stays closed). In human hydrocephalus, the capillary wall shows blood-brain barrier dysfunction with increased vesicular and vacuolar transport, open interendothelial junctions, thin and fragmented basement membranes, and discontinuous perivascular astrocytic end-feet.[29]

Difference between Hydrocephalus and IIH

The absorption of CSF is traditionally seen to occur through the arachnoid granulations, requiring a pressure difference of 4 mm Hg.[8] This pressure drop is maintained by the structural integrity of the walls of the sinuses.

In IIH, there is collapse of the sinuses below the torcular, and there is an elevation in central venous pressure due to obesity.[30] The venous collapse has the effect of raising the venous sinus pressure and initially reducing the pressure drop across the arachnoid granulations so CSF absorption stops. The CSF pressure will rise until a new equilibrium is reached, where the pressure drop across the arachnoid granulations reverts back to 4 mm Hg and CSF absorption resumes at a higher overall pressure. There is no ventricular enlargement because the pressure in the deep brain parenchyma is unchanged compared with the superficial parenchyma (ie, there is no transependymal CSF absorption, and this prevents the dilation from occurring).

In hydrocephalus, there is collapse of the sagittal sinus. This sets up a pressure differential between the superficial and deep brain parenchyma. The increase in sagittal sinus pressure stops CSF absorption across the arachnoid granulations, and the CSF pressure initially rises. If the blood-brain barrier becomes deficient in the region of the subependymal white matter, then CSF can be absorbed through this route. If the CSF transependymal absorption matches CSF production, then the CSF pressure will fall back into the normal range. A small pressure difference across the ependyma will dilate the ventricles across time. The pressure in the sagittal sinus remains slightly elevated; thus, the pressure across the arachnoid granulations is reduced, and the pressure drop remains unfavorable for CSF absorption via this route. Shunt insertion eliminates the transvenous pressure difference, and the ventricles can be reduced in size because the pressure difference between the deep capillaries and CSF reverses.

Limitations

The measurements occurred in quiet respiration in supine patients. It is difficult to predict how the pressure changes would be altered in the upright position. The pressure drop between the CSF and sagittal sinus in the Figure was estimated from a previous article. The Davson equation would predict a higher pressure difference of 5–6 mm Hg.[7] This would actually increase the pressure difference from the CSF to the subependymal capillaries, making ventricular dilation more likely.

CONCLUSIONS

The size of the sagittal sinus cross-section is smaller in hydrocephalus than in controls and IIH. The cross-sectional area of the straight sinus is not significantly different across groups. If one considered the sinus blood flow, the current modeling suggests a small increase in transvenous pressure in hydrocephalus but no change in IIH. This finding may have a bearing on the cause of ventricular dilation. Shunting reverses the change.

ACKNOWLEDGMENTS

We thank Dr O. Baledent for clarifying the technique used in the article by Fall et al.[11]

REFERENCES

1. Bondurant CP, Jimenez DF. **Epidemiology of cerebrospinal fluid shunting.** *Pediatr Neurosurg* 1995;23:254–58; discussion 259 CrossRef Medline
2. Portnoy HD, Branch C, Castro ME. **The relationship of intracranial venous pressure to hydrocephalus.** *Childs Nerv Syst* 1994;10:29–35 CrossRef Medline

3. Stephensen H, Tisell M, Wikkelsö C. **There is no transmantle pressure gradient in communicating or noncommunicating hydrocephalus.** *Neurosurgery* 2002;50:763–73; discussion 771–73 CrossRef Medline
4. Eide PK, Saehle T. **Is ventriculomegaly in idiopathic normal pressure hydrocephalus associated with a transmantle gradient in pulsatile intracranial pressure?** *Acta Neurochir (Wein)* 2010;152:989–95 CrossRef Medline
5. Bateman GA. **Vascular compliance in normal pressure hydrocephalus.** *AJNR Am J Neuroradiol* 2000;21:1574–85 Medline
6. Bateman GA, Lechner-Scott J, Copping R, et al. **Comparison of the sagittal sinus cross-sectional area between patients with multiple sclerosis, hydrocephalus, intracranial hypertension and spontaneous intracranial hypotension: a surrogate marker of venous transmural pressure?** *Fluids Barriers CNS* 2017;14:18 CrossRef Medline
7. Bateman GA, Siddique SH. **Cerebrospinal fluid absorption block at the vertex in chronic hydrocephalus: obstructed arachnoid granulations or elevated venous pressure?** *Fluids Barriers CNS* 2014;11:11 CrossRef Medline
8. Bateman GA. **Arterial inflow and venous outflow in idiopathic intracranial hypertension associated with venous outflow stenosis.** *J Clin Neurosci* 2008;15:402–08 CrossRef Medline
9. Atkins T, Escudier M. *A Dictionary of Mechanical Engineering.* Oxford: Oxford University Press; 2013;174
10. Hund S, Kameneva M, Antaki J. **A quasi-mechanistic mathematical representation for blood viscosity.** *Fluids* 2017;2:10 CrossRef
11. Sutera SP, Skalak R. **The history of Poiseuille's law.** *Annual Review of Fluid Mechanics*1993;25:1–19 CrossRef
12. Fall S, Pagé G, Bettoni J, et al. **Use of phase-contrast MRA to assess intracranial venous sinus resistance to drainage in healthy individuals.** *AJNR Am J Neuroradiol* 2017;38:281–87 CrossRef Medline
13. Lekner J. **Viscous flow through pipes of various cross-sections.** *Eur J Phys* 2007;28:521–27 CrossRef
14. Boddu SR, Gobin P, Oliveira C, et al. **Anatomic measurements of cerebral venous sinuses in idiopathic intracranial hypertension patients.** *PLos One* 2018;13:e0196275 CrossRef Medline
15. Bateman GA, Lechner-Scott J, Lea RA. **A comparison between the pathophysiology of multiple sclerosis and normal pressure hydrocephalus: is pulse wave encephalopathy a component of MS?** *Fluids Barriers CNS* 2016;22;13:18 CrossRef Medline
16. Best CH, Taylor NB, Brobeck JR. *Physiological Basis of Medical Practice***.** 12th ed. Baltimore: Williams & Wilkins; 1991:147
17. Ekstedt J. **CSF hydrodynamic studies in man, 2: normal hydrodynamic variables related to CSF pressure and flow.** *J Neurol Neurosurg Psychiatry* 1978;41:345–53 CrossRef Medline
18. Andeweg J. **The anatomy of collateral venous flow from the brain and its value in aetiological interpretation of intracranial pathology.** *Neuroradiology* 1996;38:621–28 CrossRef Medline
19. Schaller B. **Physiology of cerebral venous blood flow: from experimental data in animals to normal function in humans.** *Brain Res Brain Res Rev* 2004;46:243–60 CrossRef Medline
20. Olivero WC, Rekate HL, Chizeck HJ, et al. **Relationship between intracranial and sagittal sinus pressure in normal and hydrocephalic dogs.** *Pediatr Neurosci* 1988;14:196–201 CrossRef Medline
21. Hatt A, Cheng S, Tan K, et al. **MR elastography can be used to measure brain stiffness changes as a result of altered cranial venous drainage during jugular compression.** *AJNR Am J Neuroradiol* 2015; 36:1971–77 CrossRef Medline
22. Meder JF, Chiras J, Roland J, et al. **Venous territories of the brain.** *J Neuroradiol* 1994;21:118–33 Medline
23. Perry A, Graffeo CS, Fattahi N, et al. **Clinical correlation of abnormal findings on magnetic resonance elastography in idiopathic normal pressure hydrocephalus.** *World Neurosurg* 2017;99:695–700.e1 CrossRef Medline
24. Bateman GA. **The pathophysiology of idiopathic normal pressure hydrocephalus: cerebral ischemia or altered venous hemodynamics?** *AJNR Am J Neuroradiol* 2008;29:198–203 CrossRef Medline
25. Balédent O, Gondry-Jouet C, Meyer ME, et al. **Relationship between cerebrospinal fluid and blood dynamics in healthy volunteers and patients with communicating hydrocephalus.** *Invest Radiol* 2004;39: 45–55 CrossRef Medline
26. Kim H, Jeong EJ, Park DH, et al. **Finite element analysis of periventricular lucency in hydrocephalus: extravasation or transependymal CSF absorption?** *J Neurosurg* 2016;124:334–41 CrossRef Medline
27. Miyagami M, Nakamura S, Moriyasu N. **Ventricular enlargement in experimental occlusion of superior sagittal sinus in reference to histopathological findings (author's transl)** [in Japanese]. *No Shinkei Geka* 1975;3:947–54 Medline
28. Hladky SB, Barrand MA. **Mechanisms of fluid movement into, through and out of the brain: evaluation of the evidence.** *Fluids Barriers CNS* 2014;11:26 CrossRef Medline
29. Castejón OJ. **Submicroscopic pathology of human and experimental hydrocephalic cerebral cortex.** *Folia Neuropathol* 2010;48:159–74 Medline
30. Ahmed RM, Wilkinson M, Parker GD, et al. **Transverse sinus stenting for idiopathic intracranial hypertension: a review of 52 patients and of model predictions.** *AJNR Am J Neuroradiol* 2011;32:1408–14 CrossRef Medline

Absence of Disproportionately Enlarged Subarachnoid Space Hydrocephalus, a Sharp Callosal Angle, or Other Morphologic MRI Markers Should Not Be Used to Exclude Patients with Idiopathic Normal Pressure Hydrocephalus from Shunt Surgery

S. Agerskov, M. Wallin, P. Hellström, D. Ziegelitz, C. Wikkelsö, and M. Tullberg

ABSTRACT

BACKGROUND AND PURPOSE: Several studies have evaluated the use of MR imaging markers for the prediction of outcome after shunt surgery in idiopathic normal pressure hydrocephalus with conflicting results. Our aim was to investigate the predictive value of a number of earlier proposed morphologic MR imaging markers in a large group of patients with idiopathic normal pressure hydrocephalus.

MATERIALS AND METHODS: One hundred sixty-eight patients (mean age, 70 ± 9.3 years) with idiopathic normal pressure hydrocephalus, subjected to standardized quantification of clinical symptoms before and after shunt surgery, were included in the study. Outcome was calculated using a composite score. Preoperative T1, FLAIR, and flow-sensitive images were analyzed regarding the presence of 13 different morphologic MR imaging markers.

RESULTS: The median Evans index was 0.41 (interquartile range, 0.37–0.44). All patients had an aqueductal flow void sign present and white matter hyperintensities. The median callosal angle was 68.8° (interquartile range, 57.7°–80.8°). Dilated Sylvian fissures were found in 69%; focally dilated sulci, in 25%; and widening of the interhemispheric fissure, in 55%. Obliteration of the sulci at the convexity was found in 36%, and 36% of patients were characterized as having disproportionately enlarged subarachnoid space hydrocephalus. Sixty-eight percent of patients improved after surgery. None of the investigated MR imaging markers were significant predictors of improvement after shunt surgery.

CONCLUSIONS: Disproportionately enlarged subarachnoid space hydrocephalus, a small callosal angle, and the other MR imaging markers evaluated in this study should not be used to exclude patients from shunt surgery. These markers, though they may be indicative of idiopathic normal pressure hydrocephalus, do not seem to be a part of the mechanisms connected to the reversibility of the syndrome.

ABBREVIATIONS: DESH = disproportionately enlarged subarachnoid space hydrocephalus; EI = Evans index; iNPH = idiopathic normal pressure hydrocephalus

Idiopathic normal pressure hydrocephalus (iNPH) is a syndrome of gait and balance disturbances, cognitive dysfunction, and urinary incontinence seen predominantly in the elderly population.[1-5] Enlarged ventricles on CT or MR imaging are required for the diagnosis, and different radiologic evaluation techniques have been used to increase the diagnostic and predictive accuracy ever since its first description by Adams et al.[6] Today, MR imaging is considered the standard radiologic method, and MR imaging–based criteria for diagnosing the condition are incorporated in the international and Japanese diagnostic guidelines.[1,2]

In the international guidelines, ventriculomegaly with an Evans index (EI) of >0.3 in combination with at least 1 of 4 supportive findings is required for the diagnosis of probable iNPH.[2] The Japanese criteria instead emphasize the finding of disproportionately enlarged subarachnoid space hydrocephalus (DESH), requiring it for the diagnosis of probable iNPH if no Tap-Test or CSF drainage test is performed.[1]

The connection between radiologic findings and symptoms in iNPH has also recently been reinforced with the use of a composite scale score including morphologic CT findings in patients with iNPH.[7]

The use of morphologic MR imaging markers for selecting appropriate shunt surgery candidates has been investigated, but results vary; the use of these markers for predictive purposes is still disputed.[8-12] DESH[13] has won support as a prognostic marker,[14] most recently by Shinoda et al,[15] who developed a 10-point grading scale to aid in patient selection for surgery. Virhammar

Received September 5, 2018; accepted after revision October 24.

From the Hydrocephalus Research Unit (S.A., M.W., P.H., C.W., M.T.), Department of Clinical Neuroscience and Rehabilitation, Institute of Neuroscience and Physiology, and Department of Neuroradiology (D.Z.), Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden.

This work was supported by grants from the Edit Jacobson Foundation, the Rune and Ulla Almlövs Foundation for Neurological Research, and the Gothenburg Foundation for Neurological Research.

Please address correspondence to Simon Agerskov, MD, Hydrocephalus Research Unit, Institute of Neuroscience and Physiology, Department of Clinical Neuroscience, Sahlgrenska Academy, University of Gothenburg, Sahlgrenska University Hospital, 41345 Gothenburg, Sweden; e-mail: simon.agerskov@vgregion.se

http://dx.doi.org/10.3174/ajnr.A5910

Table 1: Demographic data of 168 patients with iNPH

Data	
Age (mean) (SD) (yr)	71 (9.3)
Male/female (%)	61:39
Symptom duration (mean) (SD) (mo)	47 (59)
Comorbidity (%)	
Hypertension	49
Cardiovascular disease	26
Diabetes mellitus	15

et al[8,16] reported that the presence of DESH, a narrow callosal angle (<63°), and dilation of the temporal horns were predictors of good surgical outcome.

The use of morphologic MR imaging markers for predicting outcome after shunt surgery in patients with iNPH requires further investigation. Hence, the aim of this study was to investigate the association between 13 morphologic MR imaging markers and postoperative outcome in a large consecutive cohort of patients with iNPH subjected to a detailed clinical evaluation.

MATERIALS AND METHODS

One-hundred sixty-eight patients consecutively diagnosed with iNPH in accordance with the international guidelines[2] who underwent shunt surgery between 2006 and 2013 were included in the study. Patients were included if they had a complete preoperative MR imaging scan including volumetric T1, FLAIR, and flow-sensitive T2 sequences. If substantial movement artifacts were present on any of the sequences, the patient was excluded. Demographic data of the patient group are shown in Table 1.

Clinical Evaluation

All patients underwent detailed clinical examinations by a neurologist and a physiotherapist preoperatively and 3–6 months postoperatively following standardized protocols.[5] To evaluate outcome after shunt surgery, we created a composite score incorporating 4 continuous measures based on 2 gait tests and 2 cognitive tests (the Timed 10-Meter Walk Test[17]; the Timed Up and Go Test[18]; the Identical Forms Test, measuring perceptual speed and accuracy; and the Bingley Memory Test[19]). Each score was standardized into a 0–100 scale with 0 representing the worst possible performance and 100 equaling the mean performance of healthy individuals at 70 years of age.[4,17,20] The composite score was calculated using the mean value of the 4 included tests, and the difference between the pre- and postoperative score constituted the outcome for each patient. Patients were classified as improved if their score increased by ≥5 points.[4,5] All shunts were functioning at the time of the postoperative clinical evaluation.

Imaging

All preoperative MR imaging scans were performed using a 1.5T Intera (Philips Healthcare, Best, the Netherlands) or a 1.5T Achieva dStream (Philips Healthcare) scanner. The imaging sequences were reformatted and analyzed using an Advantage Workstation 2.0 (GE Healthcare, Milwaukee, Wisconsin). The imaging protocol consisted of the following: 1) a sagittal T1-weighted volume sequence with TR/TE = 25/4.6 ms, flip angle = 30°, FOV = 260 × 260 × 190 mm^3, matrix = 260 × 259, and 380 overcontiguous 1-mm slices; 2) a transaxial FLAIR sequence with TR/TE = 9000/100 ms, inversion recovery delay = 2500 ms, FOV = 230 × 230 × 230 mm^3, matrix = 192 × 192, 44 sections, 3-mm slice thickness, and no gap; 3) a flow-sensitive sagittal TSE sequence with TR/TE = 10,238/300 ms, FOV = 230 × 230 × 38 mm^3, matrix = 384 × 284, thirteen sections, 2-mm slice thickness, and 1-mm gap; and 4) an aqueduct-centered turbo field echo sequence with TR/TE= 7.7/3.8 ms, flip angle = 60°, FOV = 150 × 150 × 15 mm^3, matrix = 256 × 256 with 75 overcontiguous 0.2-mm slices.

On the T1-weighted volume sequences, a line connecting the anterior and posterior commissures was defined (ie, the anterior/posterior commissure plane).[8,16,21] The T1-weighted 3D volumes were then reformatted generating coronal images perpendicular to and transaxial images parallel to the anterior/posterior commissure plane. All reformatted T1-weighted volume sequences had voxel sizes of 1 × 1 × 1 mm and were used with the FLAIR, TSE, and turbo field echo sequences in the image analyses.

In total, 13 imaging markers were analyzed (Figure).

The EI was measured on transaxial T1-weighted images as the index between the maximum diameter of the frontal horns and the maximum inner skull diameter in the slice above the foramen of Monro (Fig *A*). Then, the maximum diameter of the temporal horns was recorded bilaterally (Fig *B*). The callosal angle was analyzed on coronal T1 images at the level of the posterior commissure (Fig *C*).[16]

Coronal T1 images were also used to measure the widest diameter of the third ventricle between the anterior and posterior commissures (Fig *D*). On sagittal T1 slices, the widest anteroposterior midline diameter of the fourth ventricle was determined along a line perpendicular to the posterior border of the brain stem (Fig *E*).

Obliteration of the high-convexity sulci was assessed on transaxial T1 images and graded as obliterated if no sulci were distinguishable on the 10 most cranial slices covering the vertex (Fig *F*).

The presence of focally enlarged (transport) sulci was analyzed on transaxial and coronal T1 series. The sulci were only determined as focally widened if there were no signs of general cortical atrophy, the sulcal widening was asymmetric, and the affected sulci lacked connection with the Sylvian fissure. The number of focally widened sulci was recorded (Fig *G*).

Dilation of the Sylvian fissures was measured on coronal T1 images using an ordinal scale (Fig *H*).[16,22] Widening of the anterior part of the interhemispheric fissure was estimated on transaxial T1 images using a 3-step ordinal scale (Fig *I*). The flow void phenomenon in the cerebral aqueduct and fourth ventricle (flow void sign) was evaluated and graded using the ordinal scale developed by Algin et al[23] and later modified by Virhammar et al (Fig *J*).[8] Periventricular and deep white matter hyperintensities were analyzed on transaxial FLAIR series using the scale developed by Fazekas et al.[24]

DESH was considered present if patients showed signs of Sylvian fissure dilation (ordinal rating 1 or 2) in conjunction with obliterated sulci at the high convexity.

Analyses were performed in a retrospective manner with investigators blinded to the patients' clinical data. To ensure reproducibility, 2 authors (S.A., M.W.) analyzed all variables independently in 10 randomly selected patients and calculated interrater

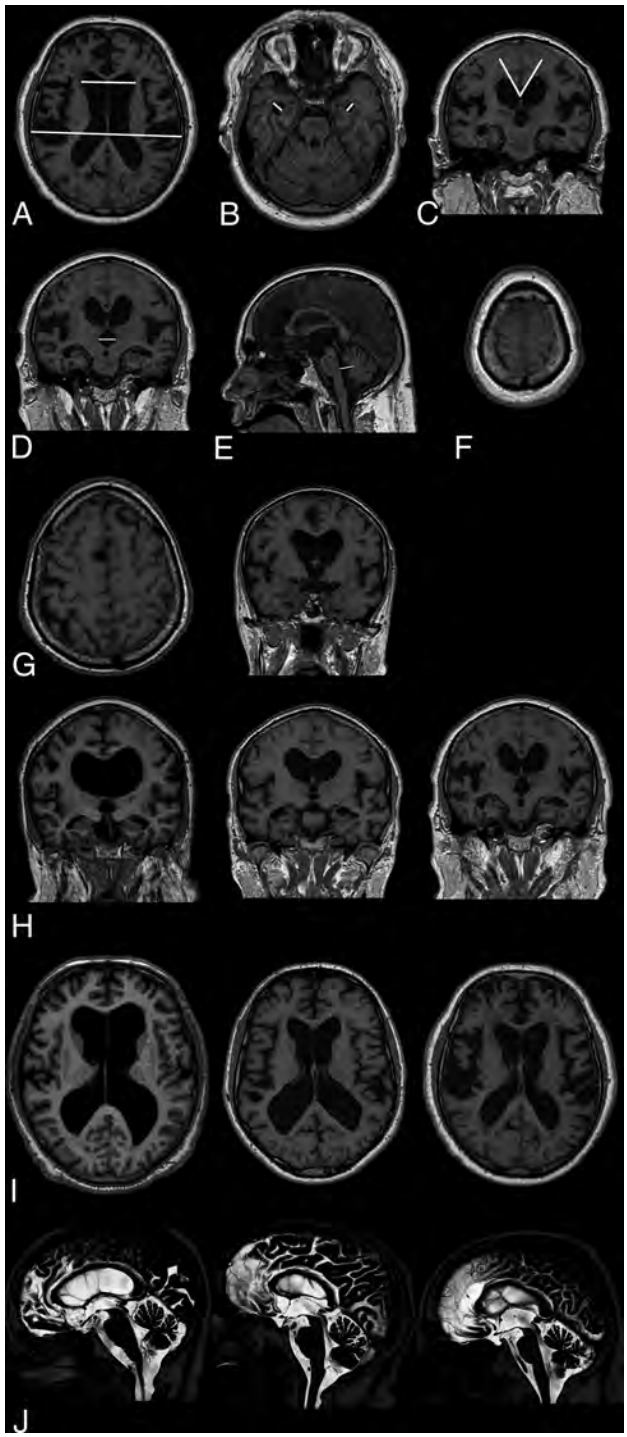


FIGURE. Morphologic MR imaging markers analyzed in 168 patients with iNPH preoperatively. *A*, Evans index. *B*, Maximum width of the temporal horns. *C*, Callosal angle. *D*, Maximum width of the third ventricle. *E*, Maximum anteroposterior diameter of the fourth ventricle. *F*, Obliteration of sulci at the vertex. *G*, Transport of sulci on transaxial and coronal images (left-right). *H*, Width of the Sylvian fissures graded 0 to 1–2 (left-right). *I*, Width of the interhemispheric fissure graded 0 to 1–2 (left-right). *J*, Flow void sign graded 1 to 2–3 (left-right, 0 not shown in the figure).

reliability. In cases in which discrepancies occurred, the variables were redefined and re-evaluated until an interrater reliability of >0.7 was achieved. The development of the image-analysis protocol was supervised by an experienced neuroradiologist (D.Z.) who also assisted in the refinement of MR imaging markers as required until the interrater reliability ratings were sufficient. In addition, 20% of patient scans were randomly selected and re-evaluated by one of the authors (S.A.) to calculate the test-retest reliability, which was >0.8 for all variables.

Table 2: Pre- and postoperative gait, cognitive, and total scores in 168 patients with iNPH

	Preoperative	Postoperative	*P*
Gait (median) (IQR)	38 (17–61)	66 (43–87)	<.001
Cognition (median) (IQR)	35 (19–66)	50 (20–75)	<.001
Total (median) (IQR)	39 (23–63)	57 (38–76)	<.001

Note:—IQR indicates interquartile range.

Statistics

All statistical tests were performed using nonparametric procedures. Differences in distributions among binary variables were tested using the McNemar test. Tests of differences between groups for ordinal and interval data were performed using the Wilcoxon rank sum test. Correlations were tested using Spearman rank correlations. Associations between outcome and analyzed MR imaging variables were assessed using logistic regression models with results presented as odds ratios with 95% confidence intervals. Interrater and test-retest reliability was calculated using intraclass correlation coefficients for continuous variables and the weighted/unweighted Cohen κ for ordinal and nominal variables, respectively. Statistical significance was $P < .05$. All calculations were performed in SPSS, Version 24.0, released in 2014 (IBM, Armonk, New York).

Ethical Considerations

The data collection was approved by the Ethics Committee for Medical Research at Gothenburg University, with written informed consent obtained from all participants or close relatives.

RESULTS

Surgical Outcome

Sixty-eight percent ($n = 115$) of the patients improved after shunt surgery, 23% ($n = 39$) were unchanged, and conditions of 8% ($n = 14$) deteriorated. Improvement was most pronounced in the gait domain (Table 2).

Preoperative MR Imaging Findings and Outcome

All patients had an EI of >0.3, a present flow void sign, white matter hyperintensities, and a callosal angle of <90° (Table 3). Dilation of the Sylvian fissures was found in 72%, while focal dilation of the supra-Sylvian sulci and obliteration of sulci at the high convexity were more uncommon (28%–36%). Thirty-nine percent had a callosal angle of <63°(responders, 39%; nonresponders, 38%; P = not significant). There were no significant differences between responders and nonresponders in the distribution of any of the morphologic MR imaging markers or in the prevalence of periventricular and deep white matter changes. Furthermore, the severity of white matter changes did not have any effect on postoperative outcome.

In the logistic regression models adjusted for age and sex, no MR imaging marker was significantly associated with postoperative improvement, neither for the total score nor for any of the subdomain scores.

Table 3: Preoperative morphologic MRI markers in 168 patients with iNPH

All Patients	No.	Median/% (IQR)	Improved (Median/%)	Not Improved (Median/%)	*P*
Evans index		0.41 (0.37–0.44)	0.40	0.39	NS
Temporal horns (mm)		9.1 (7.5–11.0)	9.0	9.1	NS
Callosal angle		68° (56°–81°)	68°	69°	NS
Third ventricle width (mm)		15.5 (13.3–18)	15.4	16.5	NS
Fourth ventricle width (mm)		14.7 (13.2–16.3)	14.7	14.6	NS
Transport sulci (%)	0	72	75	68	NS
	1	17	19	11	
	2	8	6	14	
	>2	3	0	7	
Dilated interhemispheric fissure (%)	0	42	45	38	NS
	1	49	45	54	
	2	9	10	8	
Obliterated sulci at vertex (%)		36	36	35	NS
Dilated Sylvian fissure (%)	0	28	31	22	
	1	45	45	46	
	2	27	24	32	
DESH present (%)		36	36	34	NS
Flow void sign (%)	0	0	0	0	
	1	30	30	26	
	2	42	37	47	
	3	28	35	27	
PVH (%)	0	0	0	0	NS
	1	57	58	49	
	2	26	25	29	
	3	17	17	22	
DWMH (%)	0	0	0	0	NS
	1	46	47	45	
	2	37	42	29	
	3	17	11	26	

Note:—PVH indicates periventricular hyperintensities; DWMH, deep white matter hyperintensities; NS, not significant; IQR, interquartile range.

Table 4: Correlation coefficients between clinical symptom severity and preoperative MRI findings in 168 patients with iNPH

	Total Symptom Score	Gait Score	Cognition Score
EI	−0.09	−0.04	−0.12
Temporal horns (mean)	−0.30	0.1	−0.15[a]
Interhemispheric fissure	−0.20	−0.04	−0.04
Callosal angle	0.17	0.13	0.16
Third ventricle	−0.19	−0.25[a]	0.07
Sylvian fissures (mean)	−0.1	−0.15[a]	0.08
Obliterated sulci	0.19	0.18	0.13
DESH	0.11	0.06	0.11
Flow void	0.03	0.09	−0.03
Fourth ventricle	−0.20[a]	−0.20[a]	0.16
Transport sulci	0.07	0.18	0.08
PVH	−0.26[a]	−0.22[a]	−0.23[a]
DWMH	−0.22[a]	−0.21[a]	−0.17[a]

Note:—PVH indicates periventricular hyperintensities; DWMH, deep white matter hyperintensities.
[a] $P < .05$.

Correlations with Clinical Symptoms

A few measures were significantly correlated with preoperative clinical symptoms. However, the correlations were all weak (ie, <.30) (Table 4).

DISCUSSION

In this study of 168 patients with iNPH, we analyzed a number of proposed MR imaging markers for the prediction of outcome after shunt surgery. We could not show significant associations between any of the analyzed MR imaging markers and postoperative improvement, nor were there significant differences in the presence of proposed imaging markers between improved and nonimproved patients. Specifically, a small callosal angle or the finding of DESH was not associated with a favorable outcome.

The lack of correlation between morphologic MR imaging markers and postoperative improvement corroborates some earlier studies[9,12,25] and contradicts other publications that have reported significant associations between postoperative improvement and the presence of a narrow callosal angle,[16] the DESH phenomenon,[15,16] and dilation of the temporal horns.[8]

A possible explanation for the different results regarding the predictive value of morphologic imaging markers reported here compared with many earlier studies could be the selection of patients. We based the iNPH diagnosis and decision to perform shunt surgery on clinical and radiologic criteria, and only in patients in whom the outcome of shunt surgery was considered uncertain was the CSF Tap-Test or lumbar infusion test used as a supplementary test. Lumbar puncture was performed for intracranial pressure measurement and exclusion of other disorders. Furthermore, with the exception of the EI, none of the imaging markers were specifically required for diagnosis. Diagnostic criteria requiring the presence of DESH, a positive response to CSF drainage, or an increased resistance to CSF outflow such as the Japanese guidelines,[1] entail a possible selection bias in which patients who would potentially improve after shunt surgery were excluded from studies. We believe that the inclusion of patients with iNPH in our study is more general with less selection bias compared with many earlier studies, which could explain differences in results reported. Overall, we consider the patient sample representative, and the results reported here are robust. Moreover, our results are in agreement with those of Craven et al[12]; and in a recent report by Benedetto et al[26] using a CT-based method to assess DESH, there were no differences between patients who improved and those who did not improve after shunt surgery.

Another possible cause of the contradictory results in some earlier studies might be the use of different outcome measures. In our study, the aim was to use a sensitive outcome measure similar to the iNPH scale designed by Hellström et al[4] in order to base our calculations on improvements in continuous variables that are norm-based and thus reproducible. Many of the previous studies have used the modified Rankin Scale, which was developed for use in patients with stroke and does not measure symptom severity in

iNPH but instead provides a general measure of disability.[27] Other groups have applied outcome scales based on ordinal or nominal ratings,[8,10,15] making a direct comparison with the results presented here more difficult. To maximize the sensitivity and validity of results, one should use quantitative outcome measures if possible.[4] Using outcome measurements that are blunt or potentially not measuring an actual improvement of hydrocephalic symptoms (such as the modified Rankin Scale) increases the risk of misjudging proposed imaging markers and their use as predictors of postoperative outcome.

The prevalence of DESH in this study is lower in comparison with previous publications.[8,13,28] This outcome might be because in accordance with the international diagnostic guidelines, the components of the DESH phenomenon were not part of the diagnostic criteria for the evaluated iNPH group. Furthermore, we graded sulcal compression at the vertex as obliterated or not, meaning that patients who did not show complete sulcal obliteration but still had some degree of compression in conjunction with Sylvian dilation were graded as not having DESH. However, DESH prevalence figures in patients with iNPH of around 30% have previously been reported,[12] and the same authors did not find any support for DESH as a predictive factor for shunt responsiveness, confirming our results.

While DESH is a common finding in patients with iNPH and can aid in diagnosing the disorder, this study implies that it should be used neither as an obligatory diagnostic finding nor as a predictive marker, given the risk of excluding patients from shunt surgery who might benefit from the procedure.

All patients in our study presented with an EI of >0.3, in agreement with the international guidelines.[2] In addition, the ventriculomegaly also involved the third and fourth ventricles in most patients, corroborating previous findings.[9] The enlargement of the third and fourth ventricles also correlated significantly, albeit weakly, with the gait symptom score. Although these findings could not significantly predict good postoperative outcome, they are still important to consider from a pathophysiologic aspect. Infratentorial periventricular structures might be involved in the development of clinical symptoms.[5,29,30]

All patients presented with white matter changes on preoperative MR imaging. The severity of these changes did not differ significantly between shunt responders and nonresponders in this study; this result corroborates previously published work[8,31] and reinforces the theory that the extent of white matter damage should not exclude patients from shunt surgery.

Our findings of only weak correlations between ventricular dilation or white matter changes on the one hand and symptom severity on the other differ markedly from a recent study reporting an association between 8 CT-based imaging markers and the severity of clinical symptoms.[7] The cited study used the iNPH grading scale,[4] and the statistical analysis was performed using linear regression modeling, which we were unable to reproduce given the absence of a linear relationship between our dependent and independent variables and non-normal data distribution.

Our findings support the view that clinical improvement in iNPH after shunt surgery is mainly attributed to increased metabolism and extracellular fluid flow in predominantly periventricular regions of the brain and not morphologic changes as measured on structural MR imaging. Support for this notion comes from imaging studies of perfusion[32-35] and diffusion[29,36] and CSF biomarker studies[37,38] as well as a recent study indicating reduced glymphatic clearance in patients with iNPH.[39]

The results reported here imply that morphologic MR imaging markers only correlate with symptom severity in a limited way and cannot predict postoperative outcome. To find reliable markers for selecting appropriate candidates for shunt surgery, focus should be turned to the use of higher order MR imaging analyses, such as diffusion- and perfusion-based techniques as well as combinations of MR imaging and biochemical methods. Further studies are needed in this area.

Strengths and Weaknesses

The strengths of this study are the large consecutively included patient population, a prospective data collection, and the detailed assessment of clinical outcome after shunt surgery. We also realigned all scans before the subsequent analysis, thus minimizing the effects of possible misalignment. In addition, with the exception of the EI, the analysis of all variables was performed after the diagnosis of iNPH was made, thus reducing the risk of selection bias.

The main limitations are the retrospective image analysis and the fact that a group of patients diagnosed before 2006 had to be excluded because they lacked MR imaging scans that fulfilled our inclusion criteria. The excluded patients did not differ in any demographic data nor in our outcome score compared with our study population. In all, we consider the evaluated patient sample and the results reported here to be representative.

CONCLUSIONS

DESH, a small callosal angle, and the other MR imaging markers evaluated in this study should not be used to exclude patients from shunt surgery. These markers, though they may be indicative of iNPH, do not seem to be a part of the mechanisms connected to the reversibility of the syndrome.

Disclosures: Simon Agerskov—*RELATED*: *Grant*: Rune and Ulla Amlöv Foundation, Edit Jacobson Foundation, Gothenburg Foundation for Neurological Research. Per Hellström—*UNRELATED*: *Employment*: Clinical Neuropsychologist at Sahlgrenska University Hospital. Carsten Wikkelsö—*UNRELATED*: *Board Membership*: Associate Editor of *Acta Neurologica Scandinavica*.

REFERENCES

1. Mori E, Ishikawa M, Kato T, et al; Japanese Society of Normal Pressure Hydrocephalus. **Guidelines for management of idiopathic normal pressure hydrocephalus: second edition.** *Neurol Med Chir (Tokyo)* 2012;52:775–809 CrossRef Medline
2. Relkin N, Marmarou A, Klinge P, et al**. Diagnosing idiopathic normal-pressure hydrocephalus.** *Neurosurgery* 2005;57(3 Suppl):S4–16; discussion ii-v Medline
3. Blomsterwall E, Svantesson U, Carlsson U, et al. **Postural disturbance in patients with normal pressure hydrocephalus.** *Acta Neurol Scand* 2000;102:284–91 CrossRef Medline
4. Hellström P, Klinge P, Tans J, et al. **A new scale for assessment of severity and outcome in iNPH.** *Acta Neurol Scand* 2012;126:229–37 CrossRef Medline
5. Agerskov S, Hellström P, Andrén K, et al. **The phenotype of idiopathic normal pressure hydrocephalus: a single center study of 429 patients.** *J Neurol Sci* 2018;391:54–60 CrossRef Medline

6. Adams RD, Fisher CM, Hakim S, et al. **Symptomatic occult hydrocephalus with normal cerebrospinal-fluid pressure: a treatable syndrome.** *N Engl J Med* 1965;273:117–26 CrossRef Medline
7. Kockum K, Lilja-Lund O, Larsson EM, et al. **The idiopathic normal-pressure hydrocephalus Radscale: a radiological scale for structured evaluation.** *Eur J Neurol* 2018;25:569–76 CrossRef Medline
8. Virhammar J, Laurell K, Cesarini KG, et al. **Preoperative prognostic value of MRI findings in 108 patients with idiopathic normal pressure hydrocephalus.** *AJNR Am J Neuroradiol* 2014;35:2311–18 CrossRef Medline
9. Wikkelsö C, Andersson H, Blomstrand C, et al. **Computed tomography of the brain in the diagnosis of and prognosis in normal pressure hydrocephalus.** *Neuroradiology* 1989;31:160–65 CrossRef Medline
10. Kazui H, Mori E, Ohkawa S, et al. **Predictors of the disappearance of triad symptoms in patients with idiopathic normal pressure hydrocephalus after shunt surgery.** *J Neurol Sci* 2013;328:64–69 CrossRef Medline
11. Halperin JJ, Kurlan R, Schwalb JM, et al. **Practice guideline: idiopathic normal pressure hydrocephalus—response to shunting and predictors of response: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology.** *Neurology* 2015;85:2063–71 CrossRef Medline
12. Craven CL, Toma AK, Mostafa T, et al. **The predictive value of DESH for shunt responsiveness in idiopathic normal pressure hydrocephalus.** *J Clin Neurosci* 2016;34:294–98 CrossRef Medline
13. Hashimoto M, Ishikawa M, Mori E, et al; Study of INPH on neurological improvement (SINPHONI). **Diagnosis of idiopathic normal pressure hydrocephalus is supported by MRI-based scheme: a prospective cohort study.** *Cerebrospinal Fluid Res* 2010;7:18 CrossRef Medline
14. Garcia-Armengol R, Domenech S, Botella-Campos C, et al. **Comparison of elevated intracranial pressure pulse amplitude and disproportionately enlarged subarachnoid space (DESH) for prediction of surgical results in suspected idiopathic normal pressure hydrocephalus.** *Acta Neurochir (Wien)* 2016;158:2207–13 CrossRef Medline
15. Shinoda N, Hirai O, Hori S, et al. **Utility of MRI-based disproportionately enlarged subarachnoid space hydrocephalus scoring for predicting prognosis after surgery for idiopathic normal pressure hydrocephalus: clinical research.** *J Neurosurg* 2017;127:1436–42 CrossRef Medline
16. Virhammar J, Laurell K, Cesarini KG, et al. **The callosal angle measured on MRI as a predictor of outcome in idiopathic normal-pressure hydrocephalus.** *J Neurosurg* 2014;120:178–84 CrossRef Medline
17. Steffen TM, Hacker TA, Mollinger L. **Age- and gender-related test performance in community-dwelling elderly people: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds.** *Phys Ther* 2002;82:128–37 CrossRef Medline
18. Podsiadlo D, Richardson S. **The timed "Up & Go": a test of basic functional mobility for frail elderly persons.** *J Am Geriatr Soc* 1991;39:142–48 CrossRef Medline
19. Tullberg M, Rosengren L, Blomsterwall E, et al. **CSF neurofilament and glial fibrillary acidic protein in normal pressure hydrocephalus.** *Neurology* 1998;50:1122–27 CrossRef Medline
20. Solana E, Poca MA, Sahuquillo J, et al. **Cognitive and motor improvement after retesting in normal-pressure hydrocephalus: a real change or merely a learning effect?** *J Neurosurg* 2010;112:399–409 CrossRef Medline
21. Ishii K, Kanda T, Harada A, et al. **Clinical impact of the callosal angle in the diagnosis of idiopathic normal pressure hydrocephalus.** *Eur Radiol* 2008;18:2678–83 CrossRef Medline
22. Kitagaki H, Mori E, Ishii K, et al. **CSF spaces in idiopathic normal pressure hydrocephalus: morphology and volumetry.** *AJNR Am J Neuroradiol* 1998;19:1277–84 Medline
23. Algin O, Hakyemez B, Taskapilioglu O, et al. **Morphologic features and flow void phenomenon in normal pressure hydrocephalus and other dementias: are they really significant?** *Acad Radiol* 2009;16:1373–80 CrossRef Medline
24. Fazekas F, Chawluk JB, Alavi A, et al. **MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging.** *AJR Am J Roentgenol* 1987;149:351–56 CrossRef Medline
25. Virhammar J, Laurell K, Cesarini KG, et al. **Increase in callosal angle and decrease in ventricular volume after shunt surgery in patients with idiopathic normal pressure hydrocephalus.** *J Neurosurg* 2018 Feb 2:1–6. [Epub ahead of print] CrossRef Medline
26. Benedetto N, Gambacciani C, Aquila F, et al. **A new quantitative method to assess disproportionately enlarged subarachnoid space (DESH) in patients with possible idiopathic normal pressure hydrocephalus: the SILVER index.** *Clin Neurol Neurosurg* 2017;158:27–32 CrossRef Medline
27. Wilson JT, Hareendran A, Hendry A, et al. **Reliability of the modified Rankin Scale across multiple raters: benefits of a structured interview.** *Stroke* 2005;36:777–81 CrossRef Medline
28. Ishikawa M, Oowaki H, Takezawa M, et al. **Disproportionately enlarged subarachnoid space hydrocephalus in idiopathic normal-pressure hydrocephalus and its implication in pathogenesis.** *Acta Neurochir Suppl* 2016;122:287–90 CrossRef Medline
29. Jurcoane A, Keil F, Szelenyi A, et al. **Directional diffusion of corticospinal tract supports therapy decisions in idiopathic normal-pressure hydrocephalus.** *Neuroradiology* 2014;56:5–13 CrossRef Medline
30. Lee PH, Yong SW, Ahn YH, et al. **Correlation of midbrain diameter and gait disturbance in patients with idiopathic normal pressure hydrocephalus.** *J Neurol* 2005;252:958–63 CrossRef Medline
31. Tullberg M, Jensen C, Ekholm S, et al. **Normal pressure hydrocephalus: vascular white matter changes on MR images must not exclude patients from shunt surgery.** *AJNR Am J Neuroradiol* 2001;22:1665–73 Medline
32. Ziegelitz D, Starck G, Kristiansen D, et al. **Cerebral perfusion measured by dynamic susceptibility contrast MRI is reduced in patients with idiopathic normal pressure hydrocephalus.** *J Magn Reson Imaging* 2014;39:1533–42 CrossRef Medline
33. Ziegelitz D, Arvidsson J, Hellström P, et al. **In patients with idiopathic normal pressure hydrocephalus postoperative cerebral perfusion changes measured by dynamic susceptibility contrast magnetic resonance imaging correlate with clinical improvement.** *J Comput Assist Tomogr* 2015;39:531–40 CrossRef Medline
34. Tullberg M, Hellström P, Piechnik SK, et al. **Impaired wakefulness is associated with reduced anterior cingulate CBF in patients with normal pressure hydrocephalus.** *Acta Neurol Scand* 2004;110:322–30 CrossRef Medline
35. Klinge PM, Brooks DJ, Samii A, et al. **Correlates of local cerebral blood flow (CBF) in normal pressure hydrocephalus patients before and after shunting: a retrospective analysis of [(15)O]H(2)O PET-CBF studies in 65 patients.** *Clin Neurol Neurosurg* 2008;110:369–75 CrossRef Medline
36. Tullberg M, Ziegelitz D, Ribbelin S, et al. **White matter diffusion is higher in Binswanger disease than in idiopathic normal pressure hydrocephalus.** *Acta Neurol Scand* 2009;120:226–34 CrossRef Medline
37. Jeppsson A, Höltta M, Zetterberg H, et al. **Amyloid mis-metabolism in idiopathic normal pressure hydrocephalus.** *Fluids Barriers CNS* 2016;13:13 CrossRef Medline
38. Nakajima M, Miyajima M, Ogino I, et al. **Cerebrospinal fluid biomarkers for prognosis of long-term cognitive treatment outcomes in patients with idiopathic normal pressure hydrocephalus.** *J Neurol Sci* 2015;357:88–95 CrossRef Medline
39. Ringstad G, Vatnehol SA, Eide PK. **Glymphatic MRI in idiopathic normal pressure hydrocephalus.** *Brain* 2017;140:2691–705 CrossRef Medline

Brain β-Amyloid and Atrophy in Individuals at Increased Risk of Cognitive Decline

I.K. Martikainen, N. Kemppainen, J. Johansson, J. Teuho, S. Helin, Y. Liu, S. Helisalmi, H. Soininen, R. Parkkola, T. Ngandu, M. Kivipelto, and J.O. Rinne

ABSTRACT

BACKGROUND AND PURPOSE: The relationship between brain β-amyloid and regional atrophy is still incompletely understood in elderly individuals at risk of dementia. Here, we studied the associations between brain β-amyloid load and regional GM and WM volumes in older adults who were clinically evaluated as being at increased risk of cognitive decline based on cardiovascular risk factors.

MATERIALS AND METHODS: Forty subjects (63–81 years of age) were recruited as part of a larger study, the Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability. Neuroimaging consisted of PET using ^{11}C Pittsburgh compound-B and T1-weighted 3D MR imaging for the measurement of brain β-amyloid and GM and WM volumes, respectively. All subjects underwent clinical, genetic, and neuropsychological evaluations for the assessment of cognitive function and the identification of cardiovascular risk factors.

RESULTS: Sixteen subjects were visually evaluated as showing cortical β-amyloid (positive for β-amyloid). In the voxel-by-voxel analyses, no significant differences were found in GM and WM volumes between the samples positive and negative for β-amyloid. However, in the sample positive for β-amyloid, increases in ^{11}C Pittsburgh compound-B uptake were associated with reductions in GM volume in the left prefrontal ($P = .02$) and right temporal lobes ($P = .04$).

CONCLUSIONS: Our results show a significant association between increases in brain β-amyloid and reductions in regional GM volume in individuals at increased risk of cognitive decline. This evidence is consistent with a model in which increases in β-amyloid incite neurodegeneration in memory systems before cognitive impairment manifests.

ABBREVIATIONS: AD = Alzheimer disease; *APOE* = *Apolipoprotein E*; Aβ = β-amyloid; PIB = Pittsburgh compound-B; PIB− = PIB negative; PIB+ = PIB positive

Alzheimer disease (AD), the most common form of late-life dementia, is characterized by abnormal deposits of neurofibrillary tangles of τ protein and plaques of β-amyloid (Aβ) protein in the brain, eventually leading to neurodegeneration and cognitive decline. The accumulation of Aβ in the brain is believed to be a key factor in the development of AD, and recent evidence suggests that reduction of brain Aβ in the early stages of AD may slow down cognitive and functional decline.[1] Therefore, there is a need to find biomarkers that identify individuals at risk of developing AD pathology who might benefit from therapeutic interventions before substantial irreversible neurodegeneration occurs.

Neuroimaging using PET and ligands specific for Aβ such as ^{11}C-labeled-Pittsburgh compound-B (^{11}C PIB) allows the measurement of brain fibrillary Aβ load in vivo. Previous studies have found increases in brain ^{11}C PIB uptake not only in patients with AD but also in patients at risk of AD.[2] The increases in ^{11}C PIB

Received November 9, 2017; accepted after revision October 12, 2018.

From the Department of Radiology (I.K.M.), Medical Imaging Center, Tampere University Hospital, Tampere, Finland; Division of Clinical Neurosciences (N.K., J.O.R.), Turku University Hospital, Turku, Finland; Turku PET Centre (N.K., J.J., J.T., S. Helin, J.O.R.), University of Turku, Turku, Finland; Department of Neurology (Y.L., S. Helisalmi, H.S., M.K.), University of Eastern Finland, Kuopio, Finland; Neurocenter (Y.L., H.S., M.K.), Neurology, Kuopio University Hospital, Kuopio, Finland; Department of Radiology (R.P.), University of Turku and Turku University Hospital, Turku, Finland; Department of Public Health Solutions (T.N., M.K.), Public Health Promotion Unit, National Institute for Health and Welfare, Helsinki, Finland; and Division of Clinical Geriatrics (T.N., M.K.), Center for Alzheimer Research, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden.

This study was supported by Finnish Governmental Research Funding for Turku University Hospital and Tampere University Hospital; the Finnish Medical Foundation; the Sigrid Jusélius Foundation; the Maud Kuistila Foundation; the Paulo Foundation; the Research Council for Health of the Academy of Finland (15762, 259615, 278457, 287490, 294061; and Responding to Public Health Challenges Research Program grants 129395, 129397, 129421, 129416, 129401); the La Carita Foundation; the Alzheimer's Association (grant HAT-10-173121); the Juho Vainio Foundation; the Novo Nordisk Foundation; the Finnish Social Insurance Institution; the Ministry of Education and Culture, Finland; the Swedish Research Council; the Alzheimer's Research and Prevention Foundation, United States; the AXA Research Fund; the Sheikha Salama bint Hamdan Al Nahyan Foundation; the Academy of Finland for Joint Program of Neurodegenerative Disorders–prevention (Multimodal preventive trials for Alzheimer's Disease); the Swedish Research Council; and the Swedish Research Council for Health, Working Life, and Welfare.

Please address correspondence to Ilkka K. Martikainen, MD, Department of Radiology, Medical Imaging Center, Tampere University Hospital, PO Box 2000, 33521 Tampere, Finland; e-mail: ilkka.martikainen@pshp.fi; @IKMartikainen

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

http://dx.doi.org/10.3174/ajnr.A5891

uptake in these patient samples are significantly associated with brain atrophy.[3,4] However, the evidence is mixed on the relationship of brain volumes and Aβ load in elderly subjects without clear cognitive impairment: Earlier studies have found both positive[5,6] and negative associations[7] or no associations.[8] This inconsistency may be related to differences in methodology and the stage of AD pathology of the samples.

Here, we studied the relationships between brain Aβ and *apolipoprotein E* (*APOE*) *ε4* carrier status with regional GM and WM volumes in a population-based sample of elderly individuals without manifest cognitive impairment but at high risk of developing dementia based on a cardiovascular risk factor profile. Earlier research suggests substantial regional variation in the accelerated brain atrophy related to early Aβ accumulation.[9] Therefore, we hypothesized that increases in ^{11}C PIB uptake are associated with specific patterns of brain volume loss. A better understanding of the relationship between Aβ and brain atrophy would not only elucidate AD mechanisms in at-risk subjects but also potentially help develop imaging-based identification of individuals who might benefit from early intervention.

MATERIALS AND METHODS

Subjects

The subjects were recruited as part of the Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (clinicaltrials.gov identifier NCT01041989). The study enrolled subjects 60–77 years of age with Cardiovascular Risk Factors, Aging and Dementia scores of at least 6 points[10] and cognition at a mean level or slightly lower than that expected for age. At least 1 of the following criteria was required for inclusion: 1) Word List Memory Task of ≤19 words, 2) Word List Recall of ≤75%, or 3) Mini-Mental State Examination score of ≤26/30 points. In general, the subjects are representative of the Finnish elderly population with several risk factors for dementia.[11,12] The exclusion criteria included major depression, dementia, or marked cognitive decline, Mini-Mental State Examination scores of <20, and symptomatic cardiovascular disease.

Here, we studied a subgroup of the above-mentioned sample (Turku University Hospital cohort), consisting of 40 subjects (21 men, 19 women; mean age, 71 ± 5.2 years). Before analyses, all neuroimaging data were evaluated for image quality. Written informed consent was obtained from all subjects who participated in the study. The study was approved by the Coordinating Ethics Committee of the Helsinki and Uusimaa Hospital District.

Clinical Measurements

The clinical measurements have been previously described in detail.[13] The cognitive performance was evaluated using the modified Neuropsychological Test Battery,[14] yielding a total composite *z* score and domain *z* score measures of memory, executive function, and processing speed. Total serum cholesterol and plasma glucose concentrations were determined enzymatically using commercial reagents and a clinical chemistry analyzer, Architect c8000 (Abbott Laboratories, Abbott Park, Illinois).

APOE Genotyping

Genomic DNA was extracted from venous blood samples with a chemagic Magnetic Separation Module I (Perkin Elmer, Waltham, Massachusetts) using magnetic beads. The *APOE* genotype was determined by polymerase chain reaction using TaqMan genotyping assays (Applied Biosystems, Foster City, California) for 2 single-nucleotide polymorphisms (rs429358 and rs7412) and an allelic discrimination method on the ABI 7500 platform (Applied Biosystems).[15]

Neuroimaging

^{11}C PIB [N-methyl-^{11}C-2-(4-methylaminophenyl)-6-hydroxybenzothiazole] was produced as described earlier.[16] On average, 406 ± 110 MBq of ^{11}C PIB was injected intravenously, and a scan from 60 to 90 minutes (3 × 10-minute frames) after injection was performed with an Ingenuity TF PET/MR scanner (Philips Healthcare, Best, the Netherlands). All images were reconstructed using a line-of-response row-action maximum likelihood algorithm with MR imaging–based attenuation correction using a segmentation-based algorithm with 3 tissue classes, including the head coil template used in the MR imaging protocol.[17] The data were reconstructed using 2 iterations and 33 subsets. The image matrix size was 128 × 128 × 90, with an axial FOV of 256 × 256 mm and an isotropic voxel dimension of 2 mm. All quantitative corrections for PET data were applied, including scatter, randoms, attenuation, detector deadtime, and normalization. Neither time-of-flight information nor resolution modeling was applied in this study. Sagittal T1-weighted 3D MR imaging data were acquired for the measurement of brain GM and WM volumes, with TR = 25 ms, TE = 5.5 ms, and a reconstructed isotropic voxel dimension of 1 mm.

PET and MR Imaging Data Processing

The neuroimaging data were processed using SPM8 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12). The ^{11}C PIB images were realigned and coregistered to the individual MR image and normalized to Montreal Neurological Institute space. ^{11}C PIB uptake data were extracted using the standard automated segmentations by FreeSurfer 5.0 (http://surfer.nmr.mgh.harvard.edu).[18] Regional ^{11}C PIB uptake was quantified as a region-to-cerebellar cortex ratio during the 60- to 90-minute scan duration. The ^{11}C PIB uptake values in the right and left hemispheres were averaged for data analysis.

MR imaging data processing for voxel-based morphometry analysis was performed using the VBM8 toolbox (http://dbm.neuro.uni-jena.de/vbm.html), with default parameters for image processing. This included bias regularization and tissue classification and registration using linear (affine) and nonlinear transformations within a unified model.[19] High-dimensional spatial normalization was accomplished using Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra. The analysis was performed on the volumes of GM and WM, multiplied by the nonlinear, but not linear, components derived from the normalization matrix. This procedure preserves actual local GM and WM volumes, accounting for individual brain size (modulated volume). The realigned and normalized GM and WM segments

Table 1: Sample demographics and clinical characteristics[a]

Characteristics	All	PIB−	PIB+	PIB− vs PIB+
No. (% women)	40 (48%)	24 (50%)	16 (44%)	$\chi^2_1 = 0.15, P = .70$
Age (yr)	71 ± 5	70 ± 6	72 ± 4	$t_{38} = -1.19, P = .24$
APOE ε4 carriers (%)	12 (31)	4 (17)	8 (53)	$\chi^2_1 = 5.83, P = .02$
Body mass index (kg/m^2)	27 ± 3	28 ± 3	26 ± 2	$t_{38} = 2.14, P = .04$
Systolic blood pressure (mm Hg)	137 ± 15	138 ± 15	136 ± 15	$t_{38} = 0.31, P = .76$
Diastolic blood pressure (mm Hg)	81 ± 8	80 ± 8	82 ± 9	$t_{38} = -0.65, P = .52$
Total serum cholesterol level (mmol/L)	5.2 ± 1	4.9 ± 1	5.5 ± 1	$t_{38} = 1.69, P = .10$
Education (yr)	9 (2)	9 (2.75)	8 (2)	$U = 158, P = .36$
Physical activity ≥ twice per week (%)	63	52	80	$\chi^2_1 = 3.02, P = .08$
Mini-Mental State Examination	27 (2)	27 (2)	27 (3)	$U = 158, P = .35$
Total composite *z* score	0.03 ± 0.5	0.06 ± 0.6	−0.01 ± 0.5	$F_{1,37} = 0.018, P = .90$
Memory *z* score	−0.05 ± 0.6	−0.10 ± 0.5	0.06 ± 0.7	$F_{1,37} = 2.42, P = .13$
Executive function *z* score	0.04 ± 0.6	0.20 ± 0.6	−0.10 ± 0.4	$F_{1,37} = 1.94, P = .17$
Processing speed *z* score	0.2 (0.9)	0.4 (1)	0.2 (1)	$F_{1,37} = 0.03, P = .86$
Dementia risk score	7.5 (3.75)	8 (3.75)	7 (2)	$U = 161, P = .39$
Fazekas score	1 (2)	1 (2)	2 (1.75)	$U = 144, P = .18$

[a] Data are given as mean ± SD or median (interquartile range).

were smoothed with a Gaussian kernel with a full width at half maximum size of 8 mm.

Data Analysis

The PET images were visually interpreted by 2 experienced readers, and subjects were classified as either PIB positive (PIB+) or PIB negative (PIB−) on the basis of consensus agreement. The subjects with PIB+ findings had cortical ^{11}C PIB retention in at least 1 region typically affected by β-amyloid deposition in AD, while the subjects with PIB− findings had only nonspecific ^{11}C PIB retention in the WM. Brain GM and WM volumes were compared between the PIB+ and PIB− samples and *APOE ε4* carriers and noncarriers voxel-by-voxel using an unpaired *t* test. The associations between ^{11}C PIB uptake and brain GM and WM volumes, and interactions between ^{11}C PIB uptake and *APOE ε4* carrier status with regional GM and WM volumes were analyzed using whole-brain voxel-by-voxel multiple linear regression analysis. Age and sex were covaried in all analyses. Primary analyses were conducted using a composite ^{11}C PIB uptake value, calculated as the average uptake in the following regions: anterior cingulate cortex, lateral temporal cortex, parietal cortex, posterior cingulate cortex, precuneus, and prefrontal cortex. Additionally, the ^{11}C PIB uptake values for the precuneus, prefrontal cortex, and posterior cingulate cortex were used for regression analysis because they are among the first regions to show increases in ^{11}C PIB uptake in mild cognitive impairment.[2] The associations were determined separately in the PIB+ and PIB− samples.

Voxels with GM or WM values of < 0.1 were excluded from the analysis. A height threshold of $P < .001$ (uncorrected) was used across the whole brain for searching significant differences in brain GM and WM volumes ($P < .05$, family-wise error rate–corrected for multiple comparisons at the cluster level). Extent threshold was defined by the expected number of voxels per cluster based on random field theory. Cluster sizes were adjusted for nonstationary smoothness.[20] In addition to the measures of GM and WM volume, microangiopathy-related hyperintensities in deep WM were evaluated on axial FLAIR images using a semiquantitative scale (0 = absence, 1 = punctate foci, 2 = beginning confluence of foci, 3 = large confluent areas).[21]

Statistical analyses were performed using SPSS, Version 23.0 (SPSS Statistics for Windows; IBM, Armonk, New York). Planned correlations were determined between clinical measures and GM and WM volumes in the PIB+ and PIB− samples. In these analyses, statistical significance was set at $P < .05$, and no correction for multiple tests was applied to these correlations.

RESULTS

General Characteristics and APOE *Genotype of the Subjects*

The general characteristics of the subjects are shown in Table 1. *APOE* genotype was determined in 39 subjects. This sample consisted of 12 *APOE ε4* carriers (ε2/ε4, $n = 1$; ε3/ε4, $n = 10$; ε4/ε4, $n = 1$) and 27 *APOE ε4* noncarriers (ε2/ε3, $n = 2$; ε3/ε3, $n = 25$).

Brain ^{11}C PIB Uptake and GM and WM Volumes

The average global brain GM volume was 618 ± 73 mL in the PIB+ sample and 624 ± 52 mL in the PIB− sample, and the average brain WM volume was 502 ± 76 mL in the PIB+ sample and 520 ± 61 mL in the PIB− sample. There were no differences in the global GM and WM volumes between the samples (GM: $F_{1,37} = 0.17$, $P = .68$; WM: $F_{1,37} = 1.21$, $P = .28$). In the whole-brain voxel-by-voxel analyses, no significant differences were found in the GM and WM volumes between the PIB+ and PIB− samples, even when using a very lenient search threshold ($P < .05$, uncorrected).

In line with our earlier work, significant correlations were found between the regional ^{11}C PIB uptake values and the composite ^{11}C PIB uptake value in the PIB+ sample (all bivariate ROI correlations, $P \leq .001$; Pearson $r = 0.70–0.98$). Therefore, the composite ^{11}C PIB uptake value can be used as a proxy for overall brain Aβ load in this study.

In the PIB+ sample, increases in composite ^{11}C PIB uptake were associated with smaller GM volumes in the right temporal lobe (temporal pole, parahippocampal gyrus; peak Montreal Neurological Institute coordinates at [29, 14, −30], cluster size = 1170 mm^3, $z = 4.0$, $P = .04$). At a lower search threshold ($P < .01$), this region also encompassed the right hippocampus and areas of the medial occipitotemporal gyrus (Fig *A, -B*). The increases in ^{11}C PIB uptake in the precuneus were associated with smaller GM volumes in the left prefrontal lobe (inferior frontal

gyrus; [−38, 32, 4], 262 mm^3, $z = 3.8$, $P = .02$; Fig *C*). At trend level ($P = .06$), an association was found between increases in ^{11}C PIB uptake in the prefrontal cortex and smaller GM volumes in the right temporal lobe (temporal pole, parahippocampal gyrus, amygdala; [29, 12, −30], 1390 mm^3, $z = 3.7$). There were no regions with a significant positive correlation between ^{11}C PIB uptake and GM volume. In the PIB− sample, no significant associations were found between brain ^{11}C PIB uptake and GM volume (all regions, $P > .05$). No significant associations were found between ^{11}C PIB uptake and regional WM volume in the PIB+ or PIB− samples. The main results are summarized in Table 2.

APOE ε4 Carrier Status and GM and WM Volumes

No significant differences were found in global brain GM volumes between *APOE ε4* carriers and noncarriers ($F_{1,36} = 1.44$, $P = .24$). The global WM volumes were significantly smaller in *APOE ε4* carriers compared with noncarriers ($F_{1,36} = 4.62$, $P = .04$). In the whole-brain voxel-by-voxel analysis, no significant differences were found in regional GM volumes when comparing *APOE ε4* carriers and noncarriers. Furthermore, no significant interactions were found between *APOE ε4* carrier status and composite ^{11}C PIB uptake with regional GM volume. Regarding regional WM, *APOE ε4* carriers had significant reductions in WM volumes in the right parieto-occipital region compared with *APOE ε4* noncarriers ([18, −87, 32], 4810 mm^3, $z = 4.7$, $P = .001$). No significant regions were found in the opposite contrast (*APOE ε4* carriers > *APOE ε4* noncarriers). No interactions were found between *APOE ε4* carrier status and composite ^{11}C PIB uptake with regional WM volume.

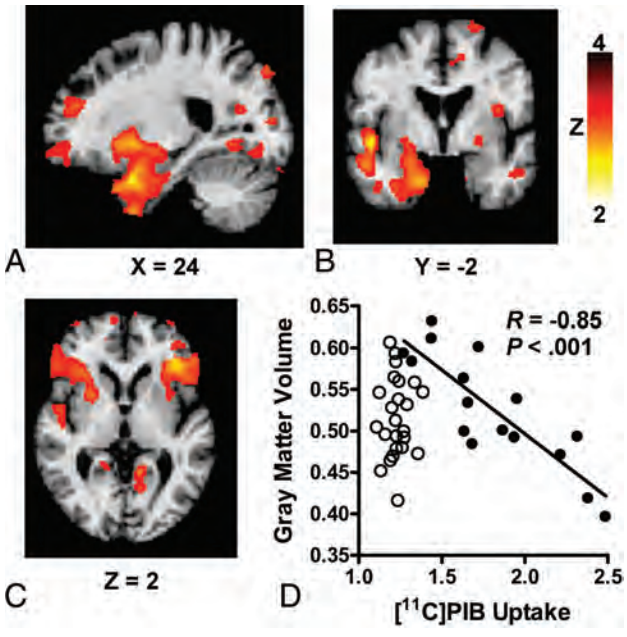


FIGURE. Associations between increases in ^{11}C PIB uptake and reductions in GM volume in subjects with PIB+ findings at increased risk of cognitive decline. In subjects classified as positive for ^{11}C PIB uptake, significant associations are found between increases in composite ^{11}C PIB uptake and reductions in GM volume in the right temporal lobe (*A* and *B*) and increases in ^{11}C PIB uptake in the precuneus and reductions in GM volume in the left prefrontal lobe (*C*). *D*, The significant negative correlation between composite ^{11}C PIB uptake and GM volume in the right temporal lobe in the PIB+ sample (*closed circles*). No correlation is found in the PIB− sample (*open circles*). The *R* and *P* values were calculated using the average GM volumes extracted at $P < .001$.

Clinical Correlations

In the overall sample, larger global GM and WM volumes were associated with younger age (GM: $r = -0.35$, $P = .01$; WM: $r = -0.37$, $P = .009$). Larger global WM volumes were also associated with higher Mini-Mental State Examination scores at trend level ($r = 0.28$, $P = .08$). Correlations between regional GM or WM volumes and clinical measures, including WM hyperintensities, were short of significance (all *P* values > .05).

DISCUSSION

PET imaging using ligands specific for Aβ, such as ^{11}C PIB, allows the evaluation of AD pathology even before clinical symptoms emerge. However, a number of patients with PET scans positive for ^{11}C PIB do not develop AD, indicating that other biomarkers are needed to accurately identify individuals who might benefit from pharmaceutical or life-style interventions.[1,12] Earlier studies in patients with AD and its prodromal states have found significant correlations between ^{11}C PIB uptake and brain volume loss,[3,4,22,23] which is more closely related to the cognitive symptoms than Aβ load.[6,8] Significant correlations have been described between brain Aβ accumulation and atrophy even in elderly subjects with no cognitive symptoms[5,6]; in fact, it has been suggested that this correlation is particularly strong at early stages of AD pathology.[24,25] Our present data extend the previous observations by showing significant associations between increases in brain Aβ and GM loss in elderly subjects at high risk of cognitive impairment. However, some studies in cognitively healthy subjects have failed to find associations between ^{11}C PIB uptake and atrophy[8] or have even suggested a positive correlation between ^{11}C PIB uptake and GM volume.[7] Considering the distinct time scales of brain Aβ accumulation and GM loss, these discrepancies may relate to differences in the stage of AD pathology among the samples.

In the PIB+ sample, a significant association was found between increases in the composite measure of brain ^{11}C PIB uptake and smaller GM volumes in the right temporal lobe, encompassing structures of the medial temporal lobe memory system. This finding is in line with previous work in cognitively healthy elderly subjects[5,6,24,25] and suggests that asymptomatic elderly individuals who are at risk of cognitive decline and have a substantial brain Aβ load show signs of impending neurodegeneration in

Table 2: Summary of associations between ^{11}C PIB uptake and brain regional GM volume

^{11}C PIB Uptake Measure	Correlation with GM Volume; Region; Peak X, Y, Z Coordinates; Cluster Size; *z* Score; *P* Value
Composite	Negative correlation, right temporal lobe (29, 14, −30), 1170 mm^3, $z = 4.0$, $P = .04$
Precuneus	Negative correlation, left prefrontal lobe (−38, 32, 4), 262 mm^3, $z = 3.8$, $P = .02$
Prefrontal cortex	Negative correlation, right temporal lobe (29, 12, −30), 1390 mm^3, $z = 3.7$, $P = .06$

the temporal lobe. Although no significant associations were found between temporal lobe volume and cognition (Mini-Mental State Examination, total *z* score, and subscores of the modified Neuropsychological Test Battery), the significance of coexisting brain Aβ load and temporal lobe atrophy in clinically healthy elderly is highlighted by data suggesting that the adverse effect of these variables on cognition is synergistic.[26] Therefore, it is likely that the subjects with PIB+ findings with temporal lobe atrophy in this study are at high risk of future cognitive decline.

In addition, increases in ^{11}C PIB uptake in the precuneus in the PIB+ sample were significantly associated with GM volume reductions in the left prefrontal lobe. While studies on brain atrophy in the context of AD pathology have generally focused on medial temporal lobe structures, AD has also been shown to be associated with atrophy in a number of other brain regions, including the frontal lobes, precuneus, and posterior cingulate cortex.[27,28] Furthermore, brain Aβ load in cognitively healthy elderly has been shown to be associated with GM volume loss in the frontal, parietal, and temporal lobes.[5,27] There is even evidence suggesting that emerging Aβ pathology in cognitively healthy elderly is particularly associated with frontoparietal atrophy, while atrophy in the temporal lobe structures accelerates later as clinical symptoms begin to manifest.[29] The mechanisms behind the regional differences in Aβ-associated atrophy are not well-understood: Potential mechanisms include differences in the afferent and efferent connections and vulnerability to Aβ-related toxicity.

Earlier studies have shown that the *APOE ε4* allele is associated with CSF Aβ levels and changes in brain GM and WM in mild cognitive impairment and AD.[30-32] Although regional reductions in WM volumes were found in *APOE ε4* carriers compared with noncarriers, the *APOE ε4* carrier status had no effect on the relationship between ^{11}C PIB uptake and GM volume, corroborating the results from previous studies conducted in clinically healthy elderly and subjects with mild cognitive impairment.[33,34] These findings are in line with evidence showing that while *APOE ε4* carrier status has a major effect on Aβ deposition, the effects on atrophy are subtle and are mediated by both Aβ-dependent and Aβ-independent mechanisms.[35]

Finally, this study has a few limitations. First, because the subjects were clinically selected to represent an elderly population with several risk factors for dementia, it is likely that they have mixed pathologies; conversely, some of the pathologies related to cognitive impairment may not have been considered in the current study. Second, the sample size was relatively small for the evaluation of associations between GM and WM volumes and clinical variables and the effects of *APOE ε4* on the relationship between ^{11}C PIB uptake and GM and WM volumes. Third, the cross-sectional data do not allow determining whether the subjects with PIB+ findings with impending temporal lobe atrophy develop cognitive impairment later on. Therefore, replication of these findings in larger samples as well as longitudinal studies are needed to determine the predictive power of ^{11}C PIB PET and GM volume in cognitive impairment in at-risk elderly individuals.

CONCLUSIONS

Our results show that elderly individuals who are at increased risk of cognitive decline based on cardiovascular risk factors and have PET scans positive for ^{11}C PIB exhibit reductions in regional GM volume in proportion to increases in brain Aβ load. Our findings are consistent with the model in which brain Aβ accumulation incites neurodegeneration before cognitive decline manifests. Furthermore, the results suggest that the brain Aβ-associated GM loss affects both the medial temporal lobe memory system and the neocortex. Together this evidence emphasizes the importance of finding biomarkers that identify individuals at risk of developing AD pathology who might still benefit from therapeutic interventions.

ACKNOWLEDGMENTS

The assistance of the personnel of the Turku PET Centre in acquiring PET and MR imaging data is gratefully acknowledged.

Disclosures: Ilkka Martikainen—*RELATED*: *Grant*: Finnish Governmental Research Funding for Tampere University Hospital.* Nina Kemppainen—*RELATED*: *Grant*: Turku University Hospital, the Finnish Medical Foundation, the Sigrid Jusélius Foundation, the Maud Kuistila Foundation, the Paulo Foundation; *UNRELATED*: *Employment*: Turku University Hospital. Hilkka Soininen—*RELATED*: *Grant*: Academy of Finland*; *UNRELATED*: *Board Membership*: AC Immune; *Consultancy*: Merck. Tiia Ngandu—*RELATED*: *Grant*: Finnish Medical Foundation.* Miia Kivipelto—*UNRELATED*: *Payment for Lectures Including Service on Speakers Bureaus*: Nestlé. Juha O. Rinne—*RELATED*: *Grant*: Sigrid Jusélius Foundation, *Comments*: unrestricted academic grant*; *UNRELATED*: *Consultancy*: Clinical Research Services Turku Ltd, *Comments*: fee for serving as a consultant neurologist. *Money paid to the institution.

REFERENCES

1. Sevigny J, Chiao P, Bussière T, et al. **The antibody aducanumab reduces Aβ plaques in Alzheimer's disease.** *Nature* 2016;537:50–56 CrossRef Medline
2. Kemppainen NM, Aalto S, Wilson IA, et al. **PET amyloid ligand [11C] PIB uptake is increased in mild cognitive impairment.** *Neurology* 2007;68:1603–06 CrossRef Medline
3. Archer HA, Edison P, Brooks DJ, et al. **Amyloid load and cerebral atrophy in Alzheimer's disease: an 11C-PIB positron emission tomography study.** *Ann Neurol* 2006;60:145–47 CrossRef Medline
4. Tosun D, Schuff N, Mathis CA, et al; Alzheimer's Disease Neuroimaging Initiative. **Spatial patterns of brain amyloid-beta burden and atrophy rate associations in mild cognitive impairment.** *Brain* 2011;134:1077–88 CrossRef Medline
5. Oh H, Madison C, Villeneuve S, et al. **Association of gray matter atrophy with age, β-amyloid, and cognition in aging.** *Cereb Cortex* 2014;24:1609–18 CrossRef Medline
6. Storandt M, Mintun MA, Head D, et al. **Cognitive decline and brain volume loss as signatures of cerebral amyloid-beta peptide deposition identified with Pittsburgh compound B: cognitive decline associated with Abeta deposition.** *Arch Neurol* 2009;66:1476–81 Medline
7. Chételat G, Villemagne VL, Pike KE, et al; Australian Imaging Biomarkers and Lifestyle Study of Ageing (AIBL) Research Group. **Larger temporal volume in elderly with high versus low beta-amyloid deposition.** *Brain* 2010;133:3349–58 CrossRef Medline
8. Jack CR Jr, Lowe VJ, Weigand SD, et al; Alzheimer's Disease Neuroimaging Initiative. **Serial PIB and MRI in normal, mild cognitive impairment and Alzheimer's disease: implications for sequence of pathological events in Alzheimer's disease.** *Brain* 2009;132:1355–65 CrossRef Medline
9. Insel PS, Mattsson N, Donohue MC, et al. **The transitional association between β-amyloid pathology and regional brain atrophy.** *Alzheimers Dement* 2015;11:1171–79 CrossRef Medline
10. Kivipelto M, Ngandu T, Laatikainen T, et al. **Risk score for the pre-**

diction of dementia risk in 20 years among middle aged people: a longitudinal, population-based study. *Lancet Neurol* 2006;5:735–41 CrossRef Medline
11. Kivipelto M, Solomon A, Ahtiluoto S, et al. **The Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER): study design and progress.** *Alzheimers Dement* 2013; 9:657–65 CrossRef Medline
12. Ngandu T, Lehtisalo J, Solomon A, et al. **A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial.** *Lancet* 2015;385:2255–63 CrossRef Medline
13. Ngandu T, Lehtisalo J, Levälahti E, et al. **Recruitment and baseline characteristics of participants in the Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER)-a randomized controlled lifestyle trial.** *Int J Environ Res Public Health* 2014;11:9345–60 CrossRef Medline
14. Harrison J, Minassian SL, Jenkins L, et al. **A neuropsychological test battery for use in Alzheimer disease clinical trials.** *Arch Neurol* 2007; 64:1323–29 CrossRef Medline
15. De la Vega FM, Lazaruk KD, Rhodes MD, et al. **Assessment of two flexible and compatible SNP genotyping platforms: TaqMan SNP Genotyping Assays and the SNPlex Genotyping System.** *Mutat Res* 2005;573:111–35 CrossRef Medline
16. Snellman A, Rokka J, López-Picón FR, et al. **Applicability of [(11)C]PIB micro-PET imaging for in vivo follow-up of anti-amyloid treatment effects in APP23 mouse model.** *Neurobiol Aging* 2017; 57:84–94 CrossRef Medline
17. Schulz V, Torres-Espallardo I, Renisch S, et al. **Automatic, three-segment, MR-based attenuation correction for whole-body PET/MR data.** *Eur J Nucl Med Mol Imaging* 2011;38:138–52 CrossRef Medline
18. Fischl B, Salat DH, Busa E, et al. **Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain.** *Neuron* 2002;33:341–55 CrossRef Medline
19. Ashburner J, Friston KJ. **Unified segmentation.** *Neuroimage* 2005;26: 839–51 CrossRef Medline
20. Hayasaka S, Phan KL, Liberzon I, et al. **Nonstationary cluster-size inference with random field and permutation methods.** *Neuroimage* 2004;22:676–87 CrossRef Medline
21. Fazekas F, Chawluk JB, Alavi A, et al. **MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging.** *AJR Am J Roentgenol* 1987;149:351–56 CrossRef Medline
22. Jack CR Jr, Lowe VJ, Senjem ML, et al. **11C PiB and structural MRI provide complementary information in imaging of Alzheimer's disease and amnestic mild cognitive impairment.** *Brain* 2008;131: 665–80 CrossRef Medline
23. Mormino EC, Kluth JT, Madison CM, et al; Alzheimer's Disease Neuroimaging Initiative. **Episodic memory loss is related to hippocampal-mediated beta-amyloid deposition in elderly subjects.** *Brain* 2009;132:1310–23 CrossRef Medline
24. Bourgeat P, Chételat G, Villemagne VL, et al; AIBL Research Group. **Beta-amyloid burden in the temporal neocortex is related to hippocampal atrophy in elderly subjects without dementia.** *Neurology* 2010;74:121–27 CrossRef Medline
25. Chételat G, Villemagne VL, Bourgeat P, et al; Australian Imaging Biomarkers and Lifestyle Research Group. **Relationship between atrophy and beta-amyloid deposition in Alzheimer disease.** *Ann Neurol* 2010;67:317–24 CrossRef Medline
26. Mormino EC, Betensky RA, Hedden T, et al. **Synergistic effect of β-amyloid and neurodegeneration on cognitive decline in clinically normal individuals.** *JAMA Neurol* 2014;71:1379–85 CrossRef Medline
27. Dickerson BC, Bakkour A, Salat DH, et al. **The cortical signature of Alzheimer's disease: regionally specific cortical thinning relates to symptom severity in very mild to mild AD dementia and is detectable in asymptomatic amyloid-positive individuals.** *Cereb Cortex* 2009;19:497–510 CrossRef Medline
28. Lerch JP, Pruessner JC, Zijdenbos A, et al. **Focal decline of cortical thickness in Alzheimer's disease identified by computational neuroanatomy.** *Cereb Cortex* 2005;15:995–1001 CrossRef Medline
29. Mattsson N, Insel PS, Nosheny R, et al. **Emerging β-amyloid pathology and accelerated cortical atrophy.** *JAMA Neurol* 2014;71:725–34 CrossRef Medline
30. Galasko D, Chang L, Motter R, et al. **High cerebrospinal fluid tau and low amyloid beta42 levels in the clinical diagnosis of Alzheimer disease and relation to apolipoprotein E genotype.** *Arch Neurol* 1998; 55:937–45 CrossRef Medline
31. Honea RA, Vidoni E, Harsha A, et al. **Impact of APOE on the healthy aging brain: a voxel-based MRI and DTI study.** *J Alzheimers Dis* 2009;18:553–64 CrossRef Medline
32. Liu Y, Paajanen T, Westman E, et al; AddNeuroMed Consortium. **Effect of APOE ε4 allele on cortical thicknesses and volumes: the AddNeuroMed study.** *J Alzheimers Dis* 2010;21:947–66 CrossRef Medline
33. Becker JA, Hedden T, Carmasin J, et al. **Amyloid-β associated cortical thinning in clinically normal elderly.** *Ann Neurol* 2011;69: 1032–42 CrossRef Medline
34. Falahati F, Ferreira D, Muehlboeck JS, et al. **Monitoring disease progression in mild cognitive impairment: associations between atrophy patterns, cognition, APOE and amyloid.** *Neuroimage Clin* 2017; 16:418–28 CrossRef Medline
35. Chételat G, Fouquet M. **Neuroimaging biomarkers for Alzheimer's disease in asymptomatic APOE4 carriers.** *Rev Neurol (Paris)* 2013; 169:729–36 CrossRef Medline

Defining the Normal Dorsal Contour of the Corpus Callosum with Time

K.L. Krause, D. Howard, D.R. Pettersson, S. Elstrott, D. Ross, J.T. Obayashi, R. Barajas Jr, A. Bonde, and J.M. Pollock

ABSTRACT

BACKGROUND AND PURPOSE: Morphological changes of the corpus callosum have been associated with a large number of congenital neurocognitive and psychiatric disorders. Focal defects or notches of the dorsal surface of the corpus callosum have not been well characterized. Our purpose was the following; 1) to characterize the dorsal contour of the corpus callosum during the life span, 2) to characterize the relationship of contour deviations to neighboring vessels, and 3) to determine whether contour deviations are congenital or acquired.

MATERIALS AND METHODS: We retrospectively reviewed normal sagittal T1-weighted brain MR images. A "notch" was defined as a concavity in the dorsal surface at least 1 mm in depth. The corpus callosum was considered to be "undulating" if there were >2 notches, including an anterior and posterior notch. The presence of a pericallosal artery and its relationship to a notch were assessed.

RESULTS: We reviewed 1639 MR imaging studies, spanning 0–89 years of age. A total of 1102 notches were identified in 823 studies; 344 (31%) were anterior, 660 (60%) were posterior, and 98 (9%), undulating. There was a positive correlation between the prevalence ($P < .001$) and depth ($P = .028$) of an anterior notch and age and a negative correlation between the prevalence of a posterior notch and age ($P < .001$). There was no difference between patient sex and corpus callosum notching ($P = .884$). Of the 823 studies with notches, 490 (60%) were associated with a pericallosal artery ($P < .001$).

CONCLUSIONS: The prevalence and depth of notches in the anterior corpus callosum increase significantly with age; this finding suggests that most notches are acquired. There is a significant positive association between the presence of a corpus callosum notch and adjacent pericallosal arteries, suggesting that this may play a role in notch formation.

ABBREVIATIONS: CC = corpus callosum; FAS = fetal alcohol syndrome

The corpus callosum (CC) is the midline commissural white matter tract that connects the 2 cerebral hemispheres. It is located at the depth of the interhemispheric fissure and forms the roof of the lateral ventricles. It is divided into 5 named sections from anterior to posterior: rostrum, genu, body, isthmus, and splenium.[1,2] The CC is the largest white matter tract in the brain, with approximately 180 million callosal fibers,[3] which allow sophisticated interhemispheric communication. With such physiologic importance, the CC has a blood supply from both the anterior and posterior circulation by way of a number of different pericallosal vessels.[4,5]

The MR imaging characteristics of the CC and neighboring vessels have been well-described[2] and are becoming more detailed as MR imaging technologies continue to advance. In turn, a number of reports also describe the MR imaging characteristics of a wide variety of congenital and acquired CC pathologies, such as fetal alcohol spectrum disorders.[1,2,6-11] Thus, the intrinsic morphology of the CC has become the subject of intense investigation and debate in its association with multiple cognitive and neuropsychiatric disorders.[9,12-14]

The MR imaging characteristics of the dorsal callosal contour and its intimate association with pericallosal blood vessels have not been well studied or described in the literature. Given that the morphology of the CC is so carefully scrutinized in a number of pathologic conditions, our objective was to investigate the normal dorsal morphology of the CC during the human life span, its anatomic relationship to neighboring vessels, and whether any contour deviation is a congenital or acquired etiology.

MATERIALS AND METHODS

Study Population

This retrospective, single-center study was approved by the institutional review board with a waiver of patient consent. Initially,

Received May 21, 2018; accepted after revision October 6.

From the Departments of Radiology (D.H., D.R.P., S.E., R.B., A.B., J.M.P.) and Neurological Surgery (K.L.K., D.R., J.T.O.), Oregon Health & Science University, Portland, Oregon.

Please address correspondence to Jeffrey M. Pollock MD, Department of Radiology, Oregon Health & Science University, 3181 SW Sam Jackson Park Rd, CR 135, Portland, OR 97239; e-mail: pollockj@ohsu.edu

http://dx.doi.org/10.3174/ajnr.A5886

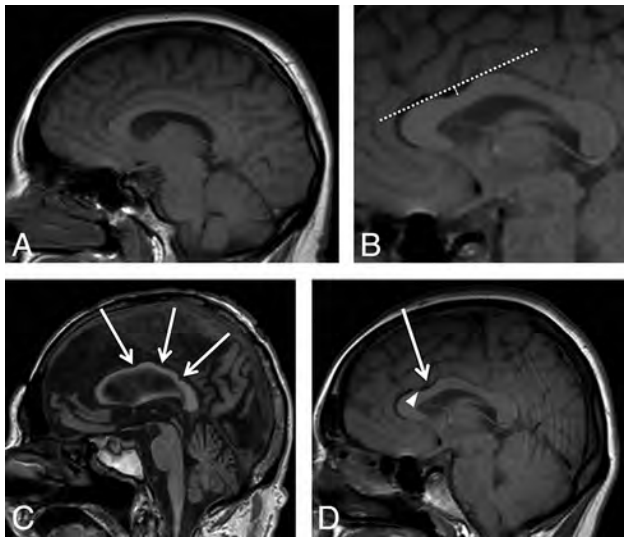


FIG 1. *A*, Midline sagittal T1-weighted image shows an example of a normal corpus callosum without evidence of anterior or posterior notching. *B*, A "notch" was defined as a depression in the dorsal surface whose depth was at least 1 mm from a tangential (*dashed*) line to the surface of the CC. The *solid arrow line* shows a 3-mm-deep anterior notch. *C*, Three areas of notching indicated by the *white arrows* correspond to an undulating configuration. *D*, Midline sagittal T1-weighted image shows that a 3-mm anterior notch (*arrow*) is present. The flow void from the pericallosal artery is visualized extending into the notch (*arrowhead*).

3800 consecutive brain MR imaging examinations with reported impressions containing the keywords "normal" or "no acute process" were reviewed. These data were then screened for all radiographically normal study findings, spanning patient ages 1 day (0 years) to 90 years, that contained a diagnostic-quality midline sagittal T1-weighted image. Studies were excluded for any pathologic process, including but not limited to developmental anomalies, hydrocephalus, intracranial hemorrhage, infarction, encephalomalacia, demyelinating disease, trauma, malignancy, any surgical intervention, or motion-limited examinations.

A team of neuroradiologists and neurosurgeons then visually inspected the dorsal surface of each CC on a midline sagittal T1-weighted image. A CC was designated "normal" if the dorsal surface contained no notching (Fig 1*A*). A "notch" was defined as a focal concavity in the dorsal surface the depth of which was at least 1 mm from a tangential line to the surface of the CC (Fig 1*B*). The notch was noted to be in either the anterior or posterior half of the body of the CC. The CC was considered "undulating" if there were >2 notches, including both an anterior and posterior notch (Fig 1*C*). The depth and location of the notch was recorded. Each image was also examined for the presence of a visible pericallosal artery flow void and whether the artery was seen within the notch (Fig 1*D*).

Statistical Analysis

Subjects were sorted by age and grouped into 5-year increments, starting at age zero. A Poisson regression model was used to determine whether the prevalence of a notch changes with time, and a Spearman ρ correlation was used to determine whether the depth of a notch changes with time. A Student *t* test was used to determine whether patient sex affects the depth of anterior and posterior notches. χ^2 tests were used to determine whether sex or the presence of a blood vessel significantly affects the presence of a notch. Logistic regression was used to determine whether age is a significant predictor variable for undulation, with undulation as a binary dependent variable. Logistic regression was also used to determine whether age is a significant predictor variable for the occurrence of anterior and posterior notches. Statistical analysis was undertaken using SPSS Statistics, Version 24 (IBM, Armonk, New York), and *P* values were considered significant at <.05.

Retrospective review summary of brain MR imaging studies, notch count, and position

Review Summary	
Brain MR imaging studies reviewed (No.)	1639
Brain MR imaging studies, with notches (No.)	823
Notches identified in 823 MR imaging studies (No.)	1102
Female (No.) (%)	919; 56
Notch identified	460; 50
Male (No.) (%)	720; 44
Notch identified	363; 50
Anterior position (No.) (%)	344; 31
Posterior position (No.) (%)	660; 60
Undulating, >2 notches (No.) (%)	98; 9
Pericallosal artery notch association (No.) (%)	490; 60

RESULTS

Retrospective review revealed 1639 unique brain MR imaging studies that met study inclusion criteria; 919 (56%) were female patients and 720 (44%) were male. Ages ranged from 0 to 89 years. There was a total of 1102 notches identified in 823 MR imaging studies; 344 (31%) were located in the anterior half, and 660 (60%) were in the posterior half of the corpus callosum. There were 98 (9%) studies that demonstrated an undulating pattern. From the total of 919 females and 720 males, 460 females (50%) and 363 males (50%) demonstrated evidence of a notch. There was no significant difference between sex and the prevalence of a notch ($P = .884$) (Table).

The prevalence of an anterior or posterior notch was significantly affected by age (Fig 2). There was a significant positive correlation between the prevalence of an anterior notch and advancing age ($P < .001$, OR = 1.095; 95% CI, 1.067–1.123), so the odds of an anterior notch increased by 1.1 times for each age category. Conversely, there was a significant negative correlation between the prevalence of a posterior notch and age ($P < .001$, OR = 0.967; 95% CI, 0.948–0.987), so the odds of a posterior notch decreased by 0.967 for each age category. Age was also found to be a statistically significant predictor of undulation of the corpus callosum ($P = .002$, OR = 1.106; 95% CI, 1.008–1.025). We then specifically examined the 0- to 1-year of age subpopulation cohort ($n = 24$) to determine the earliest prevalence of notches. This subpopulation had a 12% prevalence of an anterior notch and a 21% prevalence of a posterior notch.

Age had a variable effect on the depth of the notch (Fig 3). The average depth of the anterior notch over all age groups was 1.5 mm, and that of the posterior notch was 1.7 mm. The depth of the anterior notch significantly increased [$Rs(345) = 0.118$, $P = .028$] with increasing age. In contrast, age had no significant effect on the depth of the posterior notch [$Rs(658) = 0.019$, $P = .622$]. Sex also did not affect the depth of either the anterior [$t(343) = 1.613$, $P = .147$] or posterior [$t(656) = -0.253$, $P = .330$] notch.

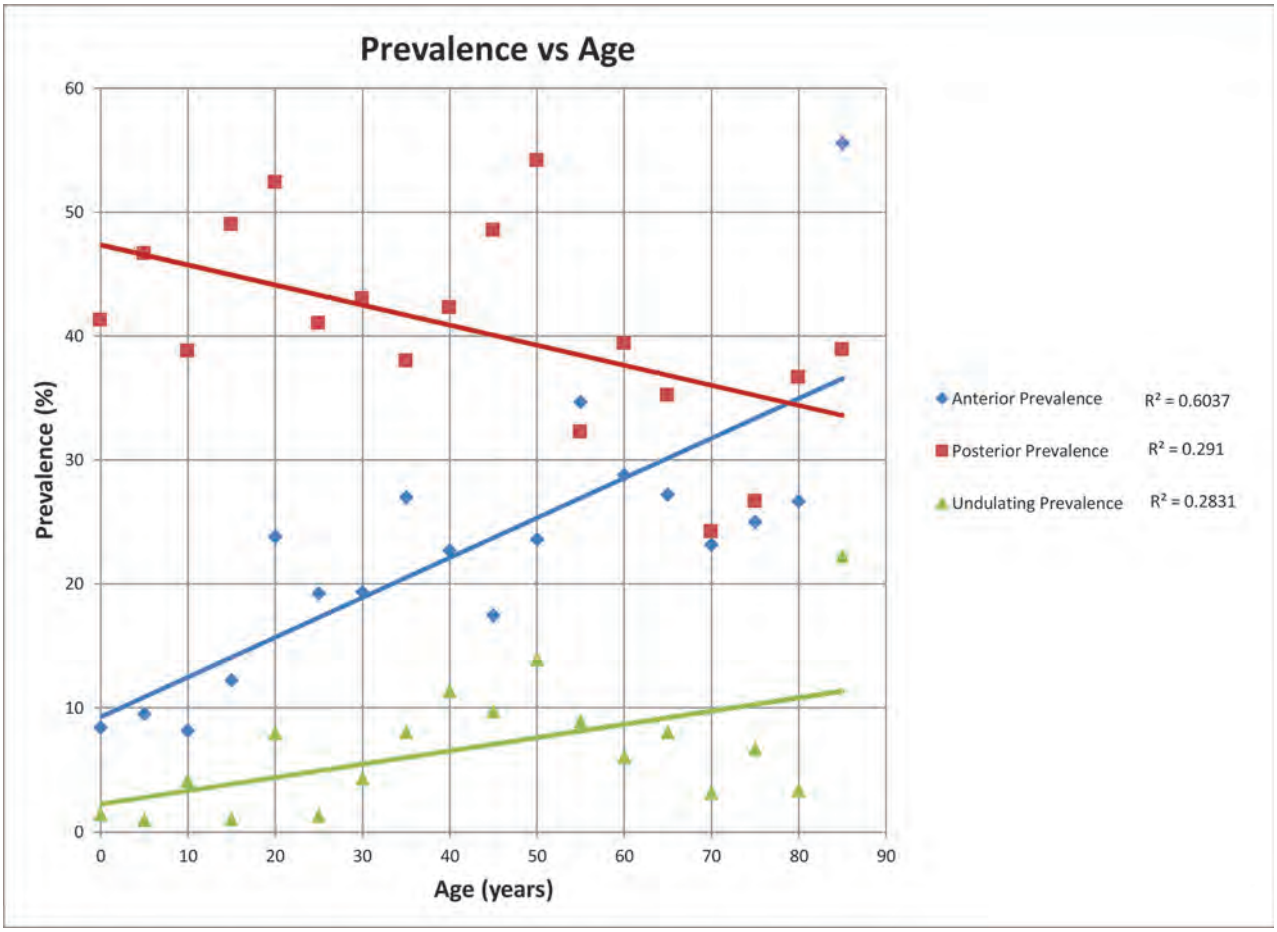


FIG 2. Scatterplot of prevalence versus age in 5-year increments. The posterior notch prevalence (*red squares* and *red line*) decreases with time, $R^2 = 0.29$. The anterior notch prevalence (*blue diamonds* and *blue line*) increases with time, $R^2 = 0.60$. The undulating pattern prevalence (*green triangles* and *green line*) also increases with time, $R^2 = 0.28$.

There were 604 (37%) studies in which a pericallosal artery flow void was identified. In 490 of the 823 (60%) studies in which a CC notch was present, a pericallosal artery flow void was significantly associated with the presence of a notch ($P < .001$) (Table).

DISCUSSION

To our knowledge, this is the first study describing the MR imaging characteristics of the dorsal CC surface across the human life span. Here we demonstrate that notching of the dorsal CC surface is present in infancy in up to 21% of individuals, suggesting that it can be a physiologic anatomic feature. This study shows that the strongest correlation is between an anterior notch prevalence and increasing age. The posterior notch and undulating pattern had a moderate correlation. These dynamic age-related changes suggest that notching is an acquired anatomic feature in most individuals.

It has been previously demonstrated that the morphology of the CC does not remain static with time. Tanaka-Arakaw et al[15] reported that the callosal growth from 0 to 25 years of age is nonlinear; there is an initial marked growth in the total cross-sectional area in the first few months to years of life, which then plateaus during later childhood and adolescence. Then, after 16 years of age, the total area of the CC progressively declines throughout the life span.[16-18] Furthermore, the individual components of the CC develop at different rates.[2] There is an initial growth of the genu beginning at birth to 5–6 years, followed by subsequent growth of the splenium until 10–12 years.[19,20] Although debated in the literature, these changes do not seem to be sex-dependent. Multiple studies have demonstrated that there is not a consistent statistical change in the absolute callosal area between men and women throughout the life span.[15,16,20-23]

Our study also supports the dynamic evolution of the CC with time. Our data suggest that as patients age, the prevalence and depth of an anterior notch in the dorsal callosal surface increase significantly. The depth of the anterior notch is most shallow during childhood and adolescence and then gradually increases with advancing age. It seems conceivable that as the CC ages and its midline cross-sectional area diminishes, the notch depth and thus its prominence become more obvious.

In contrast, however, the prevalence of the posterior notch decreased with age in our study, and the depth did not significantly differ with time. It is known that in the older population, the total white matter volume and the corpus callosum cross-sectional area decrease across time.[16-18] Concurrently, there is also a compensatory increase in the ventricular volume from brain atrophy, leading to a more rounded contour and relative flattening of the posterior callosal surface compared with the younger population.[24] Hypothetically, this flattening of the CC

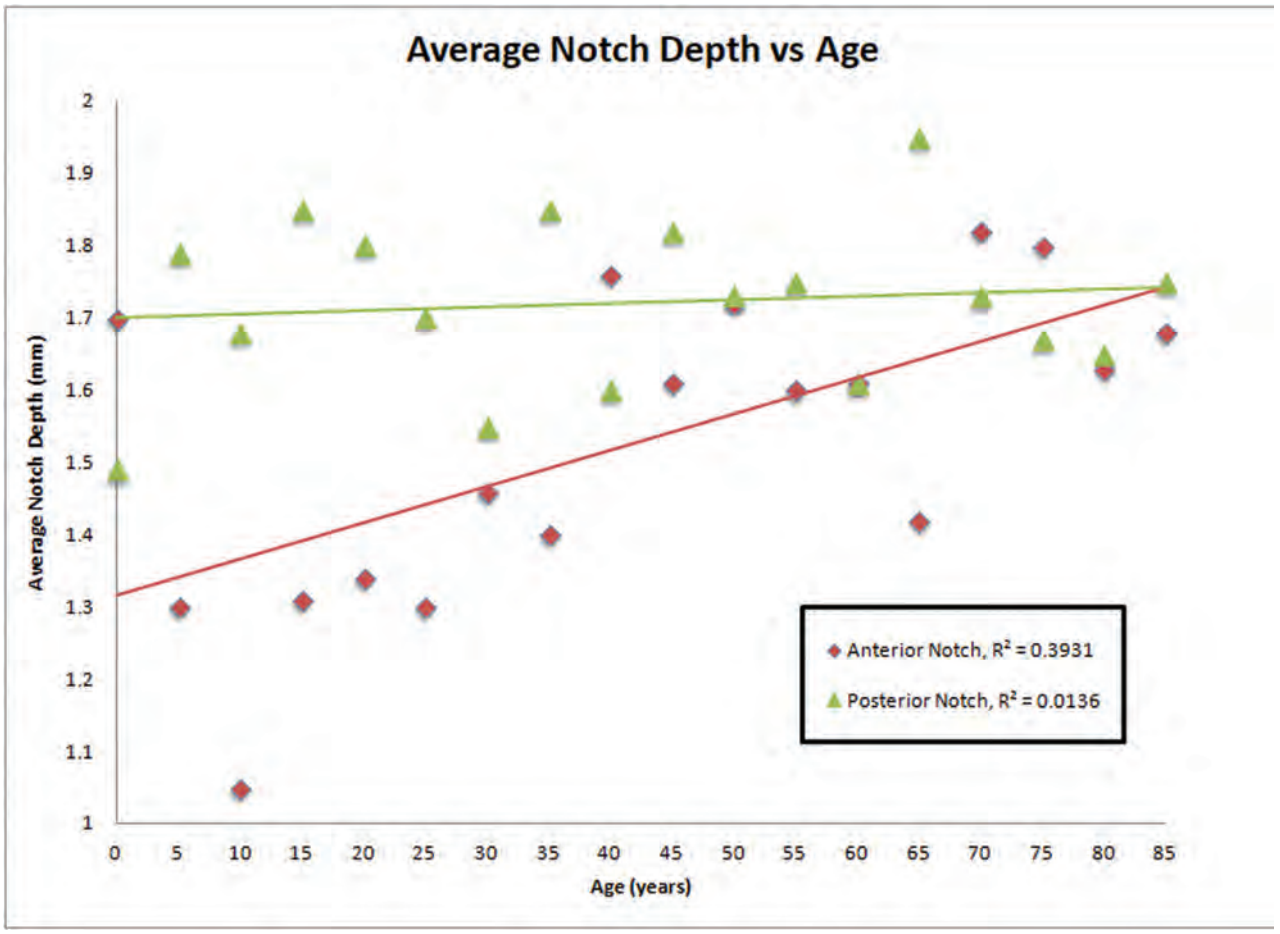


FIG 3. Scatterplot of notch depth (millimeters) versus age. No significant difference is seen in the depth of the posterior notch with increasing age ($R^2 = 0.01$). There is a significant increase in the depth of the anterior notch with increasing age ($R^2 = 0.39$).

with time could account for the decreasing prevalence of a posterior notch with increasing age. The posterior notch in the younger population was frequently a manifestation of having a thicker splenium than in the older population. In these younger patients, the thicker splenium created a measurable notch based on our methodology, which may also account for the decreasing prevalence with increasing age. Alternatively, the decrease in notching posteriorly with age might be due to atrophy of the posterior body and splenium of the CC due to biparietal pathologies.

We found that the prevalence of an anterior notch at any age is significantly correlated with the presence of a neighboring blood vessel. The CC is supplied by an extensive network of anastomosing vessels between the anterior and posterior cerebral arteries. The rostrum, genu, and body are mostly supplied by the pericallosal artery, a branch of the anterior cerebral artery. As demonstrated in a gross anatomic study by Türe et al,[4] in 60% of specimens, the pericallosal artery was entirely within the callosal sulcus, coursing along the superficial dorsal aspect of the CC. Another 30% of specimens demonstrated vessels with an irregular course through the callosal sulcus, and only 10% were not associated with the CC at all. The terminal segment of the pericallosal artery then anastomoses with the posterior pericallosal artery, a branch of the posterior cerebral artery, which supplies most of the splenium.

Given this extensive, intimate association between the pericallosal vasculature and dorsal callosal surface, our data suggest that the etiology of a callosal notch on MR imaging may be an effect of the pericallosal artery either directly indenting or focally obscuring the CC along its course or falling into an adjacent notch. Arteries are well-known for distorting adjacent white matter structures such as the pons and medulla in cases of dolichoectasia of the vertebral artery.[25] It seems plausible that with time, hypertrophy or increased tortuosity of the pericallosal artery could similarly distort the contour of the adjacent CC surface. Increased tortuosity of the vessel with time could account for the increasing prevalence of undulation of the dorsal surface as the CC accumulates additional notches. This finding may be compounded by normal callosal atrophy with aging. As the cross-sectional area of the CC diminishes with time, the vessels lose their normal support from the callosal surface. This, in turn, could lead to a vessel "falling" into an adjacent notch or "draping" of pericallosal vessels over the CC, further providing the appearance or deepening of a callosal notch on MR imaging. Similarly, the increase in anterior notching prevalence with age might be due to the much larger diameters of the proximal A2 vessels than the smaller, less pulsatile posterior vessels.

Describing the dorsal contour of the CC with time in a large population can help us better classify findings as normal, abnormal, acquired, or congenital. The morphology of the CC has been suggested to play a role in a number of different

neurocognitive and psychiatric disorders. In particular, the CC has been widely studied in fetal alcohol syndrome (FAS),[9,13,14,26,27] which affects approximately 2%–5% of the population in the United States.[28,29] An article by Riley et al,[9] from 1995, postulated that focal abnormalities or what we refer to as "notches" of the CC might be secondary to insults that occur after its complete formation. The Riley article itself does not address or describe these focal abnormalities but, ironically, remains highly cited in the medicolegal literature for establishing focal abnormalities of the CC as part of the FAS spectrum.[9] When we further analyzed our patients younger than 1 year of age to determine the percentage of our population that might have congenital notches, there was a 12% anterior notch prevalence and a 21% posterior notch prevalence. The discrepancy between the low prevalence of FAS in the population (2%–5%) and high prevalence of notches of the CC in our earliest age group (12%–21%) suggests that notching is unlikely to be directly associated with FAS. However, the significance of the prevalence discrepancy may be limited by the small sample size of patients younger than 1 year of age ($n = 24$) in our study and thus may warrant further investigation to more accurately define these percentages.

It follows that this study is intrinsically limited by its narrow focus of only examining the dorsal surface of the CC. Furthermore, this was not a longitudinal study because single patients were not followed across time to track their potential development of notches. Additionally, this patient population represents a regional sample, and a similar analysis in a different region might result in different values. This is a purely radiographic study, and although each subject was examined for any intracranial pathology, we do not have any data that describe the clinical functionality of these patients. Other more global observations that may influence CC morphology, such as callosal area or total cerebral volume, were not quantified. Furthermore, we only measured the absolute depth of the notch but did not measure the depth relative to the thickness of the CC. Studies containing small-vessel ischemic changes or studies with generalized cerebral atrophy were not excluded. These are both states that could influence CC volume. Finally, the study relied on flow voids to locate the pericallosal artery and did not have dedicated angiographic correlates. These could be areas of future study.

CONCLUSIONS

Analysis of >1600 brain MR imaging studies suggests that the presence of a dorsal callosal notch can be a common anatomic variant, affecting 50% of both men and women. The correlation of a posterior notch and an undulating pattern with age was moderate, while there was a strong correlation of anterior notch prevalence with age. The prevalence and depth of a notch are dynamic phenomena that change throughout the life span, providing evidence that notching can also be an acquired finding that, at least in some cases, is the result of normal aging. Furthermore, the presence of a notch can also be associated with a neighboring pericallosal artery.

ACKNOWLEDGMENTS

We thank Shirley McCartney PhD for her assistance with manuscript preparation and formatting.

Disclosures: Ramon Barajas—*UNRELATED*: *Grant*: National Cancer Institute, *Comments*: 3 R01 CA137488–22S1.* *Money paid to the institution.

REFERENCES

1. Battal B, Kocaoglu M, Akgun V, et al. **Corpus callosum: normal imaging appearance, variants and pathologic conditions.** *J Med Imaging Radiat Oncol* 2010;54:541–49 CrossRef Medline
2. Georgy BA, Hesselink JR, Jernigan TL. **MR imaging of the corpus callosum.** *AJR Am J Roentgenol* 1993;160:949–55 Medline
3. Devinsky O, Laff R. **Callosal lesions and behavior: history and modern concepts.** *Epilepsy Behav* 2003;4:607–17 CrossRef Medline
4. Türe U, Yaşargil MG, Krisht AF. **The arteries of the corpus callosum: a microsurgical anatomic study.** *Neurosurgery* 1996;39:1075–84; discussion 1084–85 CrossRef Medline
5. Kahilogullari G, Comert A, Ozdemir M, et al. **Arterial vascularization patterns of the splenium: an anatomical study.** *Clin Anat* 2013; 26:675–81 CrossRef Medline
6. Bourekas EC, Varakis K, Bruns D, et al. **Lesions of the corpus callosum: MR imaging and differential considerations in adults and children.** *AJR Am J Roentgenol* 2002;179:251–57 CrossRef Medline
7. Byard RW. **The corpus callosum and forensic issues: an overview.** *J Forensic Sci* 2016;61:979–83 CrossRef Medline
8. Renard D, Castelnovo G, Campello C, et al. **An MRI review of acquired corpus callosum lesions.** *J Neurol Neurosurg Psychiatry* 2014; 85:1041–48 CrossRef Medline
9. Riley EP, Mattson SN, Sowell ER, et al. **Abnormalities of the corpus callosum in children prenatally exposed to alcohol.** *Alcohol Clin Exp Res* 1995;19:1198–202 Medline
10. Uchino A, Takase Y, Nomiyama K, et al. **Acquired lesions of the corpus callosum: MR imaging.** *Eur Radiol* 2006;16:905–14 Medline
11. Hyun Yoo J, Hunter J. **Imaging spectrum of pediatric corpus callosal pathology: a pictorial review.** *J Neuroimaging* 2013;23:281–95 CrossRef Medline
12. Coger RW, Serafetinides EA. **Schizophrenia, corpus callosum, and interhemispheric communication: a review.** *Psychiatry Res* 1990;34: 163–84 CrossRef Medline
13. Donald KA, Eastman E, Howells FM, et al. **Neuroimaging effects of prenatal alcohol exposure on the developing human brain: a magnetic resonance imaging review.** *Acta Neuropsychiatr* 2015;27: 251–69 CrossRef Medline
14. Norman AL, Crocker N, Mattson SN, et al. **Neuroimaging and fetal alcohol spectrum disorders.** *Dev Disabil Res Rev* 2009;15:209–17 CrossRef Medline
15. Tanaka-Arakawa MM, Matsui M, Tanaka C, et al. **Developmental changes in the corpus callosum from infancy to early adulthood: a structural magnetic resonance imaging study.** *PLoS One* 2015;10: e0118760 CrossRef Medline
16. Allen LS, Richey MF, Chai YM, et al. **Sex differences in the corpus callosum of the living human being.** *J Neurosci* 1991;11:933–42 CrossRef Medline
17. Doraiswamy P, Figiel G, Husain M, et al. **Aging of the human corpus callosum: magnetic resonance imaging in normal volunteers.** *J Neuropsychiatry Clin Neurosci* 1991;3:392–97 CrossRef Medline
18. Hayakawa K, Konishi Y, Matsuda T, et al. **Development and aging of brain midline structures: assessment with MR imaging.** *Radiology* 1989;172:171–77 CrossRef Medline
19. Knyazeva MG. **Splenium of corpus callosum: patterns of interhemispheric interaction in children and adults.** *Neural Plast* 2013;2013: 1–12 CrossRef Medline
20. Vannucci RC, Barron TF, Vannucci SJ. **Development of the corpus callosum: an MRI study.** *Dev Neurosci* 2017;39:97–106 CrossRef Medline

21. Rajapakse JC, Giedd JN, Rumsey JM, et al. **Regional MRI measurements of the corpus callosum: a methodological and developmental study.** *Brain Dev* 1996;18:379–88 CrossRef Medline
22. Jäncke L, Staiger JF, Schlaug G. et al. **The relationship between corpus callosum size and forebrain volume.** *Cereb Cortex* 1997;7:48–56 CrossRef Medline
23. Kertesz A, Polk M, Howell J, et al. **Cerebral dominance, sex, and callosal size in MRI.** *Neurology* 1987;37:1385–88 Medline
24. Scahill RI, Frost C, Jenkins R, et al. **A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging.** *Arch Neurol* 2003;60:989–94 CrossRef Medline
25. Resta M, Gentile MA, Di Cuonzo F, et al. **Clinical-angiographic correlations in 132 patients with megadolichovertebrobasilar anomaly.** *Neuroradiology* 1984;26:213–16 CrossRef Medline
26. Sowell ER, Mattson SN, Thompson PM, et al. **Mapping callosal morphology and cognitive correlates: effects of heavy prenatal alcohol exposure.** *Neurology* 2001;57:235–44 CrossRef Medline
27. Yang Y, Phillips OR, Kan E, et al. **Callosal thickness reductions relate to facial dysmorphology in fetal alcohol spectrum disorders.** *Alcohol Clin Exp Res* 2012;36:798–806 CrossRef Medline
28. May PA, Baete A, Russo J, et al. **Prevalence and characteristics of fetal alcohol spectrum disorders.** *Pediatrics* 2014;134:855–66 CrossRef Medline
29. May PA, Gossage JP, Kalberg WO, et al. **Prevalence and epidemiologic characteristics of FASD from various research methods with an emphasis on recent in-school studies.** *Dev Disabil Res Rev* 2009;15:176–92 CrossRef Medline

Compressed Sensing–Sensitivity Encoding (CS-SENSE) Accelerated Brain Imaging: Reduced Scan Time without Reduced Image Quality

J.E. Vranic, N.M. Cross, Y. Wang, D.S. Hippe, E. de Weerdt, and M. Mossa-Basha

ABSTRACT

BACKGROUND AND PURPOSE: Compressed sensing–sensitivity encoding is a promising MR imaging acceleration technique. This study compares the image quality of compressed sensing–sensitivity encoding accelerated imaging with conventional MR imaging sequences.

MATERIALS AND METHODS: Patients with known, treated, or suspected brain tumors underwent compressed sensing–sensitivity encoding accelerated 3D T1-echo-spoiled gradient echo or 3D T2-FLAIR sequences in addition to the corresponding conventional acquisition as part of their clinical brain MR imaging. Two neuroradiologists blinded to sequence and patient information independently evaluated both the accelerated and corresponding conventional acquisitions. The sequences were evaluated on 4- or 5-point Likert scales for overall image quality, SNR, extent/severity of artifacts, and gray-white junction and lesion boundary sharpness. SNR and contrast-to-noise ratio values were compared.

RESULTS: Sixty-six patients were included in the study. For T1-echo-spoiled gradient echo, image quality in all 5 metrics was slightly better for compressed sensing–sensitivity encoding than conventional images on average, though it was not statistically significant, and the lower bounds of the 95% confidence intervals indicated that compressed sensing–sensitivity encoding image quality was within 10% of conventional imaging. For T2-FLAIR, image quality of the compressed sensing–sensitivity encoding images was within 10% of the conventional images on average for 3 of 5 metrics. The compressed sensing–sensitivity encoding images had somewhat more artifacts ($P = .068$) and less gray-white matter sharpness ($P = .36$) than the conventional images, though neither difference was significant. There was no significant difference in the SNR and contrast-to-noise ratio. There was 25% and 35% scan-time reduction with compressed sensing–sensitivity encoding for FLAIR and echo-spoiled gradient echo sequences, respectively.

CONCLUSIONS: Compressed sensing–sensitivity encoding accelerated 3D T1-echo-spoiled gradient echo and T2-FLAIR sequences of the brain show image quality similar to that of standard acquisitions with reduced scan time. Compressed sensing–sensitivity encoding may reduce scan time without sacrificing image quality.

ABBREVIATIONS: CNR = contrast-to-noise ratio; CS = compressed sensing; SENSE = sensitivity encoding; SPGR = echo-spoiled gradient echo

The excellent soft-tissue contrast resolution and specialized sequences targeting different aspects of pathophysiology make MR imaging the optimal technique for studying the brain. Despite the many advantages of brain MR imaging, MR imaging acquisition is a time-consuming endeavor compared with CT. Long image-acquisition times limit both the clinical application and practicality of MR imaging, particularly in medically unstable and pediatric patients.

MR imaging acquisition time is largely influenced by the number of data points sampled from *k*-space, the way these data points are sampled, and the way in which image reconstruction is performed. Several image-acquisition and postprocessing techniques have been developed to reduce image-acquisition time while still preserving image quality.[1,2] These include parallel MR imaging and compressed sensing (CS) MR imaging techniques, which rely on different reconstruction constraints to accelerate image production.[3] Combining these techniques can lead to image-acquisition acceleration factors that far exceed what is achievable by either parallel or CS MR imaging alone.[3-6] This combined image-acceleration technique is referred to as CS-sensitivity encoding (SENSE) MR

Received May 17, 2018; accepted after revision October 22.

From the Department of Radiology (J.E.V., N.M.C., D.S.H., M.M.-B.), University of Washington, Seattle, Washington; and Philips Healthcare (Y.W., E.d.W.), Best, the Netherlands.

Please address correspondence to Mahmud Mossa-Basha, MD, University of Washington Medical Center, 1959 NE Pacific St, Seattle, WA 98195; e-mail: mmossab@uw.edu; @mossabas

Indicates article with supplemental on-line tables.

Indicates article with supplemental on-line photo.

http://dx.doi.org/10.3174/ajnr.A5905

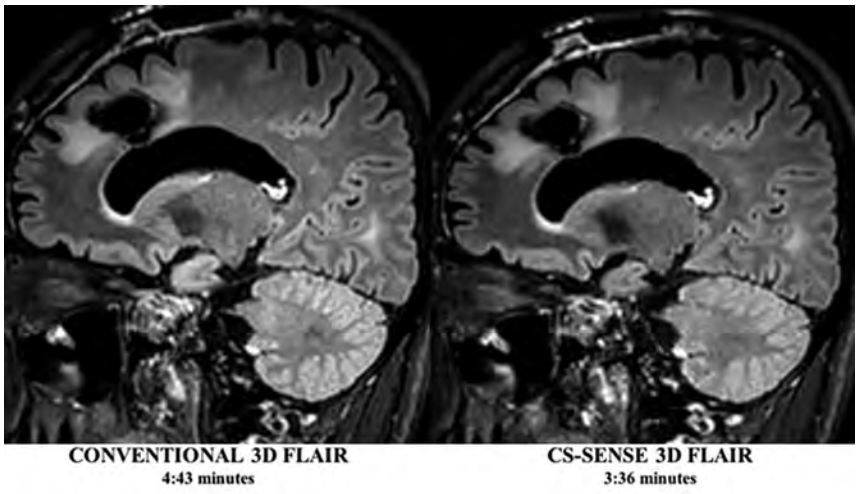


FIG 1. Conventional and CS-SENSE accelerated sagittal 3D T2-FLAIR images from the same patient demonstrate a treated primary brain tumor within the left frontal lobe. Note the sharp borders of the brain parenchymal lesion detected in both images, while CS-SENSE 3D FLAIR (*right*) was acquired with a 25% scan time reduction.

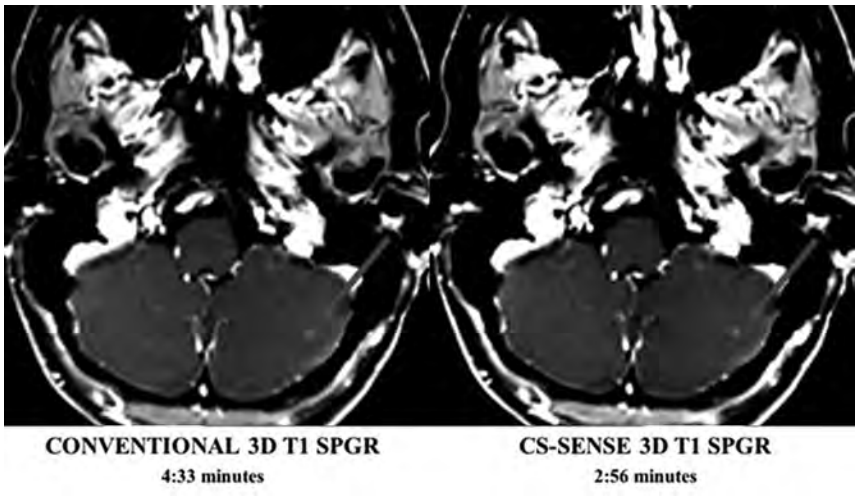


FIG 2. Conventional and CS-SENSE accelerated axial T1-SPGR images are from the same patient. The *arrow* demonstrates a small metastasis within the left cerebellar hemisphere that was detected by both sequences equally well. Acquisition of the CS-SENSE SPGR (*right*) was 35% faster than the conventional SPGR (*left*).

imaging, and it has the potential to dramatically decrease overall imaging times while still preserving image quality.

Despite the many technical advancements that have been made in accelerating MR imaging acquisition and image reconstruction, robust evaluation of these acceleration techniques in clinical practice is still warranted. Clinical verification of the ability of these accelerated image-acquisition techniques to produce diagnostic-quality images of the central nervous system is essential before broader implementation of these imaging techniques into clinical practice can occur. Only a small number of studies have investigated the performance of CS-SENSE MR imaging in limited patient populations as it relates to body imaging.[4,5] Very few studies have evaluated CS in brain MR imaging, with the studies performed focusing on the evaluation of multiple sclerosis lesions on T2-FLAIR,[7] brain MR imaging quality assessment in healthy controls,[8] and evaluating achievable acceleration, reconstruction schemes, and artifacts generated from retrospective CS.[9] To date, however, no one has critically evaluated the clinical performance of the integrated CS-SENSE algorithm for MR imaging applied to imaging of the central nervous system, to our knowledge. In addition, we present the first work to apply CS acceleration in a brain tumor clinical population. The purpose of the current study was to compare the image quality of CS-SENSE accelerated 3D T1-echo-spoiled gradient echo (SPGR) (CS-SENSE SPGR) and T2-FLAIR (CS-SENSE FLAIR) sequences with the corresponding conventional acquisitions. We hypothesized that CS-SENSE accelerated sequences will have image quality equivalent to that of conventional acquisitions while accelerating imaging.

MATERIALS AND METHODS

Patient Selection

With our institutional review board approval and after obtaining informed written consent, adult patients (18 years of age or older) were prospectively scanned between February 8, 2017, and January 19, 2018, for assessment of the MR imaging brain tumor protocols of our institution with inclusion of a conventional sequence and a corresponding CS-SENSE accelerated acquisition when the clinical schedule permitted. CS-SENSE accelerated acquisition was performed before or after the corresponding conventional acquisition in alternating order (to mitigate potential bias from ordering effects), with both sequences performed after gadolinium administration for both T2-FLAIR and SPGR.

Image Acquisition

All imaging was performed on a 3T Ingenia MR imaging scanner (Philips Healthcare, Best, the Netherlands) using a 16-channel head coil (In Vivo, Gainesville, Florida). Each patient underwent the brain tumor imaging protocol of our institution. This included the following sequences: axial DWI, axial T1-spin-echo, sagittal 3D T2-FLAIR with gadolinium, axial 3D T1-SPGR with gadolinium, and coronal and axial T1-spin-echo with gadolinium. In addition to these conventional acquisitions, each patient underwent either a CS-SENSE accelerated 3D T2-FLAIR (Fig 1) or a CS-SENSE accelerated gadolinium-enhanced 3D T1-SPGR sequence (Fig 2), which was performed during their routine MR imaging examination. The sequence scan parameters for both the conventional and CS-SENSE MR imaging sequences are listed in Table 1. The CS-SENSE FLAIR and CS-SENSE SPGR sequences had acceleration factors of 1.3 and 1.7 with scan time reduc-

Table 1: Imaging parameters for CS-SENSE accelerated and conventional MRI sequences

	3D T2-FLAIR	CS-SENSE 3D T2-FLAIR	3D T1-SPGR	CS-SENSE 3D T1-SPGR
FOV (cm)	230 × 2230 × 150	230 × 230 × 150	240 × 240 × 168	240 × 240 × 168
True voxel size (mm^3)	1 × 1 × 2	1 × 1 × 2	1.1 × 1.1 × 1.1	1.1 × 1.1 × 1.1
Interpolated voxel (mm^3)	1 × 1 × 1	1 × 1 × 1	1 × 1 × 0.5	1 × 1 × 0.5
TR/TE (ms)	25/3.6	25/3.6	4800/301	4800/301
Slices (No.)	150	150	300	300
TI (ms)	–	–	1650	1650
SENSE acceleration factor (RL, FH)	(2, 1)	(2, 1)	(2.4, 1.7)	(2.4, 1.7)
CS acceleration factor	–	1.3	–	1.7
Total acceleration factor	2	3.4	4.1	5.3
Scan time (min:sec)	4:33	3:36	4:43	2:56
Scan time reduction	–	25%		35%

Note:—RL indicates Right-to-left; FH, Foot-to-head; –, Not used/no value.

tions of 25% and 35% compared with the conventional acquisition counterparts, respectively. These CS-SENSE accelerated acquisitions used a balanced variable density incoherent undersampling acquisition scheme and iterative reconstruction to solve an inverse problem with a sparsity constraint. Specifically, the images were acquired using a random undersampling pattern with the Poisson disc style distribution. Image reconstruction was performed using a wavelet transform for the sparsity term, according to the common CS and parallel imaging problem definitions. Prior knowledge of noise decorrelation, regularization, and coil sensitivities was used to provide an optimal SNR as a starting point, allowing additional acceleration capabilities via sparsity constraining. The reconstruction algorithm was based on a modified fast iterative shrinkage/soft thresholding algorithm (FISTA) scheme,[10] which entails iterative reconstruction. Conventional clinical acquisition T2-FLAIR and T1-SPGR sequences served as imaging control sequences against which the CS-SENSE FLAIR and SPGR sequences could be compared.

Image Evaluation

Two experienced board-certified neuroradiologists (M.M.-B. and N.M.C.) blinded to the imaging technique and patient clinical information independently evaluated all CS-SENSE and corresponding conventional sequences. All imaging studies were deidentified and randomized so that each rater was unaware of whether they were reviewing a CS-SENSE or conventional acquisition. Raters evaluated overall imaging quality on the following 4-point scale: 1, nondiagnostic; 2, limited but interpretable; 3, minimally limited; and 4, optimal quality. Image SNR was rated on the following 5-point scale: 1, markedly diminished SNR that renders the images uninterpretable; 2, moderately diminished SNR that affects interpretation; 3, diminished SNR that only mildly limits interpretation; 4, mildly diminished SNR that does not affect image interpretation; and 5, optimal SNR. Image artifacts were evaluated on the following 5-point scale: 1, severe image artifacts; 2, moderate artifacts; 3, mild artifacts; 4, trace artifacts; and 5, no artifacts. Brain parenchymal lesion border sharpness was evaluated on the following 3-point scale: 1, a lesion whose borders were indistinguishable from background brain; 2, a lesion with blurry margins; and 3, sharp lesion margins. Finally, gray-white matter sharpness was assessed using the following 3-point scale: 1, indistinguishable gray-white sharpness; 2, blurry gray-white sharpness; and 3, well-defined gray-white sharpness.

In addition to qualitative image evaluation, quantitative evaluation of SNR and the contrast-to-noise ratio (CNR) was also performed for each CS-SENSE and conventional acquisition, respectively. For each patient, ROIs were drawn on representative images from each CS-SENSE and conventional acquisition. ROIs were positioned within the normal subcortical white matter, within a focal brain lesion, and outside the patient, in what was classified as image background. All ROIs were the same size and had nearly identical positioning between sequences. As reported elsewhere in the literature,[7,11] SNR and CNR were calculated as follows: $SNR = SI/SD_{noise}$ and $CNR = (SI_{lesion} - SI_{WM})/SD_{noise}$, where SI is the average signal intensity of the lesion or white matter and SD_{noise} is the SD of noise.

Statistical Analysis

While raters evaluated images using 4- or 5-point scales, they ultimately only used 2–3 levels of each scale, with the middle level being the most common. To improve interpretability, we dichotomized all scales, mainly to get the best possible balance of ratings above and below the threshold. Specifically, image quality was dichotomized as optimal image quality (5 versus 1–4), optimal SNR (5 versus 1–4), no or trace artifacts (4–5 versus 1–3), sharp gray-white matter boundaries (3 versus 1–2), and sharp lesion boundaries (3 versus 1–2).

The percentages for each image-quality metric were compared between CS-SENSE and the corresponding conventional images (FLAIR versus CS-FLAIR, SPGR versus CS-SPGR) using the nonparametric bootstrap to calculate 95% CI and *P* values for the differences. The widths of the 95% CIs were used to help assess a plausible range of differences in image quality between CS-SENSE and the corresponding conventional images. Ratings from both raters were analyzed together for the primary analysis and separately as a sensitivity analysis. Bootstrap resampling was performed by patient to account for the nonindependence of ratings by both raters of the same images and for multiple scans acquired from some patients.

Interrater agreement was assessed using the Cohen κ and by counting how often both raters, 1 rater, and neither rater rated CS-SENSE images at least as highly as conventional images. All statistical calculations were conducted with R statistical and computing software (Version 3.1.1; http://www.r-project.org/). Throughout, 2-tailed tests were used with statistical significance defined as $P < .05$.

Table 2: CS-SENSE and conventional MRI sequence qualitative evaluation

Metric	SAG FLAIR (*n* = 70 Reads)					SPGR (*n* = 68 Reads)				
	Acquisition[a]		Difference			Acquisition[a]		Difference		
	CS-SENSE	Conventional	Mean	(95% CI)	*P* Value	CS-SENSE	Conventional	Mean	(95% CI)	*P* Value
Optimal image quality	19 (27.1)	19 (27.1)	0.0%	(−510.0–10.0)	>.99	34 (50.0)	25 (36.8)	13.2%	(−5.7–31.8)	.15
Optimal SNR	6 (8.6)	5 (7.1)	1.4%	(−55.7–8.6)	.67	27 (39.7)	24 (35.3)	4.4%	(−59.7–18.2)	.55
No or trace artifacts	52 (74.3)	60 (85.7)	−511.4%	(−524.3–0.0)	.068	46 (67.6)	37 (54.4)	13.2%	(−53.0–28.6)	.11
Sharp gray-white matter boundaries	36 (51.4)	39 (55.7)	−54.3%	(−512.9–4.3)	.36	27 (39.7)	26 (38.2)	2.2%	(−58.1–12.2)	.67
Sharp lesion boundaries	57 (86.4)	55 (83.3)	2.3%	(−57.1–11.7)	.65	52 (83.9)	52 (83.9)	0.8%	(−58.6–10.0)	.84

Note:—SAG indicates sagittal.
[a] Values are No. (%) where higher percentages indicate a higher prevalence of good image quality.

Table 3: CS-SENSE and conventional MRI interrater agreement

Metric	SAG FLAIR (*n* = 35 Exams)				AX SPGR (*n* = 34 Exams)			
	CS-SENSE		Conventional		CS-SENSE		Conventional	
	κ	(95% CI)	κ	(95% CI)	κ	(95% CI)	κ	(95% CI)
Optimal image quality	0.20	(0.00–0.43)	0.17	(−0.07–0.43)	0.32	(0.00–0.59)	0.57	(0.26–0.83)
Optimal SNR	0.00	[a]	−0.07	[a]	0.12	(−0.19–0.42)	0.11	(−0.18–0.42)
Trace or no artifacts	0.41	(0.07–0.72)	−0.14	[a]	0.60	(0.26–0.86)	0.50	(0.28–0.74)
Sharp gray-white matter boundaries	0.01	[a]	0.02	(−0.14–0.16)	−0.10	(−0.30–0.04)	−0.04	(−0.19–0.07)
Sharp lesion boundaries	0.18	(0.00–0.53)	0.00	[a]	0.31	(−0.09–0.69)	0.33	(0.00–0.70)

Note:—AX indicates axial; SAG, sagittal.
[a] The 95% confidence interval was not provided when ratings were too uncommon to calculate reliably with the bootstrap.

RESULTS

Patient Data

Sixty-nine patients were reviewed. Three patients were scanned with 1 of the 2 accelerated image-acquisition sequences but were not scanned with the corresponding conventional sequence and were excluded from analysis. This step resulted in a final cohort of 66 patients. Of these 66 patients, 35 patients were imaged with CS-SENSE FLAIR for a total of 35 scans (16 women [45.7%]; 25–74 years of age; median, 47 years). Thirty-one patients were imaged with the CS-SENSE SPGR, with 1 patient being scanned twice and a second patient being scanned 3 times, each at different time points, resulting in a total of 34 CS-SENSE SPGR scans (16 women [51.6%]; 18–83 years of age; median, 63 years). Ninety-five percent of patients included in this cohort demonstrated brain lesions on MR imaging (On-line Table 1). There were a total of 89 and 56 lesions on FLAIR and SPGR, respectively. There were lesions in 34 patients (34 scans) in the FLAIR cohort, 22 of whom had multiple lesions. Lesions were also present in 29 patients (32 scans) in the SPGR cohort, 10 of whom had multiple lesions. Lesion sizes are summarized in On-line Table 2.

Qualitative Image Comparison

Pooled image-quality ratings are summarized in Table 2. For FLAIR, there were no statistically significant differences in overall image quality, SNR, gray-white matter boundary sharpness, or lesion-border sharpness between CS-SENSE and conventional sequences, with the lower bound of the 95% CIs indicating that image quality of the CS-SENSE images was within approximately 10% of the conventional images by these metrics. However, there was a trend toward more artifacts on CS-SENSE compared with conventional images (11.4%, $P = .068$).

For SPGR, there were no significant differences in any image-quality metric between CS-SENSE and conventional SPGR, though CS-SENSE images had slightly higher image-quality ratings on average than the conventional images. By each metric, image-quality ratings of CS-SENSE were within 10% of the conventional SPGR ratings based on the lower bound of the 95% CI. Differences in image quality were most noticeable between the CS-SENSE SPGR and the standard SPGR; 50% of accelerated SPGR studies demonstrated optimal image quality compared with 37% of the standard SPGR acquisitions.

The image-quality results were generally similar when examined by each rater separately with a few notable exceptions. For FLAIR, rater 1 tended to rate gray-white boundary sharpness of CS-SENSE images more highly than on conventional images (+8.6%, $P = .14$), while rater 2 rated CS-SENSE images lower than on conventional images (−17.1%, $P < .001$). For SPGR, rater 1 rated SNR and lack of artifacts in CS-SENSE images more highly than on conventional images (+20.6%, $P = .026$ and +26.5%, $P = .027$, respectively), while rater 2 rated the images more similarly (−11.8%, $P = .27$ and +0.0%, $P > .99$, respectively).

Interrater Agreement

Interrater agreement scores for the CS-SENSE and conventional sequences are listed in Table 3. Interrater agreement for each image quality was mostly poor to fair for FLAIR ($\kappa < 0.4$) but fair to moderate for SPGR ($\kappa = 0.2–0.6$). Despite some differences in absolute ratings, raters both agreed 77%–91% of the time that the image-quality metrics of the CS-SENSE FLAIR were at least as good as the those of conventional images and rarely agreed that the CS-SENSE FLAIR images were worse than conventional images (Table 4). Similarly, raters both agreed 67%–87% of the time that the image-quality metrics of the CS-SENSE SPGR were at least as good as those of the conventional images, while only agreeing 0%–6% of the time that the conventional images were better. In terms of disagreement, there was >1 disagreement on the Likert scale only for artifact severity (On-line Figure), which occurred in 7/69 comparisons. This level of disagreement did not occur for any other qualitative metrics.

Table 4: Interrater agreement on image-quality comparisons of CS-SENSE and conventional MRI[a]

Metric	SAG FLAIR (*n* = 35 Exams)			AX SPGR (*n* = 34 Exams)		
	Image Quality of CS-SENSE ≥ Conventional			Image Quality of CS-SENSE ≥ Conventional		
	Both Raters	1 Rater	Neither Rater	Both Raters	1 Rater	Neither Rater
Optimal image quality	28 (80.0)	7 (20.0)	0 (0.0)	26 (76.5)	6 (17.6)	2 (5.9)
Optimal SNR	32 (91.4)	3 (8.6)	0 (0.0)	23 (67.6)	10 (29.4)	1 (2.9)
Trace or no artifacts	27 (77.1)	5 (14.3)	3 (8.6)	29 (85.3)	3 (8.8)	2 (5.9)
Sharp gray-white matter boundaries	28 (80.0)	7 (20.0)	0 (0.0)	26 (76.5)	8 (23.5)	0 (0.0)
Sharp lesion boundaries	29 (87.9)	4 (12.1)	0 (0.0)	27 (87.1)	3 (9.7)	1 (3.2)

Note:—SAG indicates sagittal; AX, axial.
[a] Values are No. (%).

Table 5: White matter SNR and lesion-white matter CNR

Variable	SAG FLAIR (*n* = 35 Exams)					AX SPGR (*n* = 34 Exams)				
	Acquisition[a]		Difference			Acquisition[a]		Difference		
	CS-SENSE	Conventional	Value[b]	(95% CI)	*P* Value	CS-SENSE	Conventional	Value[b]	(95% CI)	*P* Value
White matter SNR	68 (62–89)	59 (51–75)	8.6	(−0.2–17.8)	.062	43 (34–69)	43 (28–69)	0.3	(−10.9–13.7)	.75
Lesion-white matter CNR	68 (49–81)	60 (44–78)	8.6	(−2.4–17.4)	.083	15 (9–32)	12 (6–24)	3.3	(−3.2–8.8)	.31

Note:—SAG indicates sagittal; AX, axial.
[a] Values are median (interquartile range).
[b] Values are the difference in medians of the two acquisitions.

Quantitative Assessment

The white matter SNR or lesion CNR measurements were high for both conventional and CS-SENSE FLAIR acquisitions (≥44 in all cases), though there was a trend toward higher SNR and CNR values on average for the CS-SENSE acquisition (Table 5). White matter SNR measurements were also relatively high for conventional and CS-SENSE SPGR acquisitions (≥28 in all cases) with little numeric difference between them on average (difference in medians, 0.3; $P = .75$). The lesion CNR measurements from the SPGR acquisitions tended to be lower and ranged from 9 to 24 overall. The CNR was slightly higher on average in CS-SENSE than on the conventional acquisitions (difference in medians, 3.3; $P = .31$), but the difference was not statistically significant (Table 5).

DISCUSSION

Long MR imaging acquisition times represent a significant limitation to widespread use of MR imaging. This is especially true for MR imaging in the evaluation of both clinically unstable and pediatric patients: Increased MR imaging scan time may expose these individuals to an increased need for sedation or result in limited diagnostic quality due to motion. Long image-acquisition times also negatively impact radiology workflow, leading to scheduling bottlenecks. Finally, long acquisition times contribute to the high cost of MR imaging. Given the significance of image-acquisition time for patient safety, clinical efficiency, image quality, and cost, technical effort has been made to decrease image-acquisition and reconstruction times. CS techniques show promise in providing imaging acceleration without significant image-quality degradation. Despite the promise these acceleration techniques hold for improving patient throughput and decreasing imaging cost, rigorous evaluation of the performance of these acceleration techniques in a clinical imaging population has yet to be undertaken. To our knowledge, this is the first study to translate CS-SENSE, which combines and integrates CS and SENSE parallel imaging, to a clinical brain tumor patient population to evaluate image quality relative to corresponding conventional MR imaging sequences.

In the current study, we hypothesized that CS-SENSE accelerated sequences would have image quality equivalent to that of standard acquisitions while accelerating imaging. To evaluate this hypothesis, we compared the clinical performance of 2 CS-SENSE accelerated MR imaging sequences with their corresponding conventional sequences in a clinical cohort undergoing brain tumor MR imaging scans. On the basis of blinded multirater evaluations of multiple clinically pertinent imaging variables, these accelerated acquisitions largely performed as well as their conventional counterparts across several image-quality metrics, including overall image quality, SNR, image artifacts, gray-white matter boundary sharpness, and parenchymal lesion border sharpness. In particular, the lower bounds of 95% CIs of the differences in image quality between CS-SENSE and conventional images indicated that the CS-SENSE images were within 10% of the conventional images for all metrics for the SPGR sequence and for 3 of 5 metrics for the FLAIR sequence. The CS-SENSE acquisitions had no significant differences in white matter SNR and lesion CNR relative to their corresponding conventional acquisitions, and in fact, there was a trend toward higher values for the CS-SENSE FLAIR relative to the conventional FLAIR. CS-SENSE showed at least comparable SNR and CNR measures relative to their conventional counterparts. While this finding is somewhat counterintuitive considering the undersampling algorithm used by CS-SENSE, increased/similar SNR is thought be a result of the denoising algorithm incorporated into CS-SENSE. CS-SENSE FLAIR and SPGR sequences decreased imaging time by 25% and 35% relative to conventional sequences, respectively. These results confirm that CS-SENSE sequences produce diagnostic-quality MR images of the brain specifically for brain tumor protocols while reducing overall image-acquisition time compared with conventional acquisitions.

To date, multiple MR imaging techniques have been developed with the goal of accelerating image-acquisition and recon-

struction times. Examples of these acceleration techniques include parallel imaging and CS. Because parallel imaging and CS rely on different pieces of ancillary information for image production, it is possible to combine and integrate these acceleration techniques,[1-3,5,6,12] with the resultant combined technique referred to as CS-SENSE MR imaging. Most interesting, integration of these 2 individual acceleration techniques is synergistic and results in an imaging-acceleration factor that exceeds the accelerations achieved by either parallel imaging or CS alone while still keeping image noise low.[3,5] Liang et al[3] illustrated this principle using MR imaging phantoms, demonstrating superior performance of CS-SENSE relative to parallel imaging or CS alone at increasing acceleration factors. Otazo et al[5] independently came to similar conclusions when they demonstrated that a combined CS-SENSE technique resulted in a 2-fold increase in acceleration over CS alone.

Despite the increases in imaging acceleration observed with in vitro phantom imaging, little has been done to evaluate the performance of CS-SENSE acceleration techniques in a clinical patient population. Chandarana et al[4] used a novel CS-SENSE accelerated T1-weighted gradient-echo technique with a golden-angle radial *k*-space sampling scheme to evaluate 8 healthy subjects undergoing multiphase liver MR imaging. The authors concluded that their CS-SENSE accelerated sequence produced diagnostic-quality images. Kayvanrad et al[8] evaluated accelerated brain MR imaging using CS, parallel imaging, or low-resolution scans in 5 healthy volunteers and found that CS significantly outperformed parallel imaging and low-resolution imaging with a number of sequences. Sharma et al[9] performed retrospective CS of fully acquired datasets in 15 neuroimaging datasets to determine appropriate reconstructions and delineate expected artifacts in these reconstructions; they found 2-fold acceleration to be appropriate with ringing and blurring artifacts the most commonly encountered artifacts. These studies evaluated small cohorts of healthy volunteers and not patient populations. Toledano-Massiah et al[7] evaluated the clinical performance of CS 3D FLAIR in a multiple sclerosis population and demonstrated that CS produces diagnostic-quality images of the brain while reducing imaging time. In contrast to this study, which evaluated CS-accelerated FLAIR image quality, we evaluated the clinical performance of CS-SENSE acceleration of both T1- and T2-weighted sequences in a brain tumor patient population that has not previously been evaluated with these acceleration techniques.

CS-SENSE has its own inherent technical limitations. While undersampling of *k*-space is essential to decreasing image-acquisition time, if too few data points are acquired, then image quality is adversely affected. This issue manifests as decreased SNR and increased image blurring.[3] Technical challenges also arise during image reconstruction. To accurately reconstruct CS-SENSE images, solution of an optimization problem with multiple mathematical constraints is required. This process proves computationally complex and time-consuming. Fortunately, immediate sensitivity coil compression to a smaller number of virtual coils can aid in reducing the time required to complete the iterative reconstruction process.[13] The iteration used in the current examination reconstructs inline in <1 minute.

Our study had several limitations and challenges. The clinical cohort evaluated was relatively limited in size. Additionally, the current study evaluated imaging acceleration using CS-SENSE in patients undergoing an MR imaging brain tumor protocol. Individuals with other central nervous system pathologies were not included in our clinical cohort, limiting generalizability to other brain MR imaging applications. Because imaging focused on the brain, these results may not be generalizable to other pathologies or anatomic sites including spine, head and neck, and body MR imaging. Further investigation is necessary to establish the value of CS-SENSE in these anatomic regions. Interrater agreement (κ) ranged from 0 to 0.6 for the current study across all evaluations. The raters reviewed the accelerated sequences randomly and independent of the corresponding conventional acquisition using qualitative scales, both of which likely contributed to the limited agreement. Although raters had disagreements on the actual rating, they usually agreed $>80\%$ of the time that CS-SENSE image-quality metrics were at least as good as those for conventional images. Last, raters did not perform a diagnostic or clinical task such as detecting or characterizing lesions or changes in lesions, so further study is needed to confirm that these tasks are not adversely affected by CS-SENSE acceleration.

CONCLUSIONS

Accelerated CS-SENSE MR imaging provides equivalent image quality compared with corresponding conventional MR imaging in patients undergoing brain MR imaging of tumor by 3D T2-FLAIR and T1-SPGR sequences based on qualitative and quantitative assessment. CS-SENSE acceleration does not appear to adversely impact overall image quality relative to the corresponding conventional acquisitions. The results of this study are consistent with our hypothesis that there are no appreciable differences between the CS-SENSE acquisitions and their conventional counterparts. On the basis of these results, use of CS-SENSE accelerated sequences in clinical practice can be considered to reduce scan time and improve throughput without loss of image quality, specifically for 3D T2-FLAIR and T1-SPGR sequences in brain tumor MR imaging protocols.

Disclosures: Daniel S. Hippe—*UNRELATED*: *Grants/Grants Pending*: National Institutes of Health (R01NS092207), GE Healthcare, Philips Healthcare, Toshiba America Medical Systems, Siemens, *Comments*: for statistical work on other studies.* Elwin de Weerdt—*UNRELATED*: *Employment*: Royal Philips. *Money paid to the institution.

REFERENCES

1. Chun IY, Adcock B, Talavage TM. **Efficient compressed sensing SENSE pMRI reconstruction with joint sparsity promotion.** *IEEE Trans Med Imaging* 2016;35:354–68 CrossRef Medline
2. Tsao J, Kozerke S. **MRI temporal acceleration techniques.** *J Magn Reson Imaging* 2012;36:543–60 CrossRef Medline
3. Liang D, Liu B, Wang J, et al. **Accelerating SENSE using compressed sensing.** *Magn Reson Med* 2009;62:1574–84 CrossRef Medline
4. Chandarana H, Feng L, Block TK, et al. **Free-breathing contrast-enhanced multiphase MRI of the liver using a combination of compressed sensing, parallel imaging, and golden-angle radial sampling.** *Invest Radiol* 2013;48:10–16 CrossRef Medline
5. Otazo R, Kim D, Axel L, et al. **Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI.** *Magn Reson Med* 2010;64:767–76 CrossRef Medline
6. Tam LK, Galiana G, Stockmann JP, et al. **Pseudo-random center placement O-space imaging for improved incoherence compressed sensing parallel MRI.** *Magn Reson Med* 2015;73:2212–24 CrossRef Medline

7. Toledano-Massiah S, Sayadi A, de Boer R, et al. **Accuracy of the compressed sensing accelerated 3D-FLAIR sequence for the detection of MS plaques at 3T.** *AJNR Am J Neuroradiol* 2018 Jan 18. [Epub ahead of print] CrossRef Medline
8. Kayvanrad M, Lin A, Joshi R, et al. **Diagnostic quality assessment of compressed sensing accelerated magnetic resonance neuroimaging.** *J Magn Reson Imaging* 2016;44:433–44 CrossRef Medline
9. Sharma SD, Fong CL, Tzung BS, et al. **Clinical image quality assessment of accelerated magnetic resonance neuroimaging using compressed sensing.** *Invest Radiol* 2013;48:638–45 CrossRef Medline
10. Zibulevsky M, Elad M. **L1–L2 optimization in signal and image processing.** *IEEE Signal Processing Magazine* 2010;27:76–88 CrossRef
11. Kaufman L, Kramer DM, Crooks LE, et al. **Measuring signal-to-noise ratios in MR imaging.** *Radiology* 1989;173:265–67 CrossRef Medline
12. Liu F, Duan Y, Peterson BS, et al. **Compressed sensing MRI combined with SENSE in partial k-space.** *Phys Med Biol* 2012;57:N391–403 CrossRef Medline
13. Hollingsworth KG. **Reducing acquisition time in clinical MRI by data undersampling and compressed sensing reconstruction.** *Phys Med Biol* 2015;60:R297–322 CrossRef Medline

Determinants of Deep Gray Matter Atrophy in Multiple Sclerosis: A Multimodal MRI Study

G. Pontillo, S. Cocozza, R. Lanzillo, C. Russo, M.D. Stasi, C. Paolella, E.A. Vola, C. Criscuolo, P. Borrelli, G. Palma, E. Tedeschi, V.B. Morra, A. Elefante, and A. Brunetti

ABSTRACT

BACKGROUND AND PURPOSE: Deep gray matter involvement is a consistent feature in multiple sclerosis. The aim of this study was to evaluate the relationship between different deep gray matter alterations and the development of subcortical atrophy, as well as to investigate the possible different substrates of volume loss between phenotypes.

MATERIALS AND METHODS: Seventy-seven patients with MS (52 with relapsing-remitting and 25 with progressive MS) and 41 healthy controls were enrolled in this cross-sectional study. MR imaging investigation included volumetric, DTI, PWI and Quantitative Susceptibility Mapping analyses. Deep gray matter structures were automatically segmented to obtain volumes and mean values for each MR imaging metric in the thalamus, caudate, putamen, and globus pallidus. Between-group differences were probed by ANCOVA analyses, while the contribution of different MR imaging metrics to deep gray matter atrophy was investigated via hierarchic multiple linear regression models.

RESULTS: Patients with MS showed a multifaceted involvement of the thalamus and basal ganglia, with significant atrophy of all deep gray matter structures ($P < .001$). In the relapsing-remitting MS group, WM lesion burden proved to be the main contributor to volume loss for all deep gray matter structures ($P \leq .006$), with a minor role of local microstructural damage, which, in turn, was the main determinant of deep gray matter atrophy in patients with progressive MS ($P \leq .01$), coupled with thalamic susceptibility changes ($P = .05$).

CONCLUSIONS: Our study confirms the diffuse involvement of deep gray matter in MS, demonstrating a different behavior between MS phenotypes, with subcortical GM atrophy mainly determined by global WM lesion burden in patients with relapsing-remitting MS, while local microstructural damage and susceptibility changes mainly accounted for the development of deep gray matter volume loss in patients with progressive MS.

ABBREVIATIONS: DD = disease duration; DGM = deep gray matter; DMT = disease-modifying treatment; EDSS = Expanded Disability Status Scale; FA = fractional anisotropy; HC = healthy controls; LL = lesion load; MD = mean diffusivity; PMS = progressive MS; QSM = Quantitative Susceptibility Mapping; rCBV = relative CBV; RRMS = relapsing-remitting MS

Deep gray matter (DGM) involvement is generally regarded as a consistent feature in MS, generating particular interest due to its clinical relevance.[1]

Indeed, the occurrence of DGM atrophy in these patients, often described as an early phenomenon,[2,3] has been proved by several volumetric MR imaging studies, with a significant correlation with clinical disability, cognition, and disease progression.[3-5]

Along with volume loss, a wide range of pathologic changes affecting the DGM of patients with MS has been also demonstrated using different advanced MR imaging techniques. In particular, DTI studies showed the presence of microstructural damage in these structures,[6-9] while PWI and Quantitative Susceptibility Mapping (QSM) studies described decreased cerebral perfusion[10-13] and a complex pattern of susceptibility changes[14-17] affecting the DGM of patients with MS, respectively.

Although atrophy most certainly reflects neuronal loss, the main causes driving its development in MS are still debated. In-

Received July 18, 2018; accepted after revision October 29.

From the Departments of Advanced Biomedical Sciences (G.P., S.C., C.R., M.D.S., C.P., E.A.V., E.T., A.E., A.B.) and Neurosciences and Reproductive and Odontostomatological Sciences (R.L., C.C., V.B.M.), University of Naples "Federico II", Naples, Italy; IRCCS SDN (P.B.), Naples, Italy; and Institute of Biostructure and Bioimaging (G.P.), National Research Council, Naples, Italy.

Giuseppe Pontillo and Sirio Cocozza contributed equally to this work.

This study received funding from the Italian Ministry of Education, University, and Research within the Progetto di Ricerca di Interesse Nazionale framework (2010XE5L2R).

Please address correspondence to Sirio Cocozza, MD, Department of Advanced Biomedical Sciences, University "Federico II", Via Pansini, 5, 80131 Naples, Italy; e-mail: sirio.cocozza@unina.it; @NeuroN_Lab

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

Indicates article with supplemental on-line appendix and tables.

http://dx.doi.org/10.3174/ajnr.A5915

Table 1: All subjects' demographics and clinical variables[a]

	MS	HC	*P* Value (MS vs HC)	PMS	RRMS	*P* Value (PMS vs RRMS)
Age	41.2 ± 11.0	43.6 ± 13.9	.34	43.7 ± 11.8	40.0 ± 10.6	.17
Sex	33 F/41 M	20 F/21 M	.54	10 F/15 M	23 F/29 M	.73
DD	12.1 ± 7.6	NA	NA	14.3 ± 7.6	11.1 ± 7.5	.09
EDSS	3.5 (2.0–7.5)	NA	NA	5.5 (3.0–7.5)	3.5 (2.0–6.0)	<.001[b]
DMT	71/77 (92.2%)	NA	NA	22/25 (88.0%)	49/52 (94.2%)	.34

Note:—NA indicates not applicable.
[a] Age and DD (in years) are expressed as mean ± SD; EDSS, as median (range).
[b] Significant difference.

deed, DGM volume loss could be either due to the occurrence of primary local pathology or secondary to WM inflammatory damage, leading to Wallerian degeneration and deafferentation.[18,19] From this background and despite the relatively wide knowledge about the multifaceted involvement of DGM in MS, a study simultaneously investigating the contribution to the development of subcortical GM atrophy of these different physiopathologic changes is, to date, lacking.

We therefore collected different MR imaging variables related to distinct aspects of DGM damage in patients with both relapsing-remitting MS (RRMS) and progressive MS (PMS), to investigate the relationship between these alterations and the development of subcortical GM atrophy, as well as the possible different physiopathologic substrates of DGM volume loss between MS phenotypes.

MATERIALS AND METHODS

Subjects

In this single-center observational study, 77 patients with MS (52 RRMS and 25 PMS according to the 2013 revised definition of MS phenotypes[20]) and 41 healthy controls (HC) were enrolled from October 2013 to July 2015 in a neuroimaging study of neuroinflammatory disorders.[21,22] All patients with MS fulfilled the 2010 revised McDonald criteria,[23] while none of the HC presented with any condition that could affect the CNS. For all patients, the Expanded Disability Status Scale (EDSS) score was determined within 1 week from MR imaging by an experienced neurologist as an index of clinical disability, along with the record of disease duration (DD) and disease-modifying treatment (DMT) type (On-line Table 1).

Demographic and clinical characteristics of all subjects included in the analysis are provided in Table 1.

Standard Protocol Approvals, Registrations, and Patient Consents

The protocol was approved by the "Carlo Romano" ethics committee for biomedical activities of "Federico II," University of Naples, and written informed consent was obtained from all participants before the beginning of the study in accordance with the Declaration of Helsinki.

MR Imaging Data Acquisition and Processing

A complete description of all the acquired sequences, along with all processing procedures, is available in the On-line Appendix.

Briefly, the acquisition protocol included a 3D T1-weighted sequence used for volumetric analyses, a 3D-FLAIR sequence for the quantification of demyelinating lesion load (LL) volume, an echo-planar imaging sequence for the dynamic susceptibility contrast–PWI analysis, an unenhanced 3D double-echo FLASH sequence for the calculation of QSM maps, and, in a subgroup of 59 patients (38 RRMS, 21 PMS) and 38 HC, an echo-planar imaging sequence for DTI analysis.

DGM segmentation was achieved using the FIRST routine (FMRIB Integrated Registration and Segmentation Tool; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FIRST) implemented in FSL, Version 5.0.10; the latter was also used to extract fractional anisotropy (FA) and mean diffusivity (MD) maps from the DTI datasets. PWI data were analyzed using Olea Sphere software, Version 2.3 (Olea Medical, La Ciotat, France) to generate relative cerebral blood volume (rCBV) and relative CBF maps. A complete description of all processing steps required for the calculation of QSM images is available in Palma et al,[24] Borrelli et al,[25] and Palma et al.[26]

For each subject, DTI, PWI, and QSM maps were coregistered to the 3D T1-weighted sequence via affine registration, and segmentation masks were used to obtain DGM volumes and mean values for each MR imaging metric (Fig 1).

Statistical Analysis

Distribution of all data was preliminarily checked with graphs and tests (ie, the Levene test for homoscedasticity and the Kolmogorov-Smirnov test for normality), and those variables showing a significantly skewed distribution (namely, LL) were normalized by log-transformation.

Group differences were probed by the Student *t* (age, DD), Pearson χ^2 (sex, DMT), and Mann-Whitney (EDSS) tests, while differences in MR imaging metrics were tested by ANCOVA analyses, including age, sex, and DD (when comparing MS subgroups) as covariates.

To determine the main contributors to subcortical GM atrophy in patients with MS, we conducted analyses as follows. For each DGM structure, the relationship between advanced MR imaging metrics and the respective normalized volume was preliminarily investigated using partial correlation analyses (age-, sex-, and DD-corrected) in the whole MS population and within the different subgroups. Variables showing a significant correlation with DGM volume loss were entered in the third and final step of each hierarchic multiple linear regression analysis, in which volume was set as the dependent variable and clinical and demographic variables (age, sex, DD, and type of DMT) were entered in the first block and LL in the second step.

Finally, to explore the potential additional value of multiple MR imaging parameters for the prediction of clinical disability

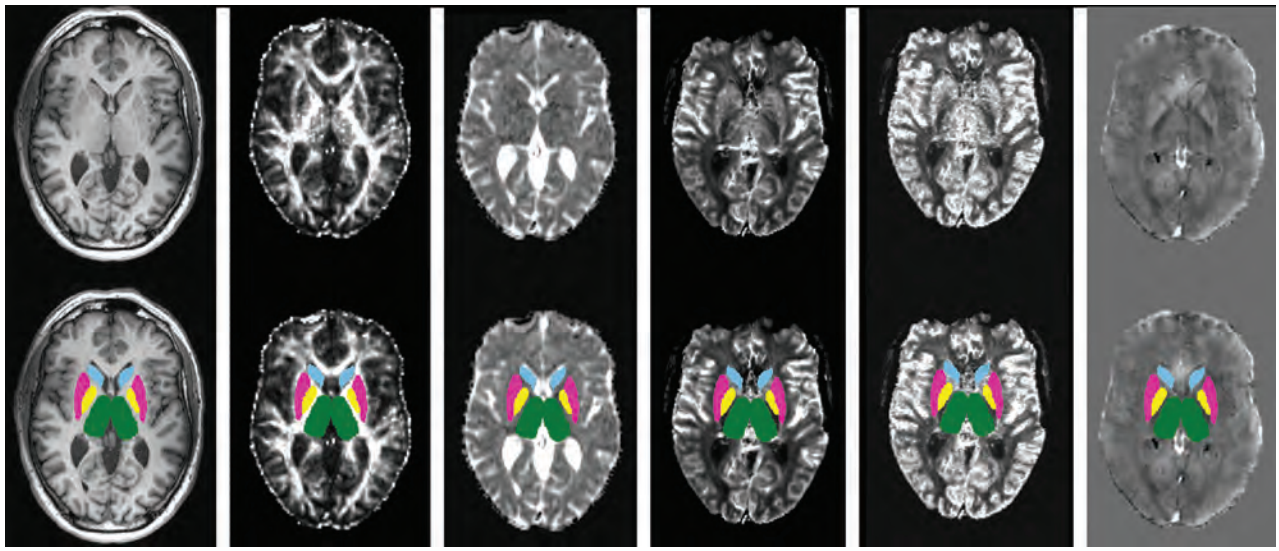


FIG 1. Results of the coregistration between different MR imaging modalities (*upper row*), with superimposed deep gray matter masks (*lower row*). From left to right, 3D T1-weighted volume, fractional anisotropy, mean diffusivity, relative cerebral blood volume, relative cerebral blood flow, and Quantitative Susceptibility Mapping images.

in all groups, we also preliminarily investigated the relationship between the EDSS score and DGM metrics that proved to be significantly different between patients with MS and HC via partial correlation analyses (age-, sex-, and DD-corrected). MR imaging metrics showing a significant correlation with EDSS were then entered in the fourth and final block of a hierarchic multiple linear regression analysis for the prediction of clinical disability, including clinicodemographic variables in the first step, LL in the second block, and DGM volumes in the third step, with the same analysis also probed within different MS subgroups.

All analyses were performed with the Statistical Package for the Social Sciences, Version 24 (IBM, Armonk, New York), with a significance level set at $P < .05$, Bonferroni-corrected for multiple comparisons. For the regression analyses, multiple-comparison correction was adopted for the models only, while independent predictors were considered significant at $P < .05$.

Data Availability

Anonymized data will be shared by request from any qualified investigator.

RESULTS

Demographic and Clinical Findings

The MS and HC groups did not significantly differ in age ($P = .34$) and sex ($P = .54$). Similarly, when we compared PMS and RRMS, the 2 subgroups did not show any significant difference in terms of age ($P = .17$), sex ($P = .73$), DD ($P = .09$), or DMT ($P = .34$), with patients with progressive MS showing a more severe clinical impairment compared with those with RRMS ($P < .001$).

Between-Group MR Imaging Analysis

A complete list of the results of the between-group analysis regarding MR imaging metrics is available in Table 2.

Compared with HC, patients with MS showed a significant reduction of all brain volumes (all with $P < .001$), with the PMS subgroup presenting with a slightly higher lesion load compared with the RRMS group ($P = .03$, not significant after Bonferroni correction), while no differences emerged between phenotypes in terms of brain tissue volumes.

A significant atrophy of both the thalamus and basal ganglia was found in MS compared with HC (all with $P < .001$), with increased MD values in patients at the level of caudate nucleus and thalamus ($P < .001$) and a trend toward increased FA in the putamen ($P = .02$, not significant after Bonferroni correction). Finally, patients showed a reduction of susceptibility values in the thalamus compared with HC ($P = .001$), while no differences emerged for PWI measures.

When possible differences between MS phenotypes were probed, PMS compared with RRMS showed a significant volumetric reduction of the thalamus ($P = .005$), caudate ($P = .007$), globus pallidus ($P = .001$), and, to a lesser extent, the putamen ($P = .02$, not significant after Bonferroni correction), with borderline higher MD ($P = .05$, not significant after Bonferroni correction), significantly reduced rCBV values ($P = .002$) at the level of the caudate nucleus, and a trend toward higher FA in the putamen ($P = .05$, not significant after Bonferroni correction).

Partial Correlation Analyses

A complete list of all the results obtained from the preliminary correlation analyses is available in the On-line Appendix and corresponding On-line Tables 2 and 3.

Relationship between Advanced MR Imaging Metrics and DGM Volumes

A list of the results of the regression analyses investigating the relations between advanced MR imaging metrics and DGM volumes in patients with MS as well as within different subgroups is reported in Tables 3 and 4, respectively.

The regression analysis conducted on the entire MS group showed that LL (all with $P \leq .02$) and microstructural GM

Table 2: MRI metrics for all subjects included in the analysis[a]

	MS	HC	*P* Value (MS vs HC)	PMS	RRMS	*P* Value (PMS vs RRMS)
LL	13.8 ± 16.2	NA	NA	18.3 ± 16.7	11.6 ± 15.7	.03[c]
NBV	1467.8 ± 86.9	1547.2 ± 67.7	<.001[b]	1440.0 ± 101.7	1481.2 ± 76.3	.17
NGMV	745.8 ± 60.4	793.5 ± 54.8	<.001[b]	724.7 ± 66.3	755.9 ± 55.3	.09
NWMV	722.0 ± 36.6	753.6 ± 27.6	<.001[b]	715.3 ± 46.4	725.3 ± 30.7	.31
DGM Volumes						
Thalamus	18.7 ± 2.6	21.9 ± 2.1	<.001[b]	17.4 ± 3.1	19.4 ± 2.1	.005[b]
Caudate	8.4 ± 1.4	9.5 ± 1.2	<.001[b]	7.7 ± 1.6	8.7 ± 1.1	.007[b]
Putamen	12.3 ± 1.8	13.8 ± 1.8	<.001[b]	11.5 ± 1.3	12.9 ± 1.5	.02[c]
Globus pallidus	4.5 ± 0.7	5.0 ± 0.6	<.001[b]	4.1 ± 0.8	4.6 ± 0.5	.001[b]
DGM Diffusion						
MD thalamus	0.95 ± 0.09	0.88 ± 0.06	<.001[b]	0.98 ± 0.11	0.96 ± 0.11	.25
MD caudate	1.00 ± 0.10	0.92 ± 0.08	<.001[b]	1.04 ± 0.10	0.98 ± 0.10	.05[c]
FA putamen	0.272 ± 0.024	0.263 ± 0.019	.02[c]	0.282 ± 0.029	0.267 ± 0.020	.05[c]
DGM Perfusion						
rCBV caudate	2.8 ± 1.0	3.1 ± 1.0	.12	2.3 ± 1.0	3.0 ± 0.9	.002[b]
DGM Susceptibility						
χ thalamus	−26.8 ± 31.4	−7.7 ± 22.4	.001[b]	−38.7 ± 31.4	−21.1 ± 30.1	.14

Note:—NBV indicates normalized brain volume; NGMV, normalized gray matter volume; NWMV, normalized white matter volume; NA, not applicable.
[a] Values are expressed as means ± SD (volume in milliliters, MD in $\times 10^{-3}$ mm^2/s, rCBV in milliliters/100 mL, FA is a scalar value between 0 and 1, χ is in parts per billion).
[b] Significant difference.
[c] Not significant after Bonferroni correction.

Table 3: Results of the analyses exploring the relationship between MRI measures and DGM volume in the entire MS group

	Model			Predictor		
DGM Volumes	R^2 (ΔR^2)	*F* (ΔF)	*P* Value	Standardized β	*T*	*P* Value
Thalamus	0.679 (0.320)	13.266 (21.902)	<.001			
DMT				−0.205	−2.189	.03
LL				−0.347	−3.383	.002
MD thalamus				−0.515	−4.738	<.001
χ thalamus				0.295	2.818	.007
Caudate	0.609 (0.297)	11.337 (19.351)	<.001			
DMT				−0.265	−2.729	.009
LL				−0.279	−2.696	.009
MD caudate				−0.422	−4.122	.001
FA caudate				−0.494	−5.255	<.001
Putamen	0.445 (0.187)	6.944 (17.546)	<.001			
DMT				−0.272	−2.429	.02
LL				−0.278	−2.367	.02
FA putamen				−0.516	−4.189	.001
Globus pallidus	0.300 (0.183)	6.078 (18.557)	<.001			
DMT				−0.358	−3.332	.001
LL				−0.481	−4.308	<.001

changes (all with $P \leq .001$) were independent predictors of volume loss for the investigated DGM structures, with an additional significant contribution of susceptibility changes to the development of thalamic volume loss ($P = .007$).

When the regression analyses were probed in the different MS phenotypes, LL proved to be a constant independent predictor of volume loss for all the DGM structures in the RRMS group (all with $P \leq .006$), with an additional and relative contribution of local microstructural GM changes in the development of thalamic ($P = .02$) and caudate ($P = .04$) atrophy. On the other hand, microstructural damage was the main determinant of DGM atrophy in patients with PMS (all with $P \leq .01$) without a significant value of global MR imaging measures but coupled with susceptibility changes in the development of volume loss at the level of the thalamus ($P = .05$).

Scatterplots of the relationships between DGM volumes and significant MR imaging predictors in patients with MS and within the different subgroups are shown in Figs 2 and 3, respectively.

Relationship between DGM Metrics and Clinical Disability

The regression analysis showed that the only independent predictor of clinical disability in the MS group was thalamic volume (standardized $\beta = -0.306$, $P = .02$), which explained, in addition to clinicodemographic variables and LL, 32.2% of the variance of the EDSS ($\Delta R^2 = 5.6\%$, $P < .001$) without incremental explanation of the variance provided by other advanced MR imaging measures (On-line Table 4). No significant differences between RRMS and PMS emerged in the same analysis.

DISCUSSION

In this study, we used a multimodal MR imaging approach to investigate different features of DGM involvement in MS, confirming the presence of a diffuse and multifaceted involvement of subcortical GM structures in this condition. We demonstrated that WM lesion burden represents the main determinant of DGM atrophy in RRMS, with a concomitant though relative role of local microstructural damage, which, in turn, proved to be the main

Table 4: Results of the analyses exploring the relationship between MRI measures and DGM volume in the MS subgroups

	Model			Predictor		
DGM Volumes	R^2 (ΔR^2)	F (ΔF)	P Value	Standardized β	T	P Value
RRMS						
Thalamus	0.750 (0.081)	6.637 (5.738)	<.001			
DMT				−0.289	−2.313	.03
LL				−0.454	−3.214	.003
MD thalamus				−0.371	−2.395	.02
Caudate	0.563 (0.061)	6.662 (4.346)	<.001			
DMT				−0.352	−2.702	.01
LL				−0.411	−2.938	.006
MD caudate				−0.287	−2.085	.04
Putamen	0.662 (0.196)	7.187 (16.073)	<.001			
DMT				−0.497	−4.166	<.001
LL				−0.502	−4.009	<.001
Globus pallidus	0.611 (0.158)	5.484 (11.608)	<.001			
DMT				−0.487	−3.868	<.001
LL				−0.450	−3.407	.001
PMS						
Thalamus	0.832 (0.585)	7.761 (19.119)	<.001			
MD thalamus				−0.566	−3.091	.01
χ thalamus				0.493	2.233	.05
Caudate	0.799 (0.576)	7.362 (18.587)	.001			
FA caudate				−0.783	−5.789	<.001
Putamen	0.580 (0.506)	3.228 (16.890)	.03[a]			
FA putamen				−0.836	−4.110	.001

[a] Not significant after Bonferroni correction.

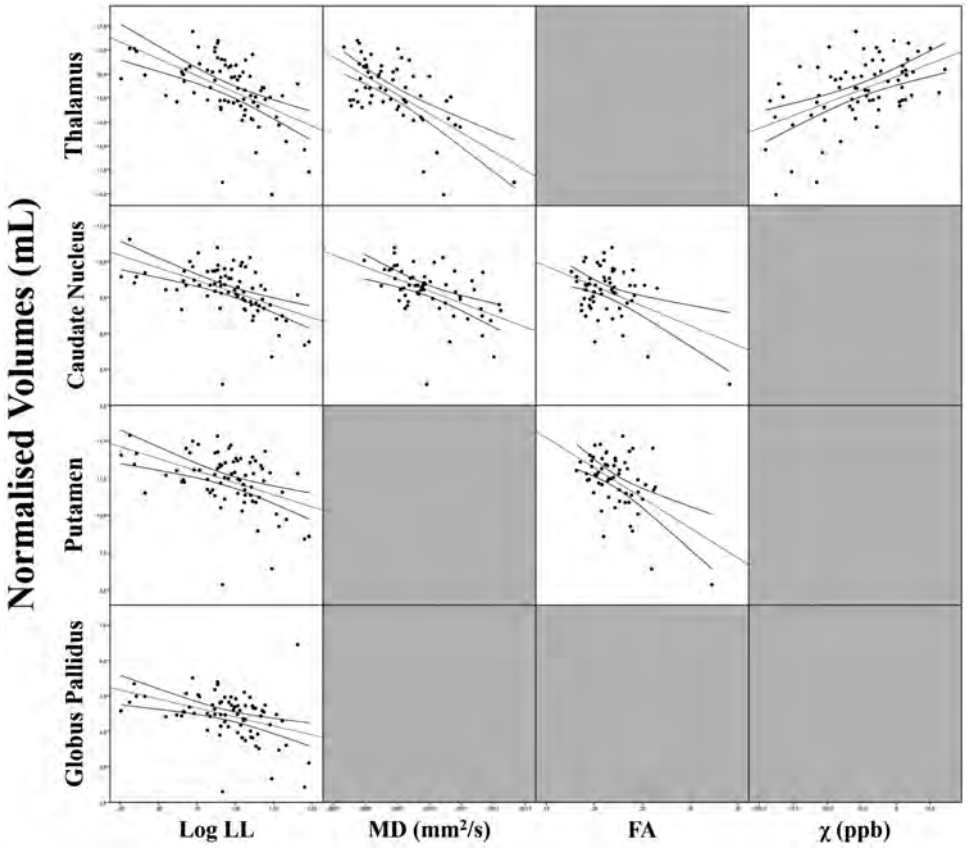


FIG 2. Scatterplot matrix showing the correlations between DGM volumes and their significant MR imaging predictors in the regression analyses in patients with MS.

contributor to the development of subcortical GM volume loss in PMS, along with thalamic susceptibility changes.

In accordance with the great corpus of scientific literature that describes subcortical GM atrophy as an early and consistent feature of MS, strictly linked with disease course and clinical progression,[2-5] we found significant DGM atrophy in our group of pa-

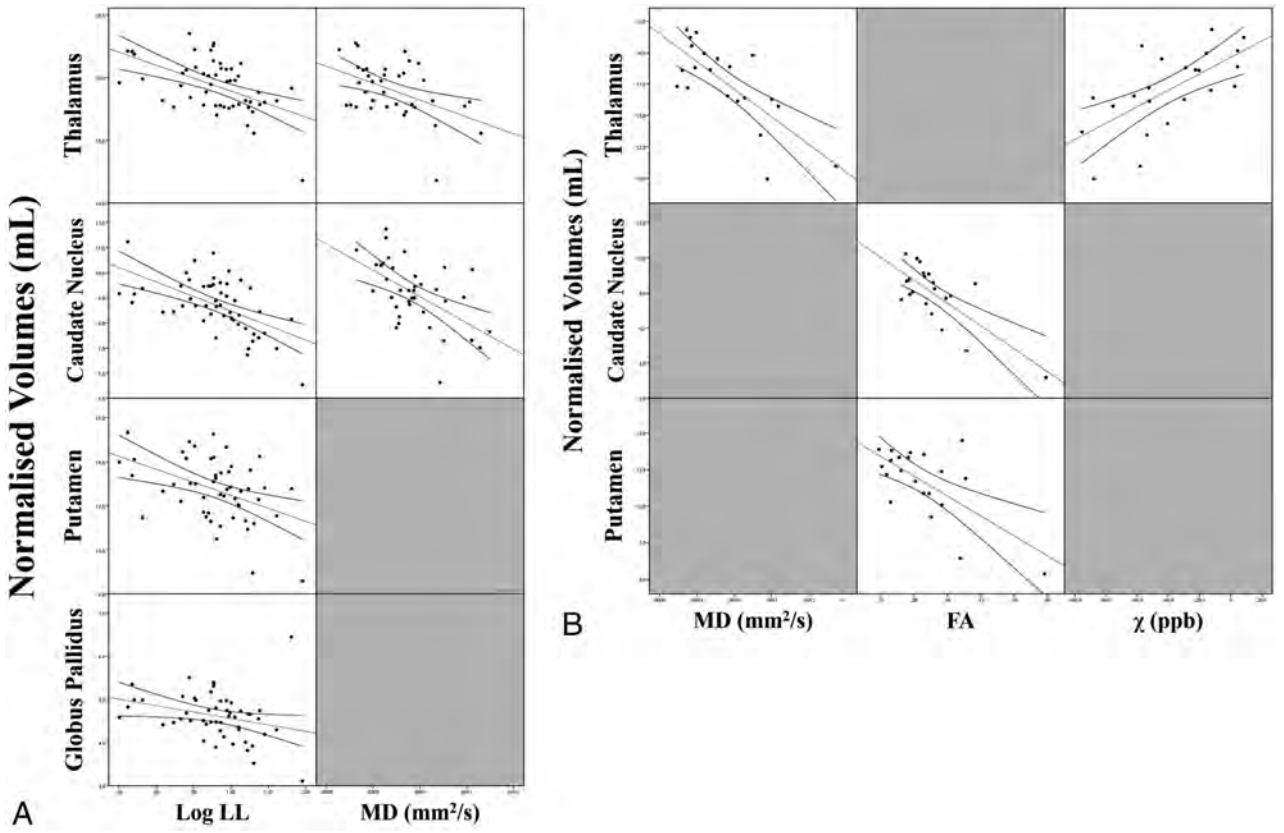


FIG 3. Scatterplot matrices of the correlations between DGM volumes and their significant MR imaging predictors in the regression analyses in patients with RRMS (*A*) and PMS (*B*).

tients. This volume loss was more pronounced in PMS, a finding also consistent with previous evidence,[2] in line with the hypothesis of a prominent role of neurodegenerative phenomena in the pathophysiology of this phenotype.[18,27]

Furthermore, our results confirm the presence of microstructural damage in the DGM of patients with MS, mostly affecting the caudate nuclei and thalami. These findings, also in substantial accordance with previous evidence,[6-9] can be considered as a reflection of microstructural damage due to demyelination and axonal injury, which cause a net loss of structural barriers facilitating water diffusion, while the increased FA might be explained by extra-axonal phenomena such as the loss of dendritic connections and/or the swelling of neuronal cell bodies.[7,8]

On the other hand, we failed to find significant alterations of PWI parameters in patients with MS compared with HC, a result in conflict with some previous studies in which decreased perfusion of these structures was described.[10-12,28,29] A possible explanation for this discrepancy could reside in the different methodologic approaches between our work (in line with those available in a study analyzing PWI data with a method similar to ours[30]) and these previous studies conducted using hand-drawn ROIs[11,28,29] (which had operator dependencies) or voxel-based approaches[10,12] (which provide different, though complementary, information). Most interesting, when we compared MS phenotypes, a reduction in rCBV values of the caudate nuclei in patients with PMS compared with RRMS was proved. Different mechanisms could explain this finding, mainly related to a decreased neuronal metabolic demand secondary to atrophy,[13,30] though primary neuronal metabolic dysfunctions and alterations of cerebral vasculature have also been proposed.[13,30]

Finally, at the QSM analysis, patients with MS showed a significant reduction of magnetic susceptibility values in the thalamus compared with HC, in line with recent quantitative MR imaging studies.[14-17] The physiopathologic basis of this altered susceptibility could reside in the variable association of reduced paramagnetic components (ie, iron) and increased diamagnetic components (ie, myelin and/or calcium). Thus, several hypotheses have been proposed, including increased myelin density due to GM loss, calcium deposition, and, in particular, iron depletion. The latter could be an indirect consequence of tissue loss and/or could result from an active process of iron removal from damaged oligodendrocytes, related to chronic microglia activation, ultimately leading to neurodegeneration.[14-17]

When investigating the contribution of different MR imaging metrics to the development DGM atrophy in MS subgroups, we found that a global MR imaging measure of WM damage (namely, the LL) was a constant significant predictor of volume loss for all DGM structures for the patients with RRMS, with an additional, though relative, contribution of local microstructural GM changes in the development of thalamic and caudate atrophy only. This result confirms a suggested possible role of WM lesions in driving atrophy of the highly connected subcortical GM structures, most likely through axonal transection leading to disconnection, with a subsequent degeneration along axonal projections.[2,31]

On the other hand, microstructural damage proved to be the

main determinant of DGM atrophy in patients with PMS, without a significant role of LL, corroborating the role of local microstructural damage as a possible primary determinant of neuronal loss and subsequent atrophy in subcortical GM.[6-7]

Furthermore, when we evaluated the determinants of thalamic atrophy in PMS, a direct effect of the reduced susceptibility values found in this structure on its volume was found, in apparent contrast with the common notion of iron increase as a possible driver of neurodegeneration in MS.[32] A possible explanation for this last result could be found in the peculiar morphofunctional architecture of the thalamus. Indeed, this structure, due to its rich connectivity profile, shows a high oligodendrocyte density, being more prone to secondary effects from remote injury in other areas of the brain.[17] Thus, a chronic microglial activation could lead to thalamic oligodendroglial damage, with subsequent iron release, generating, in turn, a vicious cycle reducing axonal protection and neuronal repair, eventually leading to neurodegeneration.[17]

All these results, taken together, support the hypothesis that different MS phenotypes could be characterized by distinct physiopathologic mechanisms, with a more prominent role of primary GM pathology in patients with PMS, which occurs, at least in part, independent from global WM lesion burden.[1,19,27]

Finally, we explored the clinical impact of these DGM alterations, proving that thalamic volume was the only significant predictor of EDSS score, without any additional value provided by the other tested MR imaging measures. This result confirms and expands the current knowledge about the clinical relevance of thalamic atrophy in MS, which could represent a common pathway through which both WM lesions and local DGM pathology contribute to clinical disability.[3-5,33] Indeed, the thalamus is involved in all the major functional circuits in the brain, providing a point of convergence across multiple cortical, limbic, brain stem, and cerebellar systems; therefore, it is easy to understand how its volume loss could represent one of the most clinically relevant biomarkers of disease in MS.[33]

Some limitations of the present study should be acknowledged. In particular, additional physiopathologic information on DGM structures could have been obtained using other advanced MR imaging techniques (eg, magnetization transfer ratio, MR spectroscopy, or functional MR imaging) or with voxelwise approaches, while a longitudinal evaluation could have helped unravel the causal relationships among different aspects of DGM pathology as well as between subcortical GM alterations and clinical disability. Thus, future studies are warranted simultaneously investigating the role of additional different aspects of DGM pathology and their evolution with time, coupled with more extensive neurologic and neuropsychological examinations, to further expand our knowledge of the physiopathology and the clinical relevance of DGM damage in MS.

CONCLUSIONS

Our study provides additional information about DGM involvement in patients with MS, showing the presence of alterations of different MR imaging metrics as a possible reflection of neurodegenerative and neuroinflammatory processes in these structures. Furthermore, our results demonstrate the presence of a different behavior of DGM atrophy in MS phenotypes, with subcortical GM volume loss mainly determined by global WM lesion burden in patients with RRMS, while local microstructural damage, along with susceptibility changes, account for the development of the significant DGM atrophy occurring in patients with PMS.

Disclosures: Sirio Cocozza—*UNRELATED*: *Payment for Lectures Including Service on Speakers Bureaus*: Genzyme. Roberta Lanzillo—*UNRELATED*: *Consultancy*: Biogen, Merck, Teva Pharmaceutical Industries, Novartis; *Payment for Lectures Including Service on Speakers Bureaus*: Teva Pharmaceutical Industries, Biogen, Merck, Novartis, Genzyme. Enrico Tedeschi—*UNRELATED*: *Payment for Lectures Including Service on Speakers Bureaus*: Scientific Press, Ars Educandi, Genzyme; *Payment for Development of Educational Presentations*: Scientific Press, Ars Educandi, Genzyme. Vincenzo Brescia Morra—*UNRELATED*: *Board Membership*: Novartis, Biogen, Teva Pharmaceutical Industries, Mylan, Bayer AG, Merck, Genzyme, Roche; *Consultancy*: Novartis, Biogen, Teva Pharmaceutical Industries, Mylan, Bayer AG, Merck, Genzyme, Roche; *Expert Testimony*: Novartis, Biogen, Teva Pharmaceutical Industries, Mylan, Bayer AG, Merck, Genzyme, Roche; *Grants/Grants Pending*: Novartis, Biogen, Teva Pharmaceutical Industries, Mylan, Bayer AG, Merck, Genzyme, Roche*; *Payment for Lectures Including Service on Speakers Bureaus*: Novartis, Biogen, Teva Pharmaceutical Industries, Mylan, Bayer AG, Merck, Genzyme, Roche; *Payment for Development of Educational Presentations*: Novartis, Biogen, Teva Pharmaceutical Industries, Mylan, Bayer AG, Merck, Genzyme, Roche; *Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed*: Novartis, Biogen, Teva Pharmaceutical Industries, Mylan, Bayer AG, Merck, Genzyme, Roche. Arturo Brunetti—*UNRELATED: Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed*: Bracco, GE Healthcare, *Comments*: Tickets and hotel accommodation were provided by Bracco for participation in a 1-day a meeting concerning MR contrast media in 2017. Hotel accommodations were provided by GE Healthcare for participation in a 1-day user meeting in 2017. No money provided directly. *Money paid to institution.

REFERENCES

1. Pirko I, Lucchinetti CF, Sriram S, et al. **Gray matter involvement in multiple sclerosis.** *Neurology* 2007;68:634–42 CrossRef Medline
2. Eshaghi A, Marinescu RV, Young AL, et al. **Progression of regional grey matter atrophy in multiple sclerosis.** *Brain* 2018;141:1665–77 CrossRef Medline
3. Zivadinov R, Havrdová E, Bergsland N, et al. **Thalamic atrophy is associated with development of clinically definite multiple sclerosis.** *Radiology* 2013;268:831–41 CrossRef Medline
4. Eshaghi A, Prados F, Brownlee WJ, et al; MAGNIMS study group. **Deep gray matter volume loss drives disability worsening in multiple sclerosis.** *Ann Neurol* 2018;83:210–22 CrossRef Medline
5. Schoonheim MM, Hulst HE, Brandt RB, et al. **Thalamus structure and function determine severity of cognitive impairment in multiple sclerosis.** *Neurology* 2015;84:776–83 CrossRef Medline
6. Hasan KM, Halphen C, Kamali A, et al. **Caudate nuclei volume, diffusion tensor metrics, and T(2) relaxation in healthy adults and relapsing-remitting multiple sclerosis patients: implications for understanding gray matter degeneration.** *J Magn Reson Imaging* 2009;29:70–77 CrossRef Medline
7. Hannoun S, Durand-Dubief F, Confavreux C, et al. **Diffusion tensor-MRI evidence for extra-axonal neuronal degeneration in caudate and thalamic nuclei of patients with multiple sclerosis.** *AJNR Am J Neuroradiol* 2012;33:1363–68 CrossRef Medline
8. Cavallari M, Ceccarelli A, Wang GY, et al. **Microstructural changes in the striatum and their impact on motor and neuropsychological performance in patients with multiple sclerosis.** *PLoS One* 2014;9:e101199 CrossRef Medline
9. Tovar-Moll F, Evangelou IE, Chiu AW, et al. **Thalamic involvement and its impact on clinical disability in patients with multiple sclerosis: a diffusion tensor imaging study at 3T.** *AJNR Am J Neuroradiol* 2009;30:1380–86 CrossRef Medline
10. Debernard L, Melzer TR, Van Stockum S, et al. **Reduced grey matter perfusion without volume loss in early relapsing-remitting multiple sclerosis.** *J Neurol Neurosurg Psychiatry* 2014;85:544–51 CrossRef Medline
11. Inglese M, Park SJ, Johnson G, et al. **Deep gray matter perfusion in multiple sclerosis: dynamic susceptibility contrast perfusion**

magnetic resonance imaging at 3 T. *Arch Neurol* 2007;64:196–202 CrossRef Medline
12. Doche E, Lecocq A, Maarouf A, et al. **Hypoperfusion of the thalamus is associated with disability in relapsing remitting multiple sclerosis.** *J Neuroradiol* 2017;44:158–64 CrossRef Medline
13. Lapointe E, Li DK, Traboulsee AL, et al. **What have we learned from perfusion MRI in multiple sclerosis?** *AJNR Am J Neuroradiol* 2018; 39:994–1000 CrossRef Medline
14. Zivadinov R, Tavazzi E, Bergsland N, et al. **Brain iron at quantitative MRI is associated with disability in multiple sclerosis.** *Radiology* 2018;289:487–96 CrossRef Medline
15. Hagemeier J, Zivadinov R, Dwyer MG, et al. **Changes of deep gray matter magnetic susceptibility over 2 years in multiple sclerosis and healthy control brain.** *Neuroimage Clin* 2018;18:1007–16 CrossRef Medline
16. Burgetova A, Dusek P, Vaneckova M, et al. **Thalamic iron differentiates primary-progressive and relapsing-remitting multiple sclerosis.** *AJNR Am J Neuroradiol* 2017;38:1079–86 CrossRef Medline
17. Schweser F, Raffaini Duarte Martins AL, Hagemeier J, et al. **Mapping of thalamic magnetic susceptibility in multiple sclerosis indicates decreasing iron with disease duration: a proposed mechanistic relationship between inflammation and oligodendrocyte vitality.** *Neuroimage* 2018;167:438–52 CrossRef Medline
18. Haider L, Simeonidou C, Steinberger G, et al. **Multiple sclerosis deep grey matter: the relation between demyelination, neurodegeneration, inflammation and iron.** *J Neurol Neurosurg Psychiatry* 2014;85: 1386–95 CrossRef Medline
19. Vercellino M, Masera S, Lorenzatti M, et al. **Demyelination, inflammation, and neurodegeneration in multiple sclerosis deep gray matter.** *J Neuropathol Exp Neurol* 2009;68:489–502 CrossRef Medline
20. Lublin FD, Reingold SC, Cohen JA, et al. **Defining the clinical course of multiple sclerosis: the 2013 revisions.** *Neurology* 2014;83:278–86 CrossRef Medline
21. Cocozza S, Canna A, Lanzillo R, et al. **Lack of correlation between extracranial venous abnormalities and multiple sclerosis: a quantitative MRI study.** *Br J Radiol* 2016 Jun 27:20160321. [Epub ahead of print] CrossRef Medline
22. Tedeschi E, Palma G, Canna A, et al. **In vivo dentate nucleus MRI relaxometry correlates with previous administration of gadolinium-based contrast agents.** *Eur Radiol* 2016;26:4577–84 CrossRef Medline
23. Polman CH, Reingold SC, Banwell B, et al. **Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria.** *Ann Neurol* 2011;69:292–302 CrossRef Medline
24. Palma G, Tedeschi E, Borrelli P, et al. **A novel multiparametric approach to 3D quantitative MRI of the brain.** *PLoS One* 2015;10: e0134963 CrossRef Medline
25. Borrelli P, Palma G, Tedeschi E, et al. **Improving signal-to-noise ratio in susceptibility weighted imaging: a novel multicomponent non-local approach.** *PLoS One* 2015;10:e0126835 CrossRef Medline
26. Palma G, Comerci M, Alfano B, et al. **3D non-local means denoising via multi-GPU.** In *Proceedings of the 2013 Federated Conference on Computer Science and Information Systems,* Kraków, Poland; September 8–11, 2013:495–98
27. Mahad DH, Trapp BD, Lassmann H. **Pathological mechanisms in progressive multiple sclerosis.** *Lancet Neurol* 2015;14:183–93 CrossRef Medline
28. Inglese M, Adhya S, Johnson G, et al. **Perfusion magnetic resonance imaging correlates of neuropsychological impairment in multiple sclerosis.** *J Cereb Blood Flow Metab* 2008;28:164–71 CrossRef Medline
29. Papadaki EZ, Mastorodemos VC, Amanakis EZ, et al. **White matter and deep gray matter hemodynamic changes in multiple sclerosis patients with clinically isolated syndrome.** *Magn Reson Med* 2012; 68:1932–42 CrossRef Medline
30. Debernard L, Melzer TR, Alla S, et al. **Deep grey matter MRI abnormalities and cognitive function in relapsing-remitting multiple sclerosis.** *Psychiatry Res* 2015;234:352–61 CrossRef Medline
31. Mühlau M, Buck D, Förschler A, et al. **White-matter lesions drive deep gray-matter atrophy in early multiple sclerosis: support from structural MRI.** *Mult Scler* 2013;19:1485–92 CrossRef Medline
32. Hametner S, Wimmer I, Haider L, et al. **Iron and neurodegeneration in the multiple sclerosis brain.** *Ann Neurol* 2013;74:848–61 CrossRef Medline
33. Minagar A, Barnett MH, Benedict RH, et al. **The thalamus and multiple sclerosis: modern views on pathologic, imaging, and clinical aspects.** *Neurology* 2013;80:210–19 CrossRef Medline

What Causes Deep Gray Matter Atrophy in Multiple Sclerosis?

Multiple sclerosis is a chronic neuroinflammatory and neurodegenerative disease of the central nervous system. Patients often experience a complex combination of physical and cognitive symptoms, both of which are strongly disabling. Unfortunately, progression of disability and cognitive decline has been difficult to understand using neuroinflammatory markers such as lesion volumes. Neurodegenerative components of MS and especially deep gray matter (DGM) atrophy continue to progress with time[1] and have a strong predictive potential for disability[2] and cognitive impairment.[3] Thalamic atrophy occurs very early[4] and continues linearly during the disease course,[5] with especially strong clinical correlations.[6] Therefore, it is clear that deep gray matter and especially thalamic atrophy is of great relevance for MS, and its measurement may even become reliable enough to include in routine neuroradiologic practice. What drives this typical neurodegenerative pattern in MS, however, remains unclear, probably including a combination of network disconnection,[7] Wallerian degeneration, and local damage.[8]

The study by Pontillo et al,[9] published in the current issue of the *American Journal of Neuroradiology*, represents a comprehensive way to investigate the possible correlates of deep gray matter atrophy. The authors apply several MR imaging measures of diffusion, perfusion, and susceptibility in the DGM in relapsing-remitting MS (RRMS, $n = 52$) and progressive MS ($n = 25$), which were compared with those in healthy controls ($n = 44$). Results show that white matter lesion burden was the main correlate of DGM atrophy in RRMS, possibly indicating a role for Wallerian degeneration of connected fiber bundles, resulting in structural network disconnection and atrophy. In progressive MS, however, the most important correlates of atrophy were local microstructural damage and thalamic susceptibility, while lesion volumes did not strongly relate to atrophy.

These results highlight an important point, namely that the cause and consequence of atrophy could vary among the different MS phenotypes and that these should be studied separately.[10] This point is supported by recent findings that while some therapeutic options that target neuroinflammation in the white matter may impact thalamic atrophy in RRMS,[11] these do not impact disease progression in progressive MS.[12] Nonetheless, recent studies have shown that thalamic atrophy rates are similar in all phenotypes,[5] indicating that neurodegeneration continues in progressive MS even when the formation of new neuroinflammatory lesions may become less apparent. These findings could reflect an entirely different local pathologic process or may indicate a second-order disconnection effect[10] induced by an accelerated cortical degeneration of important networks such as the default mode network,[2] causing additional waves of disconnection leading to a so-called network collapse.[13]

This notion of network disconnection was also supported by a recent study using experimental autoimmune encephalomyelitis, showing inflammation and demyelination in the spinothalamic tracts to be related to thalamic neuronal loss, while lesions within the thalamus itself were scarce.[14] In MS, focal lesions within DGM structures also do not seem to be that common and appear to be poorly related to DGM atrophy.[15] In fact, neuronal loss in nondemyelinated DGM tissue can be as severe as 35%.[16] Other work[17] has also indicated that the DGM has a less severe neuroinflammatory profile than the white matter. However, diffuse microglial activation within the thalamus has also been noted using PET research, especially in progressive MS,[18] which was also related to cortical thinning and clinical dysfunction,[19] again indicating a network effect. It remains unclear, however, whether microglial activation is a cause of neurodegeneration or a consequence of it, or both. Susceptibility-weighted imaging as used in the present study by Pontillo et al[9] has also been indicated to reflect both microglial activation (ie, through changes in iron levels) and myelin content, further complicating matters.[20]

As Pontillo et al[9] note, future longitudinal multimodal studies are now required to disentangle the causal chain of events for these different local and network-based pathologic processes. It seems apparent, however, that the cause and consequence of DGM atrophy will remain a complex combination of primary and second-order effects. Thus, future treatment strategies aiming to impact DGM atrophy may need to impact the disease early, to prevent the network collapse from happening altogether.

REFERENCES

1. Eshaghi A, Marinescu RV, Young AL, et al. **Progression of regional grey matter atrophy in multiple sclerosis.** *Brain* 2018;141:1665–77 CrossRef Medline
2. Eshaghi A, Prados F, Brownlee WJ, et al. **Deep gray matter volume loss drives disability worsening in multiple sclerosis.** *Ann Neurol* 2018;83:210–22 CrossRef Medline
3. Eijlers AJ, Van Geest Q, Dekker I, et al. **Predicting cognitive decline in multiple sclerosis: a 5-year follow-up study.** *Brain* 2018;141: 2605–18 CrossRef Medline
4. Zivadinov R, Havrdová E, Bergsland N, et al. **Thalamic atrophy is associated with development of clinically definite multiple sclerosis.** *Radiology* 2013;268:831–41 CrossRef Medline
5. Azevedo CJ, Cen SY, Khadka S, et al. **Thalamic atrophy in multiple sclerosis: a magnetic resonance imaging marker of neurodegeneration throughout disease.** *Ann Neurol* 2018;83:223–34 CrossRef Medline
6. Minagar A, Barnett MH, Benedict RH, et al. **The thalamus and multiple sclerosis: modern views on pathologic, imaging, and clinical aspects.** *Neurology* 2013;80:210–19 CrossRef Medline
7. Schoonheim MM, Meijer KA, Geurts JJ. **Network collapse and cognitive impairment in multiple sclerosis.** *Front Neurol* 2015;6:82 CrossRef Medline
8. Kipp M, Wagenknecht N, Beyer C, et al. **Thalamus pathology in multiple sclerosis: from biology to clinical application.** *Cell Mol Life Sci* 2015;72:1127–47 CrossRef Medline
9. Pontillo G, Cocozza S, Lanzillo R, et al. **Determinants of deep gray matter atrophy in multiple sclerosis: a multimodal MRI study.** *AJNR Am J Neuroradiol* 2019;40:99–106 CrossRef
10. Larochelle C, Uphaus T, Prat A, et al. **Secondary progression in multiple sclerosis: neuronal exhaustion or distinct pathology?** *Trends Neurosci* 2016;39:325–39 CrossRef Medline
11. Gaetano L, Häring DA, Radue EW, et al. **Fingolimod effect on gray matter, thalamus, and white matter in patients with multiple sclerosis.** *Neurology* 2018;90:e1324–32 CrossRef Medline
12. Ziemssen T, Derfuss T, de Stefano N, et al. **Optimizing treatment success in multiple sclerosis.** *J Neurol* 2016;263:1053–65 CrossRef Medline
13. Schoonheim MM, Meijer KA, Geurts JJ. **Network collapse and cognitive impairment in multiple sclerosis.** *Front Neurol* 2015;6:82 CrossRef Medline
14. Wagenknecht N, Becker B, Scheld M, et al. **Thalamus degeneration and inflammation in two distinct multiple sclerosis animal models.** *J Mol Neurosci* 2016;60:102–14 CrossRef Medline
15. van de Pavert SH, Muhlert N, Sethi V, et al. **DIR-visible grey matter lesions and atrophy in multiple sclerosis: partners in crime?** *J Neurol Neurosurg Psychiatry* 2016;87:461–67 CrossRef Medline
16. Cifelli A, Arridge M, Jezzard P, et al. **Thalamic neurodegeneration in multiple sclerosis.** *Ann Neurol* 2002;52:650–53 CrossRef Medline
17. Vercellino M, Masera S, Lorenzatti M, et al. **Demyelination, inflammation, and neurodegeneration in multiple sclerosis deep gray matter.** *J Neuropathol Exp Neurol* 2009;68:489–502 CrossRef Medline
18. Rissanen E, Tuisku J, Vahlberg T, et al. **Microglial activation, white matter tract damage, and disability in MS.** *Neurol Neuroimmunol Neuroinflamm* 2018;5:e443 CrossRef Medline
19. Herranz E, Gianni C, Louapre C, et al. **Neuroinflammatory component of gray matter pathology in multiple sclerosis.** *Ann Neurol* 2016;80:776–90 CrossRef Medline
20. Schweser F, Raffaini Duarte Martins AL, Hagemeier J, et al. **Mapping of thalamic magnetic susceptibility in multiple sclerosis indicates decreasing iron with disease duration: a proposed mechanistic relationship between inflammation and oligodendrocyte vitality.** *Neuroimage* 2018;167:438–52 CrossRef Medline

M.M. Schoonheim
J.J.G. Geurts
Department of Anatomy and Neurosciences
MS Center Amsterdam
Amsterdam Neuroscience
Amsterdam UMC, Vrije Universiteit Amsterdam
Amsterdam, the Netherlands

http://dx.doi.org/10.3174/ajnr.A5942

Clinical Feasibility of Zero TE Skull MRI in Patients with Head Trauma in Comparison with CT: A Single-Center Study

S.B. Cho, H.J. Baek, K.H. Ryu, B.H. Choi, J.I. Moon, T.B. Kim, S.K. Kim, H. Park, and M.J. Hwang

ABSTRACT

BACKGROUND AND PURPOSE: Conventional MR imaging techniques cannot produce optimal images of bone structures because bone has little water and a very short T2 life span. The aim of this study was to investigate the clinical feasibility of skull MR imaging using the zero TE sequence in patients with head trauma by assessing its diagnostic image quality and quantitative measurement compared with CT images.

MATERIALS AND METHODS: Thirteen enrolled patients with head trauma were assessed using brain CT and skull MR imaging. Image quality was graded on a 5-point Likert scale to compare the 2 modalities. To evaluate quantitative analyses between the 2 imaging modalities, we measured skull thickness and normalized bone tissue signal. Interobserver reliability was assessed using weighted κ statistics and the intraclass correlation coefficient.

RESULTS: Both imaging techniques clearly depicted skull fractures in all 13 patients. The mean scores for skull MR imaging and CT were 4.65 ± 0.56 and 4.73 ± 0.45 ($P = .157$), respectively, with substantial interobserver agreement ($P < .05$). The 2 imaging modalities showed no difference in skull thickness ($P = .092$) and had good correlation ($r^2 = 0.997$). The mean value of normalized bone tissue signal among the 3 layers of the skull was relatively consistent ($P = .401$) with high interobserver agreement ($P < .001$).

CONCLUSIONS: Zero TE skull MR imaging has diagnostic image quality comparable with that of CT images. It also provides consistent results on the quantitative measurement of cortical bone with CT images.

ABBREVIATION: ZTE = zero TE

MR imaging is a noninvasive technique that obtains excellent soft-tissue contrast and high resolution of anatomic detail in the body without radiation.[1] However, it is unsuitable for depicting cortical bone structures because of low proton density (approximately 20% water) and a very short T2 relaxation time (approximately 390 μs at 3T).[2] By contrast, CT is the most optimal technique for revealing bone structures in images with high spatial resolution, fast acquisition, and high availability; however, its capability in imaging soft tissue is poor, and radiation exposure is a major drawback.[3-8] In clinical practice, MR imaging is an essential diagnostic technique because of its inherent advantage in establishing a diagnosis and treatment plan for various intracranial diseases. Therefore, MR imaging would be the most ideal diagnostic imaging technique if it could provide reliable clinical information about bone structures and soft tissues. There is an increasing clinical need to resolve the limitations of MR imaging with regard to bone structures. MR bone imaging is increasingly becoming a focus of interest in the field of musculoskeletal radiology, MR imaging–based attenuation correction in PET, MR imaging–based radiation therapy planning, and MR imaging–guided focused sonography.[9]

Concerning the technical aspects of conventional MR imaging, optimal images of bone structures cannot be obtained because the minimum TE for the spin-echo and gradient-echo pulse sequences is too slow (approximately 8–10 and 1–2 ms, respectively) to detect a meaningful bone signal.[2] Compared with conventional sequences, the ultrashort TE sequence with a center-out *k*-space acquisition can allow sufficiently fast data acquisition of the rapidly decaying bone signal.[10,11] Long T2-suppression methods (eg, echo subtraction, saturation prepulses, or multiple sequences) are generally applied to separate bone from soft tissue, to

Received July 30, 2018; accepted after revision November 1.

From the Departments of Radiology (S.B.C., H.J.B., K.H.R., B.H.C., J.I.M., T.B.K.) and Neurosurgery (S.K.K., H.P.), Gyeongsang National University Changwon Hospital, Gyeongsang National University School of Medicine, Changwon, Republic of Korea; and MR Applications and Workflow (M.J.H.), GE Healthcare Korea, Seoul, Republic of Korea.

Please address correspondence to Hye Jin Baek, MD, PhD, Department of Radiology, Gyeongsang National University Changwon Hospital, Gyeongsang National University School of Medicine, 11 Samjeongja-ro, Seongsan-gu, Changwon 51472, Republic of Korea; e-mail: sartre81@gmail.com

http://dx.doi.org/10.3174/ajnr.A5916

achieve selective bone images using an ultrashort TE sequence.[12] The sequence for MR bone imaging was recently developed to visualize very short T2 relaxation of the object, based on 3D radial zero TE (ZTE), and it can provide high-resolution isotropic images with fast and silent scanning.[13-15] To date, MR bone imaging using ultrashort TE or ZTE sequences has been studied for PET/MR imaging attenuation correction from the perspective of the technical approach in the literature.[9,16-20]

Recently, ZTE bone MR imaging was applied to osseous shoulder imaging, which showed strong intermodality agreement between measurements and grades of the lesions from ZTE bone MR imaging and CT.[21] However, to the best of our knowledge, there is no study of the clinical application of MR bone imaging of the skull for diagnostic use. Therefore, this study aimed to investigate the clinical feasibility of using ZTE skull MR imaging for evaluating skull lesions in patients with head trauma by assessing its diagnostic image quality and quantitative measurement compared with CT images.

MATERIALS AND METHODS

Patient Characteristics

A review of the data base at Gyeongsang National University Changwon Hospital identified patients with head trauma who were admitted to the emergency department and underwent routine brain CT for the evaluation of intracranial or skull abnormalities between June 2017 and May 2018. Using electronic medical records and a PACS, we selected 16 patients who also had undergone ZTE skull MR imaging as the follow-up imaging. We ultimately enrolled 13 of 16 patients. Three patients were excluded because of poor image quality from uncontrolled motion artifacts (CT: $n = 2$; MR imaging: $n = 1$). The 13 patients who were included in this study comprised 4 women and 9 men with a mean age of 43.8 ± 12.7 years (range, 24–68 years). The average interval between the initial CT and ZTE skull MR imaging examinations was 10.1 ± 7.4 days (range, 1–23 days).

In the present study, all retrospective data collection and analyses were conducted in accordance with our local institutional review board guidelines, after obtaining its approval. The institutional review board determined that patient approval and informed consent were not required for retrospective review of images and electrical medical records. The patients' records and information were anonymized and de-identified before analysis.

ZTE Skull MR Imaging

In general, ZTE uses a nonselective radiofrequency excitation pulse and a 3D radial center-out *k*-space trajectory. The readout gradients are ramped up before the radiofrequency excitation (Fig 1*A*). The pulse is grouped into segments, and each segment contains numerous spokes. These gradients are changed slightly and are not ramped down between spokes. The minimal gradient switching between repetitions minimizes eddy currents to a negligible level and reduces acoustic noise level.[22] The imaging encoding starts immediately after the end of the radiofrequency excitation to fill the center of the *k*-space, which is the nominal ZTE. The radiofrequency duration must remain short (approximately 8–16 μs), and the flip angle is limited in ZTE to minimize the delay between the radiofrequency pulse and the transmit-to-receive switching time. In addition, to prevent disturbing the spin excitation substantially, the excitation bandwidth is limited. As a result, the contrast is a native proton-density contrast.

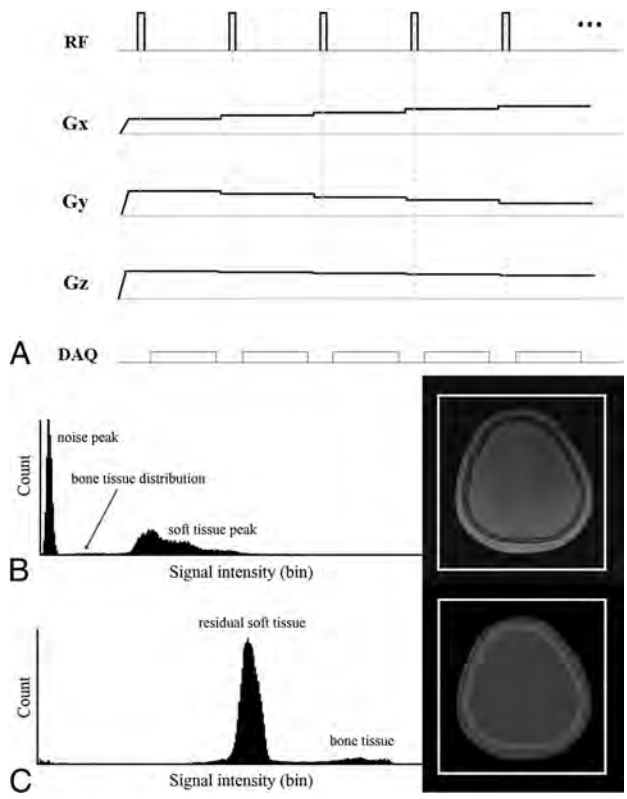


FIG 1. *A*, ZTE pulse sequence diagram of a segment with 5 spokes (the actual scan contains 384 spokes per segment). The ZTE sequence uses a hard radiofrequency (RF) pulse and switch data acquisition (DAQ) just after the RF to receive the free induction decay signal, which results in a nominal ZTE. The *k*-space fills the 3D radial center. The readout gradients (Gx, Gy, Gz) are ramped up before the RF and change in small steps. Such a small gradient change results in a fast, silent scan. *B* and *C*, Histogram-based intensity-correction has been used to generate the CT-like bone image. There are 2 typical intensity histograms and corresponding axial images. *B*, The ZTE dataset of proton-density contrast. *C*, The final dataset of the CT-like contrast image with further postprocessing. After we applied inversion logarithmic image rescaling, the CT-like contrast image reveals excellent cortical bone delineation. On each histogram, the horizontal axis represents the signal intensity and the vertical axis represents the count.

After acquiring the proton-density image using ZTE, we applied a bias-correction algorithm to neaten the soft-tissue signal intensities and correct signal inhomogeneity owing to coil geometry and variable tissue cross-sections.[9,20,21] The histogram distribution of the inverse log-scaled images was then used for the CT-like contrast images to remove the background noise while retaining bone and soft tissue (Fig 1*B*, -*C*).[9,20,21] The histogram of the proton-density image easily yields 2 groups: 1) soft tissues such as white matter, gray matter, CSF, muscle, and fat; and 2) background air. Then, the median of the tissue signals, which are magnified to the proper level, is used to determine the threshold values. The imaging value, which is higher than the bone threshold, is magnified and generates a CT-like contrast image.

ZTE Skull MR Imaging Scan Parameters

MR imaging was performed using a 3T system (Signa Architect; GE Healthcare, Milwaukee, Wisconsin) with a 48-channel head

coil. The ZTE proton-density image was acquired in the axial plane to cover the whole brain in 4 minutes 52 seconds using the following parameters: TE, 0.016 ms (nominal TE = 0); TR, 2.65 ms; FOV, 240 × 240 mm; slice thickness, 2 mm; flip angle, 1°; spokes per segment, 384; matrix size, 288 × 288; image voxel resolution, 0.8 × 0.8 × 1 mm^3; receiver bandwidth, ±31.25 kHz; and total number of scans, 3.

Brain CT Scan Parameters

Skull CT images were obtained using the routine brain protocol on 2 different CT machines with the following acquisition parameters: 1) IQon Spectral CT (Philips Healthcare, Best, the Netherlands)—120 kV(peak); 200 mAs; collimation, 16 × 0.625 mm; pitch factor, 0.985; rotation time, 0.33 seconds; FOV, 250 mm; slice thickness, 3 mm; and slice increment, 0.4 mm; and 2) Aquilion ONE/ViSION Edition CT (Toshiba Medical Systems, Tokyo, Japan)—120 kVp; 190 mAs; collimation, 80 × 0.5 mm; pitch factor, 0.985; rotation time, 0.75 seconds; FOV, 240 mm; slice thickness, 3 mm; and slice increment, 0.4 mm.

Image Analyses

All datasets were anonymized with randomization. Two readers reviewed all images using the PACS. Two attending neuroradiologists with 11 and 8 years of experience independently analyzed all ZTE skull MR images and CT scans to evaluate the image quality of skull MR imaging from a clinical feasibility perspective. They also evaluated whether a skull fracture was present in both image sets. In the review of all images, the window width and window level could be modified for evaluation. The 2 types of skull images were assessed separately to minimize bias because of the results of the other images. For the initial interpretation, each reader was provided the ZTE skull MR images, and they analyzed all images twice with an interval of 2 weeks between each analysis. After 2 weeks, the readers were provided the CT images of the enrolled patients. They analyzed all these images twice, using the same interval as in ZTE skull MR imaging interpretation. Image-quality measures for each image set were evaluated for following items: 1) the conspicuity and differentiation of 3 layers of skull structures (ie, outer table, diploic layer, and inner table), 2) clear visualization of the suture lines, 3) visualization and demarcation of the fracture line, and 4) the presence of artifacts. The readers then scored image quality on the following 5-point Likert scale: 1) nondiagnostic (ie, not acceptable for diagnostic use); 2) unacceptable (ie, not acceptable for diagnostic use); 3) sufficient (ie, acceptable for diagnostic use but with minor issues); 4) diagnostic (ie, acceptable for diagnostic use); and 5) excellent (ie, acceptable for diagnostic use).

To evaluate the geometric concordance of skull structures on the ZTE skull MR imaging and CT images, the readers measured skull thickness manually from the most inner cortex to the most outer cortex in 6 regions from each corresponding slice of the same patient on the MR imaging and CT images. For each patient, 3 slices were selected to measure skull thickness in both frontal, both parietal, both temporal, and both occipital bones.

The same readers also measured the signal intensities and Hounsfield units of the outer table, diploic layer, and inner table on the ZTE skull MR imaging and CT images. The readers defined the ROI as 15 mm^3 symmetrically in the outer cortex, diploic space, and inner cortex from each corresponding slice of the same patient on the MR imaging and CT images. The same neuroradiologists manually drew 6 ROIs in both frontal, both parietal, and both occipital bones from each patient on 3 selected slices on the MR imaging and CT images. To normalize the signal intensity of ZTE skull MR imaging, we calculated the normalized signal intensity of each skull layer by dividing its signal intensity by that of the background region and then multiplied this value by 100. We calculated the ratio of bone signal intensity of ZTE skull MR imaging to the bone density of CT images and termed this value the "normalized bone tissue signal." After the readers' independent analyses, the same neuroradiologists conducted image analysis in consensus to make a reference standard for deciding the presence of skull fracture and its location.

Statistical Analysis

The data were tested for normal distribution using the Kolmogorov-Smirnov test. Continuous variables were expressed as mean ± SD. The image-quality assessments of ZTE skull MR imaging and CT were assigned numeric values. The mean values of the readers' ratings were not directly statistically compared because these values were not strictly continuous variables. However, we decided to present a summary of the readers' ratings for ZTE skull MR imaging and CT, which are expressed as mean ± SD. The scores of each image set from the 2 readers were averaged, and the Wilcoxon signed rank test was used to compare the scores of the ZTE skull MR imaging and CT images. Interobserver agreement between 2 readers was calculated by weighted κ statistics. On the basis of the description by Landis and Koch,[23] the κ value was interpreted as follows: 0, no agreement; 0–0.19, poor agreement; 0.20–0.39, fair agreement; 0.40–0.59, moderate agreement; 0.60–0.79, substantial agreement; and 0.80–1.00, nearly perfect agreement. The skull thickness was compared and correlated between ZTE skull MR imaging and CT images using the Wilcoxon signed rank test and Spearman correlation coefficient analysis after the values of each image set from the 2 readers were averaged. The Kruskal-Wallis test was used to evaluate differences in the normalized bone tissue signal among the 3 layers of the skull. The interobserver agreement between the 2 readers was also assessed with the intraclass correlation coefficient. All statistical analyses were conducted using SPSS, Version 24.0 (IBM, Armonk, New York). A P value $< .05$ was statistically significant.

RESULTS

Skull images were successfully obtained from all patients using the ZTE technique. Skull structures were clearly depicted on the ZTE skull MR images and matched well with those depicted on the CT images. Two attending neuroradiologists independently reviewed all acquired ZTE skull MR imaging and CT images without difficulty in detecting the fracture sites. All 13 patients had a skull fracture (Figs 2 and 3). In addition, the skull suture lines were conspicuously demarcated on the ZTE skull MR imaging in all 13 patients.

Patient data and the scores of the overall image quality for ZTE skull MR imaging and CT images determined by the 2 readers are presented in the Table. The mean scores of ZTE skull MR imaging

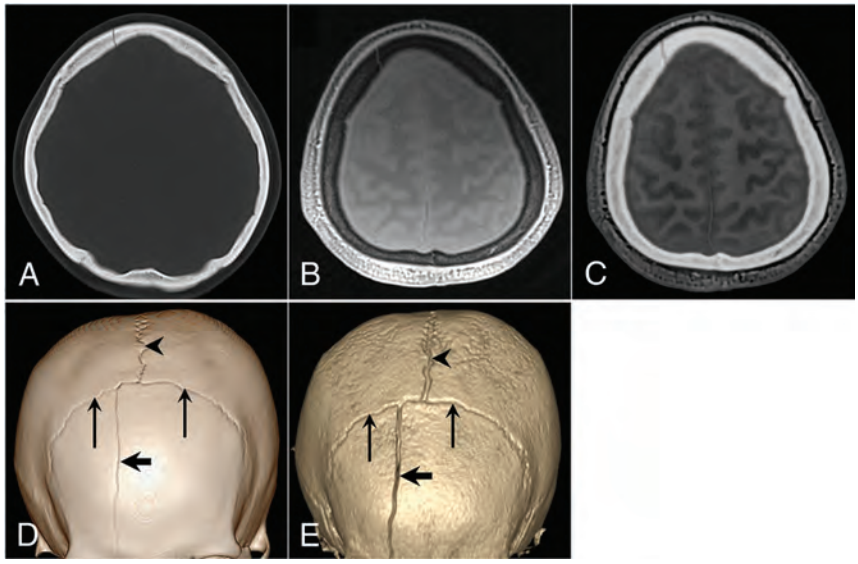


FIG 2. A 42-year-old woman with a right frontal bone fracture. *A*, Axial CT image. *B*, Axial proton-density ZTE image. *C*, Axial CT-like contrast ZTE image. 3D volume-rendered CT image (*D*) and ZTE skull MR imaging (*E*). A linear skull fracture line is visible in the right frontal bone on the CT and MR images (*thick arrows*). The coronal sutures (*thin arrows*) and sagittal sutures (*arrowheads*) are depicted.

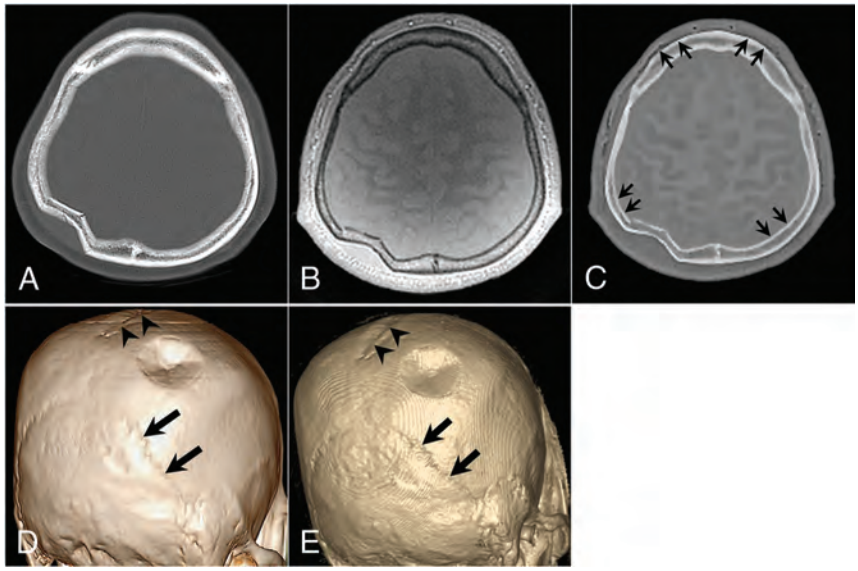


FIG 3. A 52-year-old man with a right parietal bone fracture. *A*, Axial CT image. *B*, Axial proton-density ZTE image. *C*, Axial CT-like contrast ZTE image. 3D volume-rendered CT image (*D*) and ZTE skull MR imaging (*E*). A focal depression fracture is visible in the right parietal bone. The sagittal suture (*arrowheads*) and bilateral lambdoidal sutures (*thick arrows*) show conspicuous delineation. Subtle marginal artifacts exist along the inner cortex of both parietal bones and the outer cortices of both frontal bones. The artifacts have a short segmental stepped appearance, which may be related to the postprocessing of histogram-based intensity correction (*short thin arrows*) in *C*.

and CT images were >4 points with acceptable image quality for diagnostic use. For both readers, the mean score of ZTE skull MR was slightly lower than that of CT images. However, the average mean scores for skull MR imaging and brain CT images by the readers were not statistically significant (4.65 ± 0.56 versus 4.73 ± 0.45, $P = .157$). Substantial interobserver agreement was observed for the overall image quality of skull MR imaging ($\kappa = 0.829$, $P = .001$) and brain CT ($\kappa = 0.806$, $P = .003$).

The skull thickness on ZTE skull MR imaging and CT images was not statistically different: 9.5 ± 4.3 and 9.5 ± 4.2 mm, respectively ($P = .092$). Skull thickness measured on the ZTE skull MR imaging and CT images also showed good correlation ($r^2 = 0.997$, $P < .001$). Based on the results of the intraclass correlation coefficient, interobserver agreement of the skull thickness was as follows: CT, 0.989 (95% CI, 0.984–0.992; $P < .001$) and MR imaging, 0.977 (95% CI, 0.967–0.985; $P < .001$). The mean value of the normalized bone tissue signal tended to be higher in the diploic space than in the outer and inner cortices (0.45 ± 0.13 versus 0.43 ± 0.11 versus 0.41 ± 0.01, respectively); however, there was no statistical difference among the 3 layers of the skull ($P = .401$) (Fig. 4). For interobserver agreement, the intraclass correlation coefficient for the normalized bone tissue signal was 0.885 (95% CI, 0.851–0.911; $P < .001$).

DISCUSSION

The findings of our study indicated that ZTE skull MR imaging was suitable for identifying bone structures in the skull. Its diagnostic image quality was comparable with that of CT images for evaluating traumatic skull lesions. In addition, this study provided a quantitative evaluation of ZTE skull MR images by a direct comparison with CT images and showed good correlation between the 2 skull images.

To date, MR bone imaging is challenging because the magnetization from hydrogen atoms in the cortical bone demonstrates much faster transversal relaxation than other body tissues, and the available magnetization is relatively low because of decreased proton density.[24] This phenomenon is caused by the characteristics of bone tissue, which has little water and a very short T2 life span.[2] With technical advances, MR bone imaging has proved technically feasible using ultrashort TE and ZTE sequences in previous studies.[10-15] In contrast to ultrashort TE sequences, MR bone imaging using the ZTE sequence has a greater signal-to-noise ratio with scan time efficiencies and it provides isotropic high-resolution images with multiplanar reconstruction.[13-15,21] To date, a few subsequent studies have focused on MR bone imaging in the field of PET.[2,9,11,16-20] However, MR bone imaging has not been studied from a diagnostic perspective in the field of neuroradiology, though interest in this topic is increasing.

Summary of the patient data and the scores of the 2 readers

Patient No.	Age (yr)	Sex	Fracture Location	Interval between ZTE Skull MRI and CT	ZTE Skull MRI		Brain CT	
					Reader 1	Reader 2	Reader 1	Reader 2
1	52	M	Right parietal	20 Days	5	5	5	5
2	42	F	Right frontal	17 Days	5	5	5	5
3	47	M	Right temporal	8 Days	5	5	5	5
4	68	M	Right occipital	23 Days	4	4	4	5
5	25	F	Right temporal	2 Days	3	4	4	4
6	24	M	Left occipital	3 Days	5	5	5	5
7	38	M	Right frontal	11 Days	4	4	4	4
8	41	M	Left temporal	1 Day	5	5	5	5
9	57	M	Left parietal	7 Days	5	5	5	5
10	53	M	Right frontal	14 Days	4	4	4	4
11	52	F	Right temporal	16 Days	5	5	5	5
12	31	M	Right temporal	2 Days	5	5	5	5
13	41	F	Right parietal	7 Days	5	5	5	5

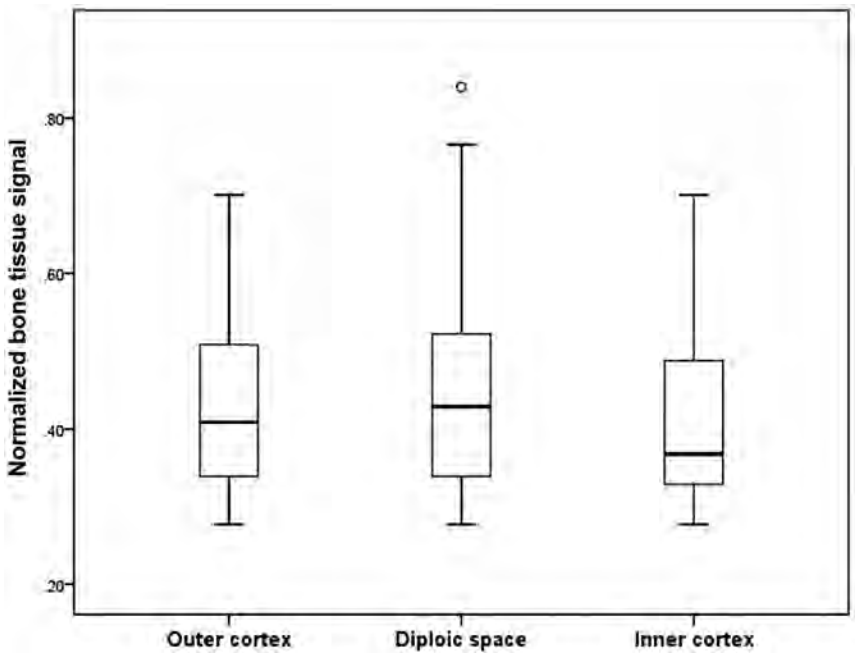


FIG 4. *Boxplot* of the normalized bone tissue signal of the outer cortex, diploic space, and inner cortex. The *line* across the box represents the median value. The box *ends* represent the first and third quartiles. The *end points* in each graph represent the smallest and largest values. The median ratio of the normalized bone tissue signal is highest in the diploic space of the skull; however, there is no significant difference among the 3 skull layers.

In the present study, we applied the ZTE sequence to obtain images of the skulls of patients with head trauma. Our findings showed results similar to those of previous studies from the perspective of its clinical feasibility in assessing the head.[9,19,20] For all 13 patients, ZTE skull MR imaging was successfully acquired, and skull suture lines with skull fractures were depicted as clearly on the ZTE skull MR imaging as on the CT images. Our results for detecting bone abnormalities were similar to those of a recent study,[21] which showed comparable agreement for evaluating osseous lesions in the shoulder between raters and modalities. In the present study, the CT-like image with positive contrast for the skull obtained by postprocessing allowed more intuitive interpretation for evaluating the skull, and it is also consistent with the previous study, despite differences in the applied anatomic sites.

We also assessed the overall imaging quality of skull MR imaging, and ours is the first study to evaluate the image quality of skull MR imaging for diagnostic use. ZTE skull MR imaging yielded acceptable image quality of >4 points with substantial interobserver agreement in the present study. There was no significant difference, though the mean score of ZTE skull MR images was slightly lower than that of CT images because moderate motion artifacts of a few slices occurred in 1 patient owing to the patient's irritability, which was closely associated with the longer scan time of ZTE skull MR imaging than CT. For clinical use, further technologic effort to reduce the scan time is essential to expand the indications of ZTE skull MR imaging for various clinical situations.

We were also able to identify subtle marginal irregularities on the only CT-like images of ZTE skull MR imaging (2/13 patients, 15.4%), though there were no demonstrable motion artifacts in the 2 patients. These artifacts had a thin, short segmental stepped appearance in the inner or outer cortex of both frontal and parietal bones (Fig 3*C*); however, these artifacts did not have a crucial impact for diagnosis by the 2 readers. The reason for this finding is unclear, but it may be related to the postprocessing of histogram-based intensity correction to separate bone tissue from other tissue. We expect that this issue of marginal stepped artifacts can be solved if the algorithm for the postprocessing of ZTE skull MR imaging is improved.

In the current study, the 2 imaging modalities showed no statistically significant differences in the skull thickness measurement, and they had good correlation in skull thickness measurements. In contrast to the finding of a previous study,[21] we measured the normalized bone tissue signal and found that the normalized bone tissue signal from each skull layer was relatively consistent on the basis of a direct comparison of the signal-to-noise ratio on ZTE skull MR imaging and Hounsfield units on CT for the same regions. In addition, the skull-thickness measure-

ments and the normalized bone tissue signal showed high interobserver agreement. The mean value of the normalized bone tissue signal was higher in the diploic space than in the other skull layers. This result could reflect the characteristics of the tissue in the diploic space, which has a fat component in the marrow cavity; this factor results in a short T1 relaxation time in MR imaging and a decreased number of Hounsfield units on the CT image. These quantitative comparison results of the 2 imaging modalities imply that ZTE skull MR imaging could reproduce images close to the skull itself and suggest that ZTE skull MR imaging could be a valid alternative to CT for skull imaging in a variety of clinical situations. Furthermore, these results also support the findings of previous studies[9,20] that suggested possible technical applications to develop attenuation correction algorithms.

With regard to the technical aspect, hard pulse sequences such as ZTE require a higher readout bandwidth to allow shorter encoding times and less T2* blurring.[9,25] The flip angle below the corresponding Ernst angle is also important to obtain native proton-density–weighted images and achieve appropriate tissue differentiation during the postprocessing.[9] In a previous study,[9] the authors performed ZTE skull MR imaging based on the default high-resolution protocols using different imaging bandwidths (±31.25 kHz, ±62.5 kHz, and ±125 kHz) and flip angles (0.6°, 1.2°, and 2.4°). They found that images with the highest bandwidth and lowest flip angle were sharpest and had the least blurring at the tissue interfaces, whereas the images with the lowest bandwidth and highest flip angle had the highest signal-to-noise ratio and soft-tissue contrast with partial T1 saturation. In the present study, we used ±31.25 kHz of bandwidth and 1° of flip angle with fewer sophisticated postprocessing steps to acquire ZTE skull MR imaging; these images were different from those used in previous studies.[9,20,21,26] We initially aimed to evaluate the clinical diagnostic feasibility of ZTE skull MR imaging for detecting skull fracture, compared with CT; therefore, we focused on obtaining sufficient image quality with an optimal signal-to-noise ratio to interpret the fracture and suture lines in the skull on the visual analysis. With this perspective of visual qualitative analysis, we were not concerned about other structures such as the paranasal sinuses in the facial bones, which require a higher bandwidth for depicting in detail.

Several limitations of this study should be considered when interpreting the results. First, there was an unavoidable selection bias because the data from all patients were evaluated retrospectively, the sample size was small, and the study was conducted in a single center. Second, in this study, we evaluated only patients with head trauma; therefore, the representation of other skull pathologies is limited. This feature may have a different effect on the image quality for interpretation. Third, we did not use the sophisticated formula in the previous study[26] to obtain the normalized bone tissue signal because we were focused on the morphologic perspective of ZTE skull MR imaging with simple and easy postprocessing. In the current study, it was sufficient to apply a simple noise threshold and a bias-correction algorithm to enhance bone-tissue signal intensities and correct signal inhomogeneity due to coil geometry; therefore, complex scaling was not required. However, this approach to obtain the normalized bone tissue signal had an inherent limitation of value consistency because it was scaled to the background noise. A relatively low signal-to-noise ratio in ZTE skull MR imaging may indeed induce inhomogeneity of the signal intensity, which may then affect the results of the normalized bone tissue signal. Fourth, manual ROI selection, which was used for normalization, was reader-dependent and small; therefore, it may have affected the results, though we analyzed the interobserver reliability. Fifth, the ZTE sequence allows fast and quiet scanning to obtain skull images; however, it takes approximately 5 minutes, which allows the possibility of motion artifacts. We expect that further studies with larger sample sizes and various targeted patients will be conducted to validate our results in the near future.

CONCLUSIONS

In the current study, ZTE skull MR imaging generated a CT-like image with positive contrast for the skull by postprocessing, and it showed diagnostic image quality comparable with that of CT images for evaluating suture lines and traumatic skull lesions. It also showed a good correlation with CT images in skull-thickness measurement, and the normalized bone tissue signal was relatively consistent. Therefore, we believe that ZTE skull MR imaging may broaden the indications of MR imaging examinations, especially for radiosensitive patients with trauma such as pediatric patients or pregnant women because of its inherent benefit of not generating radiation. In addition, ZTE skull MR imaging may be helpful in MR imaging–dependent technologies such as PET/MR imaging reconstruction or MR imaging–guided radiation therapy and in procedures that provide useful information regarding attenuation correction or anatomic details.

ACKNOWLEDGMENTS

The authors would like to thank Michael Carl and Robert D. Peters at GE Healthcare for their advice regarding technical background of ZTE sequences and its postprocessing steps during the revision process. We also thank Elsevier Language Editing Service for the English language review and editing (http://webshop.elsevier.com/languageservices/languageediting/).

REFERENCES

1. Blystad I, Warntjes JB, Smedby O, et al. **Synthetic MRI of the brain in a clinical setting.** *Acta Radiol* 2012;53:1158–63 CrossRef Medline
2. Du J, Carl M, Bydder M, et al. **Qualitative and quantitative ultrashort echo time (UTE) imaging of cortical bone.** *J Magn Reson* 2010; 207:304–11 CrossRef Medline
3. Zeman RK, Fox SH, Silverman PM, et al. **Helical (spiral) CT of the abdomen.** *AJR Am J Roentgenol* 1993;160:719–25 CrossRef Medline
4. Kalender WA, Seissler W, Klotz E, et al. **Spiral volumetric CT with single-breath-hold technique, continuous transport, and continuous scanner rotation.** *Radiology* 1990;176:181–83 CrossRef Medline
5. Crawford CR, King KF. **Computed tomography scanning with simultaneous patient translation.** *Med Phys* 1990;17:967–82 CrossRef Medline
6. Vock P, Soucek M, Daepp M, et al. **Lung: spiral volumetric CT with single-breath-hold technique.** *Radiology* 1990;176:864–67 CrossRef Medline
7. Mettler FA Jr, Bhargavan M, Faulkner K, et al. **Radiologic and nuclear medicine studies in the United States and worldwide: frequency, radiation dose, and comparison with other radiation sources—1950–2007.** *Radiology* 2009;253:520–31 CrossRef Medline
8. Schauer DA, Linton OW. **National Council on Radiation Protection**

and Measurements report shows substantial medical exposure increase. *Radiology* 2009;253:293–96 CrossRef Medline
9. Wiesinger F, Sacolick LI, Menini A, et al. **Zero TE MR bone imaging in the head.** *Magn Reson Med* 2016;75:107–14 CrossRef Medline
10. Robson MD, Gatehouse PD, Bydder M, et al. **Magnetic resonance: an introduction to ultrashort TE (UTE) imaging.** *J Comput Assist Tomogr* 2003;27:825–46 CrossRef Medline
11. Robson MD, Bydder GM. **Clinical ultrashort echo time imaging of bone and other connective tissues.** *NMR Biomed* 2006;19:765–80 CrossRef Medline
12. Du J, Bydder M, Takahashi AM, et al. **Short T2 contrast with three-dimensional ultrashort echo time imaging.** *Magn Reson Imaging* 2011;29:470–82 CrossRef Medline
13. Madio DP, Lowe IJ. **Ultra-fast imaging using low flip angles and FIDs.** *Magn Reson Med* 1995;34:525–29 CrossRef Medline
14. Grodzki DM, Jakob PM, Heismann B. **Ultrashort echo time imaging using pointwise encoding time reduction with radial acquisition (PETRA).** *Magn Reson Med* 2012;67:510–18 CrossRef Medline
15. Idiyatullin D, Corum C, Park JY, et al. **Fast and quiet MRI using a swept radiofrequency.** *J Magn Reson* 2006;181:342–49 CrossRef Medline
16. Keereman V, Fierens Y, Broux T, et al. **MRI-based attenuation correction for PET/MRI using ultrashort echo time sequences.** *J Nucl Med* 2010;51:812–88 CrossRef Medline
17. Catana C, van der Kouwe A, Benner T, et al. **Toward implementing an MRI-based PET attenuation-correction method for neurologic studies on the MR-PET brain prototype.** *J Nucl Med* 2010;51:1431–38 CrossRef Medline
18. Johansson A, Karlsson M, Nyholm T. **CT substitute derived from MRI sequences with ultrashort echo time.** *Med Phys* 2011;38:2708–14 CrossRef Medline
19. Delso G, Zeimpekis K, Carl M, et al. **Cluster-based segmentation of dual-echo ultra-short echo time images for PET/MR bone localization.** *EJNMMI Phys* 2014;1:1–13 CrossRef Medline
20. Delso G, Wiesinger F, Sacolick LI, et al. **Clinical evaluation of zero-echo-time MR imaging for the segmentation of the skull.** *J Nucl Med* 2015;56:417–22 CrossRef Medline
21. Breighner RE, Endo Y, Konin GP, et al. **Technical developments: zero echo time imaging of the shoulder—enhanced osseous detail by using MR imaging.** *Radiology* 2018;286:960–66 CrossRef Medline
22. Alibek S, Vogel M, Sun W, et al. **Acoustic noise reduction in MRI using Silent Scan: an initial experience.** *Diagn Interv Radiol* 2014;20:360–63 CrossRef Medline
23. Landis JR, Koch GG. **The measurement of observer agreement for categorical data.** *Biometrics* 1977;33:159–74 CrossRef Medline
24. Johnson EM, Vyas U, Ghanouni P, et al. **Improved cortical bone specificity in UTE MR imaging.** *Magn Reson Med* 2017;77:684–95 CrossRef Medline
25. Grodzki DM, Jakob PM, Heismann B. **Correcting slice selectivity in hard pulse sequences.** *J Magn Reson* 2012;214:61–67 CrossRef Medline
26. Wiesinger F, Bylund M, Yang J, et al. **Zero TE-based pseudo-CT image conversion in the head and its application in PET/MR attenuation correction and MR-guided radiation therapy planning.** *Magn Reson Med* 2018;80:1440–51 CrossRef Medline

Coil Embolization in Patients with Recurrent Cerebral Aneurysms Who Previously Underwent Surgical Clipping

S.-T. Kim, J.W. Baek, S.-C. Jin, J.H. Park, J.S. Kim, H.Y. Kim, H.W. Jeong, and Y.G. Jeong

ABSTRACT

BACKGROUND AND PURPOSE: Surgical revision of recurrent cerebral aneurysms is technically difficult. Therefore, coil embolization has been used as an alternative in these cases. The aim of this study was to evaluate the clinical and angiographic outcomes of coil embolization in patients with recurrent cerebral aneurysms after microsurgical clipping.

MATERIALS AND METHODS: Between May 1999 and February 2016, nineteen patients with 19 recurrent aneurysms who previously underwent surgical clipping were treated by coil embolization.

RESULTS: Nine patients presented with subarachnoid hemorrhage (47.4%). The interval between surgical clipping and coil embolization was 143.5 ± 66.1 months (range, 43–276 months). Single- or double-catheter coil embolization was performed in 16 patients. A balloon (n = 1) and stents (n = 2) were used to assist the coil embolization in 3 patients. Immediate radiologic findings after coil embolization showed complete occlusion in 10 patients, a residual neck in 8 patients, and a residual sac in 1 patient. Procedure-related permanent morbidity occurred in 1 patient. The mean clinical follow-up was 58.3 ± 38.8 months. Poor clinical outcomes (modified Rankin Scale score = ≥3) at the end of the clinical follow-up were reported in 5 patients (26.3%). Angiographic follow-up was available for 12 patients (63.2%). Major recurrence was detected in 5 patients (41.7%), and a tendency for aneurysm regrowth rather than coil compaction was noted in all cases.

CONCLUSIONS: In our series, coil embolization for recurrent aneurysms after surgical clipping was feasible but had a high recurrence rate and tended to result in aneurysm regrowth rather than coil compaction.

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

Microsurgical revision of recurrent cerebral aneurysms after surgical clipping is technically difficult because adhesions between the aneurysm and neighboring structures increase the possibility of injury to the adjacent normal structures. However, endovascular coil embolization, which does not require approaching the adhesions of previous surgical wounds, could be technically feasible as a retreatment option for recurrent cerebral aneurysms after surgical clipping. Some reports exist on the usefulness of coil embolization for recurrent cerebral aneurysms after surgical clipping.[1-5] The durability of coil embolization of recurrent cerebral aneurysms previously treated by surgical clipping is not well-known. The aim of this study was to evaluate the clinical and angiographic outcomes of coil embolization in patients with recurrent cerebral aneurysms after microsurgical clipping.

MATERIALS AND METHODS

Patients

Between May 1999 and February 2016, nineteen patients (5 men and 14 women) with 19 recurrent cerebral aneurysms who had been primarily treated by surgical clipping were treated with endovascular coiling and included in our study. All recurrent aneurysms in our study showed morphologic changes in the aneurysms that had been treated by clipping. Residual aneurysms that showed no morphologic change on follow-up angiography after clipping were excluded from our study. The medical records and radiographic studies of these patients were retrospectively reviewed to obtain clinical and radiographic information. In cases of loss to clinical follow-up, we obtained the clinical status by

Received July 17, 2018; accepted after revision October 10.

From the Departments of Neurosurgery (S.-T.K., J.H.P., Y.G.J.) and Diagnostic Radiology (J.W.B., H.W.J.), Busan Paik Hospital, Inje University, College of Medicine, Busan, Republic of Korea; and Department of Neurosurgery (S.-C.J., J.S.K., H.Y.K.), Haeundae Paik Hospital, Inje University, College of Medicine, Busan, Republic of Korea.

Sung-Tae Kim and Jin Wook Baek contributed equally to this work.

This work was supported by a 2016 Inje University Busan Paik Hospital research grant.

Please address correspondence to Sung-Chul Jin, MD, Department of Neurosurgery, Inje University Haeundae Paik Hospital, 875, Haeun-daero, Haeundae-gu, Busan, 612-896, Republic of Korea; e-mail: kusmal@hanmail.com

Indicates article with supplemental on-line table.

http://dx.doi.org/10.3174/ajnr.A5909

telephone. This retrospective study was approved by the institutional review board (Inje University Haeundae Paik Hospital).

Clinical outcomes were measured before and after coiling, at discharge, and at the end point of clinical follow-up using the modified Rankin Scale score. Each patient's clinical status at the final clinical follow-up evaluation was recorded as the follow-up clinical outcome. The angiographic outcome was measured by the Raymond-Roy occlusion classification.[6,7] In cases with a dome-to-neck ratio of <2 or an aneurysm neck of >4 mm, the aneurysm was defined as a wide-neck aneurysm.

Recurrence of the aneurysms after coiling was classified as either regrowth of the aneurysms or coil compaction. "Regrowth" of aneurysms was defined as morphologic changes in the aneurysms in comparison with the initial angiographic morphology of the aneurysms. "Coil compaction" of the aneurysms was defined as changes in the coil configuration in comparison with the initial appearance of the coil mass after coil embolization without morphologic changes in the aneurysm. The classification of regrowth or coil compaction of the aneurysms was determined by 2 neuroradiologists (H.W.J. and J.W.B.).

Follow-up imaging studies generally consisted of conventional angiography because CTA or MRA had poor image resolution due to artifacts produced by the clips and coils. Angiography was generally performed annually after coil embolization, and the results were classified into 3 categories: stable or improved occlusion (defined as no change or a decrease of the residual aneurysm), minor recurrence (defined as regrowth or coil compaction of the aneurysmal neck portion), and major recurrence (defined as regrowth or coil compaction of the aneurysmal sac that required retreatment).

Endovascular Strategies

Coil embolization was performed with the patient under general anesthesia. A biplane angiographic unit was used. In cases of rupture, oral antiplatelet agents were not used before coil embolization, and intravenous heparin (a bolus of 3000 IU) was administered after the aneurysm was secured with the coil. In unruptured cerebral aneurysms, dual oral antiplatelet agents (75 mg of clopidogrel and 100 mg of aspirin) were used for 5 days before coil embolization or a loading dose was administered (300 mg of clopidogrel and 300 mg of aspirin). After attaining access to the femoral artery, a bolus of 3000 IU of heparin was administered intravenously at the beginning of the procedure. An additional 1000-IU bolus of heparin was administered hourly to maintain an activated clotting time of >250 seconds. If an antiplatelet drug was necessary, aspirin (100 mg) or dual-antiplatelet agents (100 mg aspirin and 75 mg clopidogrel) were administered for 3–6 months after coil embolization.

Statistical Analysis

The statistical analysis was performed using SPSS for Windows (Version 24; IBM, Armonk, New York). For the statistical analysis, the clinical outcome was dichotomized into good (mRS = 0–2) and poor outcomes (mRS = 3–6). The factors related to major recurrence after coil embolization were evaluated. Univariate analysis was performed with age (60 years or older versus younger than 60 years), sex, initial presentation, presentation at coil embolization, multiple aneurysms, aneurysm size, wide neck, procedural complications, rebleeding after coil embolization, stent-assisted embolization, clinical outcome, and immediate radiologic outcome as factors using the Fisher exact test due to the expected frequency. Variables with a value of $P < .2$ were included in the multivariate logistic regression analysis. Statistical significance was set at $P < .05$ for a 95% confidence interval.

RESULTS

The demographic characteristics, characteristics of aneurysms, clinical factors before coil embolization, and factors related to clinical and angiographic outcome are summarized in the On-line Table. The mean age of the patients included in our study was 61.4 ± 9.3 years, with a range of 45–76 years. Seventeen patients initially presented with subarachnoid hemorrhage. All 19 patients had primarily undergone surgical clipping. At coil embolization, 9 patients (47.4%) presented with SAH due to rerupture of an aneurysm that had been treated by surgical clipping, and 10 patients presented with unruptured aneurysms, which were detected as recurrence of the aneurysms. The locations of the aneurysms included the posterior communicating artery (PcomA, n = 10), middle cerebral artery (n = 5), anterior communicating artery (n = 2), basilar tip (n = 1), and anterior choroidal artery (n = 1). The mean size of aneurysms was 6.5 ± 2.7 mm, ranging from 3.3 to 12 mm. The mean neck size was 4.9 ± 2.2 mm, ranging from 2.4 to 10 mm. Seventeen aneurysms had a wide neck, and 9 patients had multiple aneurysms. The mean interval between clipping and coil embolization was 143.5 ± 66.1 months, ranging from 43 to 276 months. Coil embolization was performed using single or double catheters in 16 patients. A balloon (n = 1) and stents (n = 2) were used to assist coil embolization in 3 patients. The 2 stent-assisted coil embolization aneurysms were located in the anterior communicating artery and PcomA, and the stents were Enterprise self-expanding stents (Codman & Shurtleff, Raynham, Massachusetts).

A control angiogram obtained after coil embolization showed complete occlusion of the aneurysms in 10 patients, a residual neck in 8 patients, and a residual sac in 1 patient. Procedure-related thromboembolic complications occurred in 2 patients, 1 of which resulted in permanent right hemiparesis (grade 3).

Angiographic follow-up was available for 12 patients (63.2%). Among the other 7 patients who did not undergo cerebral angiography, 2 patients were in poor condition (mRS ≥ 4), and the other 5 patients refused the procedure. Major recurrence was detected in 5 patients, resulting in a major recurrence rate of 41.7% (5/12).

At discharge, 9 patients had mRS scores of 0, two patients had mRS scores of 1, five patients had mRS scores of 2, two patients had mRS scores of 4, and 1 patient had an mRS score of 5. The mean clinical follow-up after coil embolization was 58.3 ± 38.8 months and ranged from 3 to 181 months. The follow-up mRS scores were 0 in 10 patients, 1 in 1 patient, 2 in 3 patients, 3 in 1 patient, 4 in 2 patients, and 6 in 2 patients. Rebleeding developed in 2 patients at 18 and 180 months after coil embolization. All patients who had poor follow-up clinical outcomes (mRS ≥ 3) presented with SAH at coil embolization. Worsening of the mRS compared with the score at discharge after coil embolization occurred in 3 patients. One patient who had an aggravated clinical

outcome after coil embolization was diagnosed with vascular dementia. Another patient exhibited rebleeding that caused clinical aggravation at 180 months after coil embolization. Another patient died due to sepsis caused by hospital-acquired pneumonia at the rehabilitation hospital.

Statistically significant results between major recurrence and the variables could not be achieved because of the small sample size. However, younger age ($P = .293$) and SAH presentation at coil embolization ($P = .242$) seemed to be somewhat related to major recurrence (Table 1).

All 5 cerebral aneurysms that recurred after coil embolization showed aneurysm regrowth (Figs 1–5). All patients had initially presented with SAH at microsurgical clipping. Thereafter, 4 patients (80%) presented with SAH due to aneurysm rerupture at coil embolization, and 1 patient underwent coil embolization due to regrowth of the aneurysm on serial follow-up angiography. Despite the risk of aneurysm regrowth, our endovascular strategy for ruptured cerebral aneurysms is considered a simple technique without flow diversion effect because of the possibility of procedural thromboembolism caused by insufficient antiplatelet medication. Additionally, we favor a simple technique as our endovascular strategy for unruptured cerebral aneurysms because stent-assisted coiling or flow diverters have shown insufficient evidence of procedural safety thus far. Therefore, we decided to retreat all patients by additional coil embolization using a simple technique without a stent (Table 2).

Table 1: Factors related to major recurrence in univariate analysis

Variables	Categories	Major Recurrence	*P* Value[a]
Age	Younger than 60 yr ($n = 7$)	4 (57.1%)	.293
	60 yr or older ($n = 5$)	1 (20.0%)	
Sex	Male ($n = 4$)	2 (50.0%)	1.0
	Female ($n = 8$)	3 (37.5%)	
Presentation at surgery	Ruptured ($n = 12$)	5 (41.7%)	1.0
	Unruptured ($n = 0$)	0 (0%)	
Presentation at coiling	Ruptured ($n = 6$)	4 (66.7%)	.242
	Unruptured ($n = 6$)	1 (16.7%)	
Multiple aneurysms	Yes ($n = 6$)	2 (33.3%)	1.0
	No ($n = 6$)	3 (50.0%)	
Size	≤10 mm ($n = 11$)	4 (36.4%)	.417
	>10 mm ($n = 1$)	1 (100%)	
Wide neck	Narrow ($n = 2$)	0 (0%)	.470
	Wide ($n = 10$)	5 (50.0%)	
Stent use	Yes ($n = 1$)	0 (0%)	1.0
	No ($n = 11$)	5 (45.5%)	
Procedural complications	Yes ($n = 1$)	1 (100%)	.417
	No ($n = 11$)	4 (36.4%)	
Clinical outcome (good, mRS ≤ 2)	Good ($n = 9$)	3 (33.3%)	.523
	Bad ($n = 3$)	2 (66.7%)	
Rebleeding after coiling	Yes ($n = 2$)	2 (100%)	.152
	No ($n = 10$)	3 (30.0%)	
Immediate radiologic results	Complete ($n = 8$)	3 (20%)	1.0
	Remnant neck ($n = 4$)	2 (37.5%)	
	Remnant sac ($n = 0$)	–	

Note:—mRS indicates modified Rankin scale.
[a] All *P* values are significant.

One patient, who had a PcomA aneurysm (patient 1), underwent additional coil embolization twice due to regrowth of the aneurysm at the 41st and 84th month after the primary coil embolization. Another patient who had a basilar tip aneurysm (patient 8) had a minor recurrence after additional coil embolization. Therefore, the patient was observed for 49 months after the primary coil embolization, and further follow-up seemed to be required. Two patients who experienced rebleeding events after coil embolization (patients 13 and 14) died or were bedridden after additional coil embolization. It took 180 months for rerupture of the middle cerebral artery aneurysm to occur after the primary coil embolization in 1 patient (patient 13) and 18 months for rerupture in another patient with a PcomA aneurysm (patient 14). The other patient who had a PcomA aneurysm (patient 16) underwent

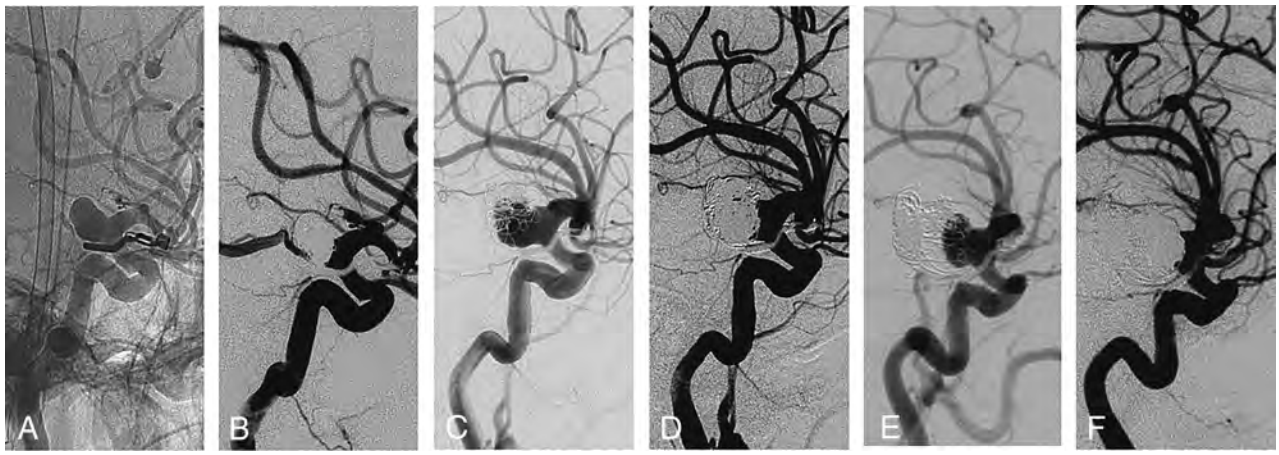


FIG 1. First case of our series, posterior communicating artery aneurysm. *A*, Working angle view angiography before the first coil embolization. *B*, After the first coil embolization, the final angiography evaluation reveals complete occlusion of the aneurysm. *C*, After 41 months, follow-up angiography reveals major recurrence of the aneurysm. *D*, After the second coil embolization, the residual neck is detected in the final angiography evaluation. *E*, After 43 months, follow-up angiography reveals major recurrence. *F*, Final image of the third coil embolization.

FIG 2. Eighth case of our series, basilar tip aneurysm. *A*, Working angle view angiography before first coil embolization. *B*, After the first coil embolization, the final angiography evaluation reveals complete occlusion of the aneurysm. *C*, After 29 months, follow-up angiography reveals major recurrence of the aneurysm. *D*, After the second coil embolization, the residual neck is detected in the final angiography evaluation. *E*, Follow-up angiography reveals minor recurrence.

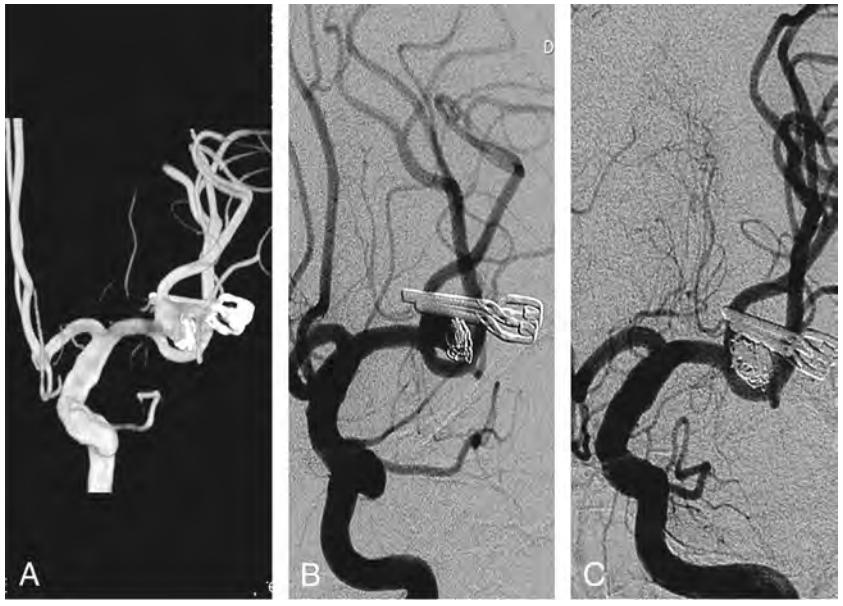


FIG 3. Thirteenth case in our series, left middle cerebral artery bifurcation aneurysm. The first coil embolization of this aneurysm was performed at another hospital. Therefore, we confirmed the presence of a residual neck in the medical record, without imaging. *A*, At 180 months after the first coil embolization, 3D rotational angiography reveals recurrence of a middle cerebral artery bifurcation aneurysm. *B*, Working angle angiography before the second coil embolization. *C*, The final angiography evaluation reveals complete occlusion of the aneurysm.

an additional coil embolization at 47 months after the primary coil embolization due to regrowth of the aneurysm.

DISCUSSION

Several reports on endovascular coiling for recurrent or residual cerebral aneurysms after microsurgical clipping have shown that the procedure is feasible.[1-4] However, because recurrent cerebral aneurysms after surgical clipping may have more aggressive clinical and radiologic features than residual cerebral aneurysms after surgical clipping, we evaluated the radiologic and clinical outcomes of recurrent cerebral aneurysms after surgical clipping while excluding residual cerebral aneurysms.

In our series, procedural morbidity occurred in 1 case of thromboembolism (5.3%), resulting in hemiparesis; no procedural mortality occurred. Because there are no prospective-controlled studies, a direct comparison of the technical feasibility of revised microsurgical clipping and coil embolization is difficult for recurrent cerebral aneurysms treated by surgical clipping. However, considering previous reports on this issue, revised microsurgical clipping in this condition seems to be a complicated and an eventually challenging process.[8,9] In addition, our procedural morbidity rate was not higher than that for the initial coil embolization of ruptured or unruptured cerebral aneurysms.[7,10] Accordingly, coil embolization may be a feasible treatment option for recurrent cerebral aneurysms treated primarily by surgical clipping.

In our series, the recurrence and rebleeding rates of coil embolization for the recurrent cerebral aneurysms after microsurgi-

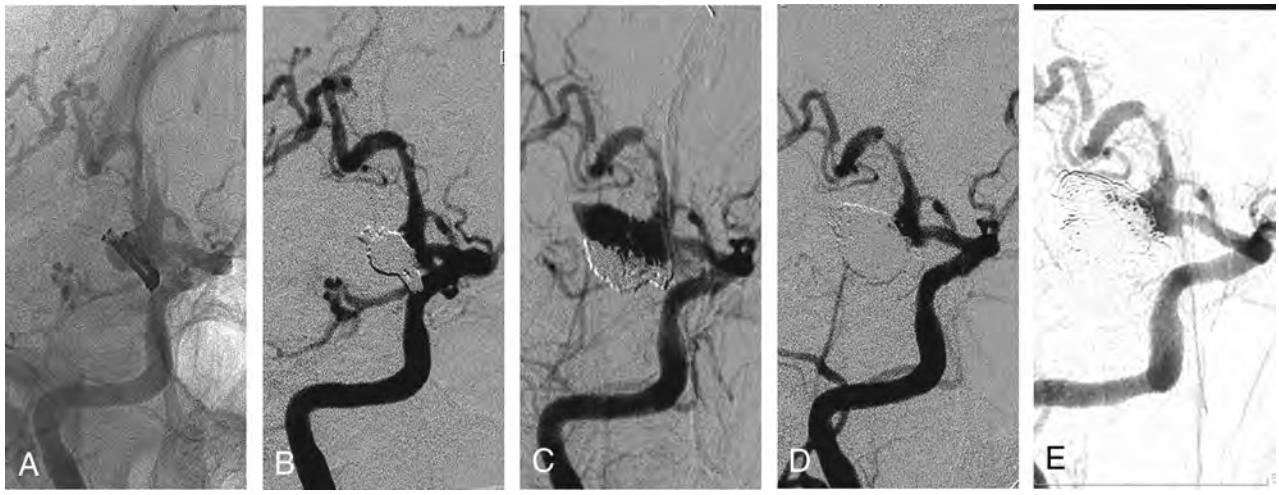


FIG 4. Fourteenth case of our series, posterior communicating artery aneurysm. *A*, Working angle view angiography before the first coil embolization. *B*, After the first coil embolization, the final angiography evaluation reveals the residual neck of the aneurysm. *C*, After 18 months, follow-up angiography reveals major recurrence of the aneurysm. *D*, After the second coil embolization, a residual neck is detected in the final angiography evaluation. *E*, Follow-up angiography reveals minor recurrence.

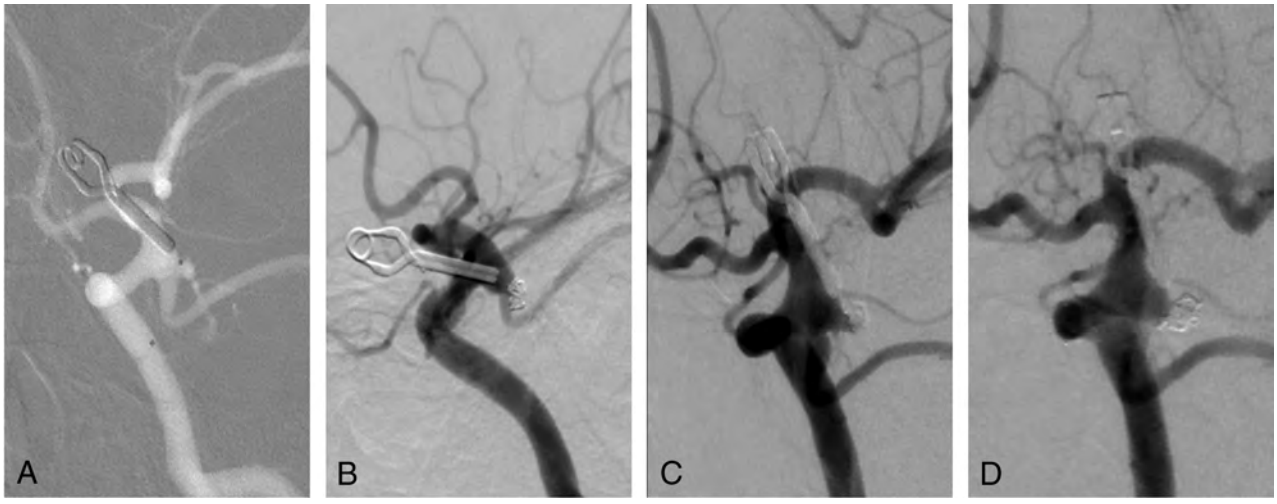


FIG 5. Sixteenth case of our series, posterior communicating artery aneurysm. *A*, Working angle view angiography before the first coil embolization. *B*, After the first coil embolization, the final angiographic evaluation reveals the residual neck of the aneurysm. *C*, After 47 months, follow-up angiography reveals major recurrence of the aneurysm. *D*, After the second coil embolization, the final angiography evaluation reveals complete occlusion of the aneurysm.

Table 2: Characteristics and treatment results of recurrent aneurysms after coil embolization

No.	Location	Presentation at Clipping	Presentation at Coiling	Interval to Retreatment	Size (H × W) (mm)	Neck Diameter (mm)	Immediate Radiologic Results	Reason for Recurrence	Follow-Up mRS
1	PcomA	R	R/-/-	41/43 mo	8.7 × 6.3	5.1	C/N/N	Regrowth	2
8	Basilar tip	R	-/-	29 mo	3.8 × 7.2	6	C/N	Regrowth	0
13	MCA	R	R/R	180 mo	12 × 10	10	N/C	Regrowth	6
14	PcomA	R	R/R	18 mo	7.8 × 8.1	8.1	N/N	Regrowth	4
16	PcomA	R	R/-	47 mo	3.3 × 3.0	2.4	N/C	Regrowth	0

Note:—C indicates complete occlusion; H, height; N, residual neck; R, ruptured; W, width.

cal clipping were 41.7% (5/12) and 16.6% (2/12), respectively. These rates are high compared with the general consensus and previous reports.[1-5] This finding may be because our series included only recurrent aneurysms after surgical clipping, unlike previous reports that examined recurrent and residual aneurysms after surgical clipping. Due to the high recurrence and rebleeding rates, strict follow-up with conventional angiography should be mandatory after coil embolization of recurrent cerebral aneurysms treated by surgical clipping.

Aneurysm regrowth is thought to be an aggressive angiographic characteristic that may be more susceptible to recurrence after repeat endovascular coiling than after coil compaction.[11,12] In our series, all the aneurysms that recurred after coil embolization showed a tendency toward regrowth rather than coil com-

paction. Furthermore, the recurred aneurysms after coil embolization showed repeat recurrence of up to 3 times, and some aneurysms are currently still growing. These aneurysms seem to have more aggressive characteristics than the usual cerebral aneurysms. Accordingly, simple coil embolization without flow diversion in most of our patients ($n = 17$, 89.5%) showed limited treatment potential in terms of durability for the aneurysms that showed regrowth because this method may not be able to recover the pathologically altered part of the vessel wall adjacent to the aneurysm, potentially leading to regrowth of the aneurysm.

Accordingly, it would be better to consider aggressive treatment options for recurrent cerebral aneurysms after surgical clipping unless the treatment option might increase the patient's risk considerably. Providing flow-diversion effects such as stent-assisted coiling or a flow diverter rather than simple coiling might improve the durability of the treatment for recurrent cerebral aneurysms that were initially treated by surgical clipping. In fact, we observed favorable outcomes using a flow diverter for recurrent cerebral aneurysms after surgical clipping or endovascular coiling.[13,14] Furthermore, the stents intended for stent-assisted coiling but not for flow diversion also have a certain flow-diversion effect.[15-17] However, those results were based on relatively short-term results or observed in an experimental environment such as with computational fluid dynamics simulation. Accordingly, we should consider the technical feasibility of applying flow diversion in each patient individually when treating this kind of lesion.

This report is a retrospective study of data from patients who underwent coil embolization at 2 centers and does not include surgical results. The study also has a small sample size and a relatively low rate of follow-up angiography after coil embolization. Therefore, the inability to obtain statistically significant results for all data analyzed is a limitation of this study.

CONCLUSIONS

In our series, coil embolization of recurrent aneurysms after surgical clipping was feasible but resulted in a high recurrence rate. Aneurysm regrowth rather than coil compaction was a major factor in the recurrence of cerebral aneurysms in this series. Strict imaging follow-up is required after coil embolization of these lesions due to the high recurrence and rebleeding risks.

REFERENCES

1. Chung J, Park IS, Park H, et al. **Endovascular coil embolization after clipping: endovascular treatment of incompletely clipped or recurred cerebral aneurysms.** *J Cerebrovasc Endovasc Neurosurg* 2014;16:262–67 CrossRef Medline
2. Kim BM, Kim DJ, Kim DI, et al. **Clinical presentation and outcomes of coil embolization of remnant or recurred intracranial aneurysm after clipping.** *Neurosurgery* 2010;66:1128–33; discussion 1133 CrossRef Medline
3. Kim ST, Jeong HW, Jeong YG, et al. **Coiling as retreatment in intracranial aneurysm of de novo formation or regrowth: case report.** *Neurointervention* 2013;8:46–51 CrossRef Medline
4. Rabinstein AA, Nichols DA. **Endovascular coil embolization of cerebral aneurysm remnants after incomplete surgical obliteration.** *Stroke* 2002;33:1809–15 CrossRef Medline
5. Schaafsma JD, Sprengers ME, van Rooij WJ, et al. **Long-term recurrent subarachnoid hemorrhage after adequate coiling versus clipping of ruptured intracranial aneurysms.** *Stroke* 2009;40:1758–63 CrossRef Medline
6. Raymond J, Guilbert F, Weill A, et al. **Long-term angiographic recurrences after selective endovascular treatment of aneurysms with detachable coils.** *Stroke* 2003;34:1398–403 CrossRef Medline
7. Roy D, Milot G, Raymond J. **Endovascular treatment of unruptured aneurysms.** *Stroke* 2001;32:1998–2004 CrossRef Medline
8. Hokari M, Kazumara K, Nakayama N, et al. **Treatment of recurrent intracranial aneurysms after clipping: a report of 23 cases and a review of the literature.** *World Neurosurg* 2016;92:434–44 CrossRef Medline
9. Kobayashi S, Moroi J, Hikichi K, et al. **Treatment of recurrent intracranial aneurysms after neck clipping: novel classification scheme and management strategies.** *Oper Neurosurg (Hagerstown)* 2017;13:670–78 CrossRef Medline
10. Park HK, Horowitz M, Jungreis C, et al. **Periprocedural morbidity and mortality associated with endovascular treatment of intracranial aneurysms.** *AJNR Am J Neuroradiol* 2005;26:506–14 Medline
11. Hasan DM, Nadareyshvili AI, Hoppe AL, et al. **Cerebral aneurysm sac growth as the etiology of recurrence after successful coil embolization.** *Stroke* 2012;43:866–68 CrossRef Medline
12. Hoppe AL, Raghavan ML, Hasan DM. **Comparison of the association of sac growth and coil compaction with recurrence in coil embolized cerebral aneurysms.** *PLoS One* 2015;10:e0123017 CrossRef Medline
13. Ding D, Starke RM, Evans AJ, et al. **Endovascular treatment of recurrent intracranial aneurysms following previous microsurgical clipping with the Pipeline Embolization Device.** *J Clin Neurosci* 2014;21:1241–44 CrossRef Medline
14. Dornbos D, Karras CL, Wenger N, et al. **Pipeline embolization device for recurrence of previously treated aneurysms.** *Neurosurg Focus* 2017;42:E8 CrossRef Medline
15. Wang C, Tian Z, Liu J, et al. **Flow diverter effect of LVIS stent on cerebral aneurysm hemodynamics: a comparison with Enterprise stents and the Pipeline device.** *J Transl Med* 2016;14:199 CrossRef Medline
16. Choi HH, Ha EJ, Lee JJ, et al. **Comparison of clinical outcomes of intracranial aneurysms: procedural rupture versus spontaneous rupture.** *AJNR Am J Neuroradiol* 2017;38:2126–30 CrossRef Medline
17. Tremmel M, Xiang J, Natarajan SK, et al. **Alteration of intra-aneurysmal hemodynamics for flow diversion using Enterprise and Vision stents.** *World Neurosurg* 2010;74:306–15 CrossRef Medline

Y-Stent-Assisted Coiling of Wide-Neck Bifurcation Intracranial Aneurysms: A Meta-Analysis

F. Cagnazzo, N. Limbucci, S. Nappini, L. Renieri, A. Rosi, A. Laiso, D. Tiziano di Carlo, P. Perrini, and S. Mangiafico

ABSTRACT

BACKGROUND: Y-stent-assisted coiling for wide-neck intracranial aneurysms required further investigation.

PURPOSE: Our aim was to analyze outcomes after Y-stent placement in wide-neck aneurysms.

DATA SOURCES: We performed a systematic search of 3 data bases for studies published from 2000 to 2018.

STUDY SELECTION: According to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines, we included studies reporting Y-stent-assisted coiling of wide-neck aneurysms.

DATA ANALYSIS: Random-effects meta-analysis was used to pool the following: aneurysm occlusion rate, complications, and factors influencing the studied outcomes.

DATA SYNTHESIS: We included 27 studies and 750 aneurysms treated with Y-stent placement. The immediate complete/near-complete occlusion rate was 82.2% (352/468; 95% CI, 71.4%–93%; $I^2 = 92\%$), whereas the long-term complete/near-complete occlusion rate was 95.4% (564/598; 95% CI, 93.7%–97%; $I^2 = 0\%$) (mean radiologic follow-up of 14 months). The aneurysm recanalization rate was 3% (20/496; 95% CI, 1.5%–4.5%; $I^2 = 0\%$), and half of the recanalized aneurysms required retreatment. The treatment-related complication rate was 8.9% (63/614; 95% CI, 5.8%–12.1%; $I^2 = 44\%$). Morbidity and mortality after treatment were 2.4% (18/540; 95% CI, 1.2%–3.7%; $I^2 = 0\%$) and 1.1% (5/668; 95% CI, 0.3%–1.9%; $I^2 = 0\%$), respectively. Crossing Y-stent placement was associated with a slightly lower complication rate compared with the kissing configuration (56/572 = 8.4%; 95% CI, 5%–11%; $I^2 = 46\%$ versus 4/30 = 12.7%; 95% CI, 3%–24%; $I^2 = 0\%$). Occlusion rates were quite comparable among Enterprise, Neuroform, and LVIS stents, whereas the Enterprise stent was associated with lower rates of complications (8/89 = 6.5%; 95% CI, 1.6%–11%; $I^2 = 0\%$) compared with the others (20/131 = 14%; 95% CI, 5%–26%; $I^2 = 69\%$ and 9/64 = 11%; 95% CI, 3%–20%; $I^2 = 18\%$).

LIMITATIONS: This was a small, retrospective series.

CONCLUSIONS: Y-stent-assisted coiling yields high rates of long-term angiographic occlusion, with a relatively low rate of treatment-related complications. Y-stent placement with a crossing configuration appears to be associated with better outcomes. Although Y-configuration can be obtained using many types of stents with comparable occlusion rates, the Enterprise stent is associated with lower complication rates.

ABBREVIATIONS: IQR = interquartile range; PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses; SAC = stent-assisted coiling; Y-SAC = Y-stent-assisted coiling

With the improvement of angiographic imaging, operator experience, and widespread use of more complex techniques, an increased number of intracranial aneurysms can be effectively treated with endovascular techniques. However, aneurysms with wide necks, unfavorable anatomic configurations, and partial incorporation of bifurcation branches are still challenging lesions for endovascular treatment, and each technique presents specific limitations. Balloon-assisted coiling may not be suitable in case of very wide-neck aneurysms with complex anatomy.[1]

Received August 28, 2018; accepted after revision October 15.

From the Interventional Neuroradiology Unit (F.C., N.L., S.N., L.R., A.R., A.L., S.M.), Careggi University Hospital, Florence, Italy; and Department of Neurosurgery (D.T.d.C., P.P.), Cisanello Hospital, University of Pisa, Pisa, Italy.

Please address correspondence to Federico Cagnazzo, MD, Interventional Neuroradiology Unit, AOU Careggi, CTO, Largo P Palagi 1, Florence 50134, Italy; e-mail: f.cagnazzo86@gmail.com

Indicates article with supplemental on-line tables.

Indicates article with supplemental on-line photos.

http://dx.doi.org/10.3174/ajnr.A5900

Flow diversion appears to be a straightforward treatment, but the fate of the covered side branches in case of bifurcation aneurysms may be a concern.[2,3] Intrasaccular flow disruption seems to be an effective tool for wide-neck aneurysms, and treatment-related outcomes continue to improve due to better case selection, technology, and technical improvement, despite long-term angiographic occlusion rates that are still lower than those of stent-assisted coiling (SAC).[4] Finally, PulseRider (Cerenovus, New Brunswick, New Jersey) and pCONus aneurysm implant (phenox, Bochum, Germany) devices appear to be promising new strategies to treat wide-neck lesions, though long-term studies are needed to confirm the efficacy of these devices.[5,6]

Y-stent-assisted coiling (Y-SAC), first proposed by Chow et al,[7] in 2004, is performed by placing 2 stents from the parent artery into each of the bifurcation vessels, creating a new bifurcation point that provides a mechanical scaffold, preventing coil protrusion.[8] To date, the feasibility of this procedure has been confirmed by small retrospective series, but there has not been sufficient evaluation of the safety and efficacy of this technique. Improved understanding of treatment-related outcomes after Y-SAC can help practitioners in the selection of lesions that can be effectively treated with this strategy. Our meta-analysis examined occlusion rates and procedure-related complications, focusing on the influence of aneurysm and patient features and treatment characteristics on the studied outcomes.

MATERIALS AND METHODS

Literature Search

A comprehensive literature search of PubMed, Ovid MEDLINE, and Ovid EMBASE was conducted for studies published from January 2000 to August 2018. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)[9] guidelines were followed. The key words and detailed search strategy are reported in On-line Table 1, and the studies included in our review are reported in On-line Table 2. The inclusion criteria were the following: studies reporting series with >5 patients with bifurcation intracranial aneurysms treated with Y-SAC. Exclusion criteria were the following: 1) case reports, 2) review articles, 3) studies published in languages other than English, 4) in vitro/animal studies, and 5) series reporting aneurysms treated with X-stent placement. In cases of overlapping patient populations, only the series with the largest number of patients or the most detailed data were included. Two independent readers screened articles in their entirety to determine eligibility for inclusion. A third author solved discrepancies.

Data Collection

We extracted the following data: 1) technical success rate, 2) occlusion rate, 3) treatment-related complications, and 4) clinical outcome. Occlusion and complication rates were analyzed on the basis of the influence of the following: 1) aneurysm location; 2) ruptured-versus-unruptured aneurysms; 3) aneurysm size: small saccular (<5 mm) and medium-sized aneurysms (between 5.0 and 9.9 mm) versus large (between 10 and 20 mm) and very large and giant aneurysms (>20 mm); 4) patient age (younger than 60 years versus older than 60 years); 5) crossing-versus-kissing Y-configuration; 6) type of stent used; and 7) hybrid-versus-nonhybrid configuration. In the hybrid configuration, an open-cell stent (Neuroform; Stryker Neurovascular, Kalamazoo, Michigan) was used with a closed-cell stent (usually Enterprise self-expanding stent; Codman & Shurtleff, Raynham, Massachusetts).

Complete/near-complete aneurysm occlusion was defined on the basis of the following: Raymond-Roy classification[10] (class 1–2), or when "complete occlusion" and "neck remnant" were used in the study. Treatment-related complications were divided into the following: 1) periprocedural/early events (within 30 days) and delayed events (after 30 days); 2) transient (asymptomatic events or complete neurologic recovery) and permanent complications (symptomatic events with permanent deficits); and 3) ischemic and hemorrhagic complications. Finally, good outcome was defined as a modified Rankin Scale score of 0–2 or a Glasgow Outcome Score of 4–5, or it was assumed if the study used terms "no morbidity," "good recovery," or "no symptoms."

Outcomes

The primary objectives of this study were to define the safety (treatment-related complications, mortality rate, and neurologic outcomes) and efficacy (technical success rate, immediate and long-term occlusion) of Y-SAC. The secondary objectives were to define the influence of aneurysm, patient, and treatment characteristics on the analyzed outcomes.

Quality Scoring

The Newcastle-Ottawa Scale[11] was used for the quality assessment of the included studies (details in On-line Table 3). The quality assessment was performed by 2 authors independently, and a third author solved discrepancies.

Statistical Analysis

We estimated, from each cohort, the cumulative prevalence (percentage) and 95% confidence interval for each outcome. Heterogeneity of the data was assessed by the Higgins index (I^2), and subsequently, the DerSimonian and Laird random-effects model was applied. The graphic representation was shown by a forest plot. To evaluate the heterogeneity and bias, the meta-regression and funnel plot followed by the Egger linear regression test were analyzed, respectively. To compare the percentages and to calculate the *P* values, a *Z*-test for 2 proportions was used. Differences were considered significant at $P < .05$. Meta-analysis was performed with ProMeta-2 (Internovi; Cesena, Italy) and OpenMeta[Analyst] (http://www.cebm.brown.edu/openmeta/).

RESULTS

Literature Review

Studies included in our meta-analysis are summarized in On-line Table 2. The search flow diagram is shown in On-line Fig 1.

Twenty-seven studies and 750 aneurysms treated with Y-SAC were included in our review.

Quality of Studies

Studies included in our review were retrospective series: Two studies were retrospective multicentric series, whereas 25 articles were single-center retrospective series. Details of the rating of the included studies are reported in On-line Table 3.

Treatment-related outcomes after Y-stent-assisted coiling of bifurcation aneurysms

Variables	Results of Systematic Review and Meta-Analysis	No. of Articles	Statistic (95% CI) (I^2)
Angiographic outcomes			
Rate of successful stent deployment	550/569 = 97%	20	(96–99) (I^2 = 0%)
Immediate aneurysm occlusion rate (RR 1–2)	352/468 = 82.2%	15	(71.4–93) (I^2 = 92%)
Long-term aneurysm occlusion rate (RR 1–2)	564/598 = 95.4%	24	(93.7–97) (I^2 = 0%)
Aneurysm recanalization	20/496 = 3%	19	(1.5–4.5) (I^2 = 0%)
Recanalization among anterior circulation	2/89 = 2%	9	(0.9–9) (I^2 = 0%)
Recanalization among posterior circulation	11/130 = 4%	12	(2–9) (I^2 = 0%)
Mean size of recanalized vs nonrecanalized aneurysms (mm)	11 ± 2.4 vs 7 ± 1.18	7	
Rate of retreatment of recanalized aneurysms	10/20 = 50%	19	(29–70)
Treatment-related complications and clinical outcomes			
Overall treatment-related complications	63/614 = 8.9%	23	(5.8–12.1) (I^2 = 44%)
Periprocedural/early complications (within 30 days)	46/574 = 6.7%	21	(4–9) (I^2 = 35%)
Delayed complications (after 30 days)	9/574 = 2.1%	21	(1–3) (I^2 = 0%)
Transient complications	33/540 = 6.5%	20	(2.2–6) (I^2 = 21%)
Symptomatic complications without permanent deficits	22/540 = 3%	20	(1.3–4) (I^2 = 35%)
Permanent complications	18/540 = 2.4%	20	(1.2–3.7) (I^2 = 0%)
Treatment-related mortality	5/668 = 1.1%	23	(0.3–1.9) (I^2 = 0%)
Overall rate of good neurologic outcome	291/317 = 94.5%	15	(92–97) (I^2 = 7%)
Type of complications			
Thromboembolic complications	44/594 = 6.5%	22	(3–7.6) (I^2 = 29%)
Hemorrhagic complications	11/594 = 2%	22	(0.7–3) (I^2 = 0%)
Acute in-stent thrombosis	20/577 = 2.1%	22	(1.6–6) (I^2 = 0%)
Aneurysm rupture after treatment	0/650 = 0%	22	
Chronic in-stent stenosis	10/295 = 2.3%	11	(0.6–4) (I^2 = 0%)

Note:—RR indicates Raymond-Roy grade.

Patient Population and Aneurysm Characteristics

Overall, 744 patients and 750 aneurysms were treated with Y-SAC (On-line Table 4). The mean age of patients was 56.6 years (range, 22–80 years), and the proportion of male patients was 36% (95% CI, 32%–40%). Aneurysms included in our study were wide-neck saccular bifurcation aneurysms. Overall, 61% (366/601; 95% CI, 56%–64%) were anterior circulation aneurysms, whereas 39% (235/601; 95% CI, 35%–43%) were posterior circulation lesions. Most of the aneurysms were located at the basilar tip (235/601 = 39%; 95% CI, 35%–43%), middle cerebral artery bifurcation (206/601 = 34.4%; 95% CI, 30%–38%), and anterior communicating artery (120/601 = 20%; 95% CI, 16%–23%). Mean aneurysm size was 9.6 mm (median, 9.9 mm; interquartile range [IQR] = 8–10.5 mm; range, 3–25 mm). The proportion of acutely ruptured aneurysms treated with Y-stent placement was 11% (66/592; 95% CI, 8.5%–14%).

Treatment Characteristics

The most common stent used was the Enterprise (476/1060 = 45%; 95% CI, 42%–47%), followed by the Neuroform (332/1060 = 31.3%; 95% CI, 28%–40%), LVIS stents (MicroVention, Tustin, California) (132/1060 = 12.5%; 95% CI, 10%–14%), Solitaire (Covidien, Irvine, California) (66/1060 = 6.2%; 95% CI, 4.9%–7.8%), and Acclino flex Stent (Acandis, Pforzheim, Germany) (54/1060 = 5%; 95% CI, 3.9%–6.6%). Nonhybrid (closed/closed-cell and open/open-cell stents) and hybrid (closed/open-cell stents) techniques were used in 87.5% (512/585; 95% CI, 84%–98%) and 12.5% (73/585; 95% CI, 10%–15%) of aneurysms, respectively. Y-stent placement with a crossing configuration was the most common technique (688/750 = 92%; 95% CI, 89%–93%), followed by kissing stent placement (60/750 = 8%; 95% CI, 6%–10%). The mean radiologic (digital subtraction angiography) follow-up was 14 months (range, 6–24 months; median, 12 months; IQR = 10–18 months), and the mean clinical follow-up was 17 months (range, 3–30 months; median, 16.5 months; IQR = 7.8–21 months).

Angiographic Outcomes

The technical success rate was 97% (550/569; 95% CI, 96%–99%; I^2 = 0%) (Table). Immediate angiographic occlusion after treatment was obtained in 82.2% (352/468; 95% CI, 71.4%–93%; I^2 = 92%) of aneurysms. The rate of long-term complete/near-complete occlusion was 95.4% (564/598; 95% CI, 93.7%–97%; I^2 = 0%). Meta-regression showed a nonsignificant variation of the effect size (P = .056), and the funnel plot, followed by the Egger linear regression test, excluded publication bias (P = .358) (On-line Fig 2).

Aneurysm recanalization occurred in 3% (20/496; 95% CI, 1.5%–4.5%; I^2 = 0%) of cases, and half of the recanalized aneurysms required retreatment. The mean size of recanalized and nonrecanalized aneurysms was 11 ± 2.4 versus 7 ± 1.18 mm (P = .0001). The aneurysm recanalization rate was 2% (2/89; 95% CI, 0.9%–9%; I^2 = 0%) and 4% (11/130; 95% CI, 2%–9%; I^2 = 0%) among the anterior and posterior circulation, respectively (P = .4).

Treatment-Related Complications

The overall complication rate was 8.9% (63/614; 95% CI, 5.8%–12.1%; I^2 = 44%) (Table). Meta-regression showed a nonsignificant variation of the effect size (P = .443), and the funnel plot, followed by the Egger linear regression test, excluded publication bias (P = .574) (On-line Fig 3). Periprocedural/early complications were 6.7% (46/574; 95% CI, 4%–9%; I^2 = 35%). Delayed complications were 2.1% (9/574; 95% CI, 1%–3%; I^2 = 0%) (3 cases of in-stent occlusion, 5 ischemic events, 1 thromboembolism after discontinuation of the antiplatelet therapy). Transient

(with and without related symptoms) and permanent complications were 6.5% (33/540; 95% CI, 2.2%–6%; $I^2 = 21\%$) and 2.4% (18/540; 95% CI, 1.2%–3.7%; $I^2 = 0\%$), respectively. The rate of symptomatic complications without permanent deficits (transient events with related symptoms) was 3% (22/540; 95% CI, 1.3%–4%; $I^2 = 35\%$).

Overall, ischemic/thromboembolic, and hemorrhagic events were 6.5% (44/594; 95% CI, 3.7%–6%; $I^2 = 29\%$) and 2% (11/594; 95% CI, 0.7%–3%; $I^2 = 0\%$), respectively. The rate of acute in-stent thrombosis was 2.1% (20/577; 95% CI, 1.6%–6%; $I^2 = 0\%$) and chronic in-stent stenosis (>50%) was 2.3% (10/295; 95% CI, 0.6%–4%; $I^2 = 0\%$). There were no cases of aneurysm rupture after treatment during follow-up.

Treatment-related mortality was 1.1% (5/668; 95% CI, 0.3%–1.9%; $I^2 = 0\%$), and the rate of good neurologic outcome was 94.5% (291/317; 95% CI, 92%–97%; $I^2 = 7\%$).

Factors Related to Aneurysm Occlusion

Overall, the occlusion rate was comparable among ruptured-versus-unruptured aneurysms ($P = .2$), anterior-versus-posterior circulation ($P = .27$), patients younger-versus-older than 60 years ($P = .5$), crossing-versus-kissing techniques ($P = .2$), and hybrid-versus-nonhybrid techniques ($P = .2$). Long-term occlusion was significantly higher among small and medium-sized versus large or very large/giant aneurysms (98/107 = 92%; 95% CI, 86%–96%; $I^2 = 0\%$ versus 33/43 = 79%; 95% CI, 67%–91%; $I^2 = 25\%$) ($P = .02$) and among aneurysms treated with Y-stent placement used as the first treatment versus retreatment (119/128 = 94%; 95% CI, 89%–98%; $I^2 = 0\%$ versus 40/48 = 82%; 95% CI, 71%–92%; $I^2 = 0\%$) ($P = .01$) (On-line Table 5). MCA aneurysms had higher occlusion rates (52/54 = 94.9%; 95% CI, 89%–97%; $I^2 = 0\%$) compared with basilar tip aneurysms (126/145 = 88.8%; 95% CI, 83%–93%; $I^2 = 0\%$) and anterior cerebral artery location (42/46 = 87%; 95% CI, 78%–96%; $I^2 = 0\%$) ($P > .05$).

Factors Related to Complications after Treatment

The complication rate was higher for ruptured (11/52 = 18%; 95% CI, 7%–28%; $I^2 = 21\%$) compared with unruptured aneurysms (29/236 = 8%; 95% CI, 4.5%–13%; $I^2 = 36\%$) ($P = .02$). There was no statistically significant difference in complication rates in relation to aneurysm location, patient age, aneurysm size, and technical factors (On-line Table 5).

Relationship between Type of Stent Used and Treatment-Related Outcomes

Treatment with the Enterprise stent was associated with 96% (98/101; 95% CI, 93%–98%; $I^2 = 0\%$) long-term occlusion and 6.5% (8/99; 95% CI, 1.6%–11%; $I^2 = 0\%$) complications. Aneurysm occlusion during follow-up and treatment-related complications after Neuroform stent placement were 94% (119/130; 95% CI, 90%–98%; $I^2 = 14\%$) and 14% (20/131; 95% CI, 5%–26%; $I^2 = 69\%$), respectively. Finally, treatment with the LVIS and LVIS Jr stents yielded 92% (52/56; 95% CI, 85%–99%; $I^2 = 0\%$) and 11% (9/64; 95% CI, 3%–20%; $I^2 = 18\%$) long-term occlusion and complications rates, respectively.

Study Heterogeneity

Substantial heterogeneity was reported in the following outcomes: technical success rate, overall rate of treatment-related complications, occlusion rate among patients younger than 60 years, and complications after treatment with the Neuroform stent.

DISCUSSION

Angiographic Outcomes

Our meta-analysis stressed several important findings related to the treatment of complex wide-neck bifurcation aneurysms requiring Y-SAC. Y-stent placement always needs the catheterization of both bifurcation branches, which sometimes may be complex in case of steep angles. Moreover, Y-stent placement with the crossing technique, which is more common than the kissing one, needs the passage of a second stent through the interstices of the first deployed device, necessitating microcatheterization through a recently placed stent, with the risk of damage/dislodgement of the first stent.[12] However, our study showed a high technical success rate (97%), demonstrating that Y-SAC is a feasible technique in experienced centers.

While 82.2% of aneurysms were occluded immediately after treatment, nearly 95% of the lesions had complete/near-complete occlusion during the 14 months of radiologic follow-up. The immediate occlusion rate reported in the literature is quite variable, ranging from 50% to 100%,[8,12,13] whereas long-term occlusion rates appear homogeneous. Assessing the heterogeneity of the data, we found a low rate of I^2 for the long-term occlusion rate and a very high rate of I^2 for the immediate occlusion rate. It is likely that the flow-diversion effect of the 2 Y-configured stents influenced the aneurysm thrombosis, and lesions with a lower occlusion rate immediately after treatment were completely occluded at the last angiographic follow-up. Cekirge et al[14] reported a series of 8 bifurcation aneurysms treated by placing 2 stents in a Y-configuration with no accompanying endosaccular packing. Given the flow-diversion effect of the stents, 6 aneurysms had complete occlusion and 2 lesions showed residual filling during about 12 months of follow-up. Because we analyzed a period of about 13 years, it is likely that the evolution of the devices (low-profile stents compatible with 0.0165-inch inner microcatheters) and the improvement of the operator experience would have influenced the outcomes. Accordingly, although it was not statistically significant, we found a trend ($P = .056$) toward a higher occlusion rate over the analyzed period (On-line Fig 2). Given the high rate of long-term occlusion, only 3% of aneurysms were recanalized during follow-up, without statistically significant differences between the anterior and posterior circulation. These results appear comparable with those reported in a series of wide-neck aneurysms treated with PulseRider[6] and pCONus[5] devices, showing rates of Raymond-Roy 1–2 occlusion close to 95%, with 5% recanalization. Among large meta-analyses of aneurysms treated with the Woven EndoBridge (WEB; Sequent Medical, Aliso Viejo, California) devices, long-term complete/near-complete occlusion ranged from 80% to 85%, with a nearly 7% retreatment rate.[4,15]

Long-term occlusion of large and giant aneurysms is notoriously challenging after selective endovascular treatment.[17] Our study found that an occlusion rate after Y-stent placement was

lower among large/giant aneurysms (79% versus 92%, $P = .02$), and the mean size of recanalized aneurysms was higher (11 mm) compared with nonrecanalized lesions (7 mm) ($P = .0001$).

Y-SAC can be used as a first treatment or in the retreatment of recanalized aneurysms. Jeon et al,[17] in a series of 25 patients with basilar tip aneurysms treated using Y-stent placement with closed-cell stents, reported 70% complete occlusion among recanalized aneurysms, and 100% occlusion among those directly treated with Y-stent placement. Our meta-analysis confirmed higher occlusion rates after Y-stent placement used as a first treatment (94% versus 82%, $P = .01$), probably because recanalized aneurysms are generally those with more complex anatomy.

Treatment-Related Complications

One of the main concerns of Y-SAC is the risk of ischemic complications related to the increased metal density of the overlapped double stents that promotes platelet aggregation. In our meta-analysis, Y-stent placement was associated with an 8.9% complication rate, with a moderate statistical heterogeneity over the included studies. Most complications occurred in the periprocedural/early period after treatment (6.7%), with a low rate of treatment-related morbidity (2.4%) and mortality (1%). Delayed complications were 2.1%, and most were associated with asymptomatic in-stent occlusion.[18-20] Overall, thromboembolism was the most common complication (6.5%), whereas acute in-stent occlusion was reported among 2% of cases. Hemorrhagic events were uncommon and were generally related to technical complications during the crossing maneuver through the interstices of the first stent (aneurysm or vessel perforation).[8] No lesions ruptured (delayed hemorrhage) after treatment during follow-up, demonstrating that the aneurysms were successfully secured.

In a meta-analysis of nearly 2500 wide-neck aneurysms treated endovascularly, the permanent complication rates were 2% and 3% for coiling and SAC, respectively.[21] Flow-diverter stents have become a suitable tool for complex wide-neck aneurysms. However, an important concern is the patency of the covered side branches when the stent is deployed at the bifurcation points. Recently, a meta-analysis of approximately 250 MCA aneurysms treated with flow-diverter stents showed an overall complication rate of 20%, with 10% treatment-related morbidity. In addition, nearly 10% of jailed arteries were occluded during follow-up.[3] Intrasaccular flow disruption with WEB devices appeared to be relatively safe, with 8% thromboembolic complications, which were lower (6%) among studies published after 2013, due to the improvement in the operators' experience.[4] Similarly, although current evidence is still limited, PulseRider can be a suitable option for wide-neck aneurysms: A recent review showed 8% intraprocedural complications, which is quite comparable with the treatment-related outcomes of Y-SAC.[22]

According to a large series and meta-analysis of very large and giant aneurysms treated endovascularly,[16] we found higher complication rates among large-giant lesions (17%) compared with small- or medium-sized aneurysms (9%). In addition, patients who underwent retreatment with Y-stent placement of a recanalized aneurysm experienced a higher complication rate (14% versus 7.6%). This is likely a reflection of a more complex anatomic configuration and larger neck of the recanalized aneurysms; in this situation, retreatment can be technically challenging.

In our study, Y-stent placement for MCA aneurysms showed better results, especially if compared with anterior cerebral artery lesions, which were associated with complication rates close to 20%. Anterior communicating artery aneurysms usually have a complex anatomy, while pericallosal lesions are characterized by distal access and small arterial caliber, increasing the risk of complications.[20,23,24]

Crossing-versus-Kissing Technique

In general, the crossing-Y-stent placement is the most commonly used technique; the second stent is deployed through the cells of the first stent. In our study, 92% of patients were treated with the crossing technique, and 8%, with the kissing approach. The crossing technique necessitates passing a wire and stent through a recently placed stent, thus risking damage/dislodgement of the first device.[19] The kissing technique avoids this problem due to the use of 2 stents placed side by side in the parent artery with the distal segment of each stent going into 1 bifurcation vessel.[19,25] However, no studies compared these 2 techniques. In our review (On-line Table 5) complete/near complete occlusion rate (Raymon-Roi grade 1 and 2) was quite comparable, and it was achieved in 95% and 90% of aneurysms treated with crossing and kissing Y-stent placement, respectively. However, the complication rate appeared to be slightly higher after kissing stent placement (12.7% versus 8.4%), though the difference was not statistically significant, probably due to the small number of patients in the kissing group.

Brassel et al[27] reported 14% ischemic complications in a series of 7 complex aneurysms treated with the kissing-Y-technique with the closed-cell Acclino flex Stent. Complete occlusion was reported in all 7 aneurysms. Jankowitz et al[19] reported an 80% complete occlusion after Y-stent placement with the kissing technique in 11 wide-neck aneurysms. One patient had a procedural aneurysm rupture, and 1 patient experienced delayed in-stent thrombosis. Accordingly, the efficacy and safety of kissing Y-stent placement probably needs larger series to be proved.

Closed-Cell and Open-Cell Stents

Y-stent placement embolization can be achieved by using various combinations of different types of stent: closed/closed-cell (such as the Enterprise), open/open-cell (Neuroform), or a hybrid combination between closed and open-cell stents. Limbucci et al[8] reported a high rate of long-term occlusion (97%) and a relatively low rate of complications (4%) after Y-stent placement performed with 2 Enterprise stents in a crossing fashion. Laser-cut closed-cell devices had good navigability, and they are usually easily deployed because they can be partially recapturable. In addition, the linear profile of 2 closed-cell stents may allow a better scaffold and protection of the parent artery during coiling.[8] However, an important concern using 2 closed-cell stents is that the first device may limit the expansion of the second one, resulting in in-stent stenosis. In our study, Y-stent placement with 2 Enterprise devices was associated with high rates of occlusion (96%) and a low rate of complications (6.5%) (On-line Table 6).

Although infrequently specified in the included studies, most Neuroform stents were the older generation devices. Castaño et al[26] described a series of patients treated with different generations of Neuroform stents, including the Neuroform Atlas Stent System in past years, reporting 95% occlusion and a very low rate of treatment-related complications (2.5%) after Y-stent placement. The authors reported that in their experience, open-cell stents had good conformability to tortuous anatomies, minor rectification of the vessels, high vessel wall apposition, and easier microcatheterization through the mesh of the first stent. However, if one investigated the literature, complications after Y-stent placement with the Neuroform stent appear quite variable, ranging from 2% to 25%. Spiotta et al[27] reported 5 periprocedural events during Y-stent placement in 19 patients with the Neuroform stent. Our meta-analysis showed 14% complications associated with Y-stent placement using the Neuroform stent. However, most were transient events without clinical sequalae.

Recently, braided stents have been used to perform Y-SAC. Compared with other stents that are laser-cut from nitinol hypotubes, LVIS stents are braided from a single nitinol wire with a closed-cell design. Theoretically, they have a smaller cell size and higher metal coverage and flow-diversion effect than other conventional self-expandable devices.[24,28] We found 92% occlusion and 11% complications using LVIS stents.

Finally, 4 studies[12,19,29,30] reported results after Y-stent placement performed with a combination (hybrid design) of closed/open-cell stents (Enterprise/Neuroform). Usually the Neuroform stent was chosen as a first stent to stabilize the structure; a second Enterprise stent can be better deployed through the interstices of the Neuroform stent that are wider than those in closed-cell devices. In our study, the overall occlusion rate was 90%, with 7.5% treatment-related complications.

Strengths and Limitations

Our study has limitations. The series are retrospective studies or small single-institution experiences. Because of the small number of cases in some subgroup analyses, the comparisons among them may not provide power to show a statistically significant difference among the studied outcomes. However, publication bias was excluded, and our review is the largest to date.

CONCLUSIONS

Although Y-SAC is usually performed in complex aneurysms, high rates of long-term angiographic occlusion are achieved with a relatively low rate of treatment-related complications. Y-stent placement with a crossing configuration appears associated with better outcomes. Even though a Y-configuration can be obtained with many types of stents with comparable occlusion rates, the Enterprise stent appears to be associated with lower rates of complications.

Disclosures: Federico Cagnazzo—*UNRELATED: Employment*: University of Florence. Antonio Laiso—*UNRELATED: Fees for Participation in Review Activities such as Data Monitoring Boards, Statistical Analysis, Endpoint Committees. Paolo Perrini—UNRELATED:* Employment: *University of Pisa.*

REFERENCES

1. Pierot L, Wakhloo AK. **Endovascular treatment of intracranial aneurysms: current status.** *Stroke* 2013;44:2046–54 CrossRef Medline
2. Cagnazzo F, Lefevre PH, Mantilla D, et al. **Patency of the supraclinoid internal carotid artery branches after flow-diversion treatment: a meta-analysis.** *J Neuroradiol* 2018 Aug 9. [Epub ahead of print] CrossRef Medline
3. Cagnazzo F, Mantilla D, Lefevre PH, et al. **Treatment of middle cerebral artery aneurysms with flow-diverter stents: a systematic review and meta-analysis.** *AJNR Am J Neuroradiol* 2017;38:2289–94 CrossRef Medline
4. Lv X, Zhang Y, Jiang W. **Systematic review of Woven EndoBridge for wide-necked bifurcation aneurysms: complications, adequate occlusion rate, morbidity, and mortality.** *World Neurosurg* 2018;110:20–25 CrossRef Medline
5. Gory B, Aguilar-Pérez M, Pomero E, et al. **One-year angiographic results after pCONus stent-assisted coiling of 40 wide-neck middle cerebral artery aneurysms.** *Neurosurgery* 2017;80:925–33 CrossRef Medline
6. Gory B, Spiotta AM, Di Paola F, et al. **PulseRider for treatment of wide-neck bifurcation intracranial aneurysms: 6-month results.** *World Neurosurg* 2017;99:605–09 CrossRef Medline
7. Chow MM, Woo HH, Masaryk TJ, et al. **A novel endovascular treatment of a wide-necked basilar apex aneurysm by using a Y-configuration, double-stent technique.** *AJNR Am J Neuroradiol* 2004;25:509–12 Medline
8. Limbucci N, Renieri L, Nappini S, et al. **Y-stent assisted coiling of bifurcation aneurysms with Enterprise stent: long-term follow-up.** *J Neurointerv Surg* 2016;8:158–62 CrossRef Medline
9. Moher D, Liberati A, Tetzlaff J, et al. **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.** *Int J Surg* 2010;8:336–41 CrossRef Medline
10. Roy D, Milot G, Raymond J. **Endovascular treatment of unruptured aneurysms.** *Stroke* 2001;32:1998–2004 CrossRef Medline
11. Wells G, Shea B, O'Connell D. *The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonradomized Studies in Meta-Analyses.* Ottawa: Ottawal Hospital Research Institute. 2011 http://www.evidencebasedpublichealth.de/download/Newcastle_Ottowa_Scale_Pope_Bruce.pdf. Accessed August 20, 2018
12. Fargen KM, Mocco J, Neal D, et al. **A multicenter study of stent-assisted coiling of cerebral aneurysms with a Y configuration.** *Neurosurgery* 2013;73:466–72 CrossRef Medline
13. Yavuz K, Geyik S, Cekirge S, et al. **Double stent-assisted coil embolization treatment for bifurcation aneurysms: immediate treatment results and long-term angiographic outcome.** *AJNR Am J Neuroradiol* 2013;34:1778–84 CrossRef Medline
14. Cekirge HS, Yavuz K, Geyik S, et al. **A novel "Y" stent flow diversion technique for the endovascular treatment of bifurcation aneurysms without endosaccular coiling.** *AJNR Am J Neuroradiol* 2011;32:1262–68 CrossRef Medline
15. Asnafi S, Rouchaud A, Pierot L, et al. **Efficacy and safety of the Woven EndoBridge (WEB) device for the treatment of intracranial aneurysms: a systematic review and meta-analysis.** *AJNR Am J Neuroradiol* 2016;37:2287–92 CrossRef Medline
16. Cagnazzo F, Mantilla D, Rouchaud A, et al. **Endovascular treatment of very large and giant intracranial aneurysms: comparison between reconstructive and deconstructive techniques—a meta-analysis.** *AJNR Am J Neuroradiol* 2018;39:852–58 CrossRef Medline
17. Jeon P, Kim BM, Kim DJ, et al. **Y-configuration double-stent-assisted coiling using two closed-cell stents for wide-neck basilar tip aneurysms.** *Acta Neurochir (Wien)* 2014;156:1677–86 CrossRef Medline
18. Bartolini B, Blanc R, Pistocchi S, et al. **"Y" and "X" stent-assisted coiling of complex and wide-neck intracranial bifurcation aneurysms.** *AJNR Am J Neuroradiol* 2014;35:2153–58 CrossRef Medline
19. Jankowitz BT, Thomas A, Jovin T, et al. **Y stenting using kissing**

stents for the treatment of bifurcation aneurysms. *J Neurointerv Surg* 2012;4:16–21 CrossRef Medline
20. Ko JK, Han IH, Cho WH, et al. **Crossing Y-stent technique with dual open-cell stents for coiling of wide-necked bifurcation aneurysms.** *Clin Neurol Neurosurg* 2015;132:54–60 CrossRef Medline
21. Zhao B, Yin R, Lanzino G, et al. **Endovascular coiling of wide-neck and wide-neck bifurcation aneurysms: a systematic review and meta-analysis.** *AJNR Am J Neuroradiol* 2016;37:1700–05 CrossRef Medline
22. Aguilar-Salinas P, Brasiliense LBC, Walter CM, et al. **Current status of the PulseRider in the treatment of bifurcation aneurysms: a systematic review.** *World Neurosurg* 2018;115:288–94 CrossRef Medline
23. Lee K, Park H, Park I, et al. **Y-configuration stent-assisted coil embolization for wide-necked intracranial bifurcation aneurysms.** *J Cerebrovasc Endovasc Neurosurg* 2016;18:355–62 CrossRef Medline
24. Samaniego EA, Mendez AA, Nguyen TN, et al. **LVIS Jr device for Y-stent-assisted coil embolization of wide-neck intracranial aneurysms: a multicenter experience.** *Interv Neurol* 2018;7:271–83 CrossRef Medline
25. Brassel F, Melber K, Schlunz-Hendann M, et al. **Kissing-Y stenting for endovascular treatment of complex wide necked bifurcation aneurysms using Acandis Acclino stents: results and literature review.** *J Neurointerv Surg* 2016;8:386–95 CrossRef Medline
26. Castaño C, Terceno M, Remollo S, et al. **Endovascular treatment of wide-neck intracranial bifurcation aneurysms with 'Y'-configuration, double Neuroform stents-assisted coiling technique: experience in a single center.** *Interv Neuroradiol* 2017;23:362–70 CrossRef Medline
27. Spiotta AM, Gupta R, Fiorella D, et al. Mid-term results of endovascular coiling of wide-necked aneurysms using double stents in a Y configuration. *Neurosurgery* 2011;69:421–29
28. Santillan A, Boddu S, Schwarz J, et al. **LVIS Jr. stent for treatment of intracranial aneurysms with parent vessel diameter of 2.5 mm or less.** *Interv Neuroradiol* 2018;24:246–53 CrossRef Medline
29. Akgul E, Balli T, Aksungur EH. **Hybrid, Y-configured, dual stent-assisted coil embolization in the treatment of wide-necked bifurcation aneurysms.** *Interv Neuroradiol* 2015;21:29–39 CrossRef Medline
30. Takano N, Suzuki M, Irie R, et al. **Usefulness of non-contrast-enhanced MR angiography using a silent scan for follow-up after Y-configuration stent-assisted coil embolization for basilar tip aneurysms.** *AJNR Am J Neuroradiol* 2017;38:577–81 CrossRef Medline
31. Thorell WE, Chow MM, Woo HH, et al. **Y-configured dual intracranial stent-assisted coil embolization for the treatment of wide-necked basilar tip aneurysms.** *Neurosurgery* 2005;56:1035–40; discussion 1035–40 Medline
32. Zhao KJ, Yang PF, Huang QH, et al. **Y-configuration stent placement (crossing and kissing) for endovascular treatment of wide-neck cerebral aneurysms located at 4 different bifurcation sites.** *AJNR Am J Neuroradiol* 2012;33:1310–16 CrossRef Medline
33. Lee WJ, Cho CS. **Y-stenting endovascular treatment for ruptured intracranial aneurysms: a single-institution experience in Korea.** *J Korean Neurosurg Soc* 2012;52:187–92 CrossRef Medline
34. Chalouhi N, Jabbour P, Gonzalez LF, et al. **Safety and efficacy of endovascular treatment of basilar tip aneurysms by coiling with and without stent assistance: a review of 235 cases.** *Neurosurgery* 2012;71:785–94 CrossRef Medline
35. Johnson AK, Heiferman DM, Lopes DK. **Stent-assisted embolization of 100 middle cerebral artery aneurysms.** *J Neurosurg* 2013;118:950–55 CrossRef Medline
36. Möhlenbruch M, Herweh C, Behrens L, et al. **The LVIS Jr. microstent to assist coil embolization of wide-neck intracranial aneurysms: clinical study to assess safety and efficacy.** *Neuroradiology* 2014;56:389–95 CrossRef Medline
37. Heller RS, Rahal JP, Malek AM. **Y-Stent embolization technique for intracranial bifurcation aneurysms.** *J Clin Neurosci* 2014;21:1368–72 CrossRef Medline
38. Johnson AK, Munich SA, Tan LA, et al. **Complication analysis in nitinol stent-assisted embolization of 486 intracranial aneurysms.** *J Neurosurg* 2015;123:453–59 CrossRef Medline
39. Melber K, Meila D, Draheim P, et al. **Vascular angular remodeling by kissing-Y stenting in wide necked intracranial bifurcation aneurysms.** *J Neurointerv Surg* 2017;9:1233–37 CrossRef
40. Bang JS, Kim CH, Kwon BJ, et al. **The difficulties and risks of Y-stent-assisted coiling: a comparison of first and second stenting procedures.** *World Neurosurg* 2016;88:146–53 CrossRef Medline
41. Cheung NK, Chiu AH, Cheung A, et al. **Long term follow-up of bifurcation aneurysms treated with braided stent assisted coiling and complex T- and Y-stent constructs.** *J Neurointerv Surg* 2018;10:560–65 CrossRef Medline

Osseous versus Nonosseous Spinal Epidural Arteriovenous Fistulas: Experiences of 13 Patients

Y. Song, S.H. Cho, D.W. Lee, J.J. Sheen, J.H. Shin, and D.C. Suh

ABSTRACT

BACKGROUND AND PURPOSE: Spinal epidural arteriovenous fistulas are rare vascular malformations. We present 13 patients with spinal epidural arteriovenous fistulas, noting the various presenting symptom patterns, imaging findings related to bone involvement, and outcomes.

MATERIALS AND METHODS: Among 111 patients with spinal vascular malformations in the institutional data base from 1993 to 2017, thirteen patients (11.7%) had spinal epidural arteriovenous fistulas. We evaluated presenting symptoms and imaging findings, including bone involvement and mode of treatment. To assess the treatment outcome, we compared initial and follow-up clinical status using the modified Aminoff and Logue Scale of Disability and the modified Rankin Scale.

RESULTS: The presenting symptoms were lower back pain ($n = 2$), radiculopathy ($n = 5$), and myelopathy ($n = 7$). There is overlap of symptoms in 1 patient (No. 11). Distribution of spinal epidural arteriovenous fistulas was cervical ($n = 3$), thoracic ($n = 2$), lumbar ($n = 6$), and sacral ($n = 2$). Intradural venous reflux was identified in 7 patients with congestive venous myelopathy. The fistulas were successfully treated in all patients who underwent treatment (endovascular embolization, $n = 10$; operation, $n = 1$) except 2 patients who refused treatment due to tolerable symptoms. Transarterial glue ($n = 7$) was used in nonosseous types; and transvenous coils ($n = 3$), in osseous type. After 19 months of median follow-up, the patients showed symptom improvement after treatment.

CONCLUSIONS: Although presenting symptoms were diverse, myelopathy caused by intradural venous reflux was the main target of treatment. Endovascular treatment was considered via an arterial approach in nonosseous types and via a venous approach in osseous types.

ABBREVIATIONS: ALS = Aminoff and Logue Scale of Disability; SDAVF = spinal dural arteriovenous fistula; SEDAVF = spinal epidural arteriovenous fistula; TAE = transarterial embolization; TVE = transvenous embolization

Spinal epidural arteriovenous fistula (SEDAVF) is a rare vascular disease of the spine that manifests with benign symptoms, such as radiculopathy.[1,2] Because of advances in imaging techniques, such as 3D rotational spinal angiography, and improvement in the understanding of the anatomy of spinal vascular lesions, SEDAVFs are diagnosed with increasing frequency.[3,4]

SEDAVFs are fistulas located in the epidural space between the branches of the radicular artery (mostly dorsal somatic branches) and the epidural venous plexus,[4,5] whereas the typical location of more common spinal dural arteriovenous fistulas (SDAVFs) is within the dural sleeve between the radiculomeningeal artery and the radicular vein or bridging vein.[4] In most cases of SEDAVFs, the venous flow drains only to the epidural and paravertebral plexuses, and the lesion is generally asymptomatic until dilation of arterialized veins causes compressive symptoms or intradural reflux occurs.[5] Differentiation from other spinal vascular pathology, including SDAVF, is mandatory to localize the exact cause of presenting symptoms.[6,7]

SEDAVFs can be categorized into 2 distinct presenting symptom patterns, depending on the involvement of the intradural vein.[8] With intradural venous reflux, they may mimic SDAVFs with symptoms of congestive myelopathy such as lower leg weakness and sphincter dysfunction.[6,9-11] In SEDAVFs without intradural venous drainage, compressive radiculopathy or myelopathy caused by epidural vascular lesions may be the main problem.[12]

Received June 14, 2018; accepted October 12.

From the Departments of Radiology and Research Institute of Radiology (Y.S., D.W.L., J.J.S., J.H.S., D.C.S.), University of Ulsan, College of Medicine, Asan Medical Center, Seoul, Republic of Korea; and Department of Neurosurgery (S.H.C.), Ulsan University College of Medicine, Asan Hospital, Gang-reung, Republic of Korea.

This work was supported by the National Research Foundation of Korea grant funded by the Korean government (No. 2018R1A2B6003143).

Please address correspondence to Dae Chul Suh, MD, PhD, Department of Radiology, Asan Medical Center, University of Ulsan, College of Medicine, 88, Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea; e-mail: dcsuh@amc.seoul.kr

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

Indicates article with supplemental on-line table.

http://dx.doi.org/10.3174/ajnr.A5904

Endovascular embolization, an operation, and combined therapy have been the main treatment methods for SEDAVFs. Although various endovascular treatment options in terms of approach routes and embolic materials have been reported,[2] the best treatment method has not been determined because of the complex anatomy of the target lesion. An understanding of the manifestations and anatomy of the lesion is the key to accurate diagnosis and proper management.[13] We present 13 cases of SEDAVF from Asan Medical Center, studying their various clinical manifestations, radiologic features, and outcomes after endovascular treatment.

MATERIALS AND METHODS

Patients

We retrospectively reviewed a data base of our institution, a tertiary hospital, that was maintained prospectively since 1993 to identify all spinal vascular malformations (ie, spinal arteriovenous malformation, spinal dural/epidural arteriovenous fistula, and vertebrovertebral fistula). Among 111 spinal vascular malformations diagnosed at our neurointervention suite from January 1993 to April 2017, we included 13 patients with SEDAVFs. The diagnosis was confirmed by high-resolution selective spinal angiography. Patient information regarding the clinical history, neurologic symptoms and signs, and follow-up results was obtained from the data base connected to the electronic medical record. Patients' symptoms were classified as myelopathy and radiculopathy according to the initial neurologic symptoms and signs. "Myelopathy" was defined as spinal cord symptoms with gait disturbance, motor/sensory deficits, or urinary/bowel sphincter dysfunction; "radiculopathy" was defined as a series of symptoms along the specific dermatomal distribution. We evaluated the presence of functional disability using the modified Aminoff and Logue Scale of Disability (ALS)[14] and the modified Rankin Scale.[15] The scores were based on the records of the patients' symptoms and neurologic examinations, which were described by neurologists not directly involved in treatment.

Imaging Diagnosis

SEDAVFs were defined as arteriovenous shunts located in the epidural space, fed by multiple epidural arterial branches and draining first to the epidural venous pouch and then into the paravertebral vein and, in some cases, into the intradural vein. Direct, high-flow, single-hole fistulas between the vertebral artery and the vertebral venous plexus (also known as vertebrovertebral fistulas) or paravertebral fistulous lesions in the extradural space were excluded from this study.[6]

The venous drainage pattern of the SEDAVFs was analyzed on the basis of the 2D angiographic findings to determine whether there was retrograde intradural reflux through the radicular vein. MR imaging, CT, or dual 3D rotational angiography or a combination of these was used to demonstrate bone involvement of the SEDAVFs.[13] The "osseous type" was defined as an SEDAVF with bone involvement, such as bone erosion, osteolytic destruction, or compression fracture (Fig 1).

Spinal MR imaging was performed for initial assessment of the disease. Axial and sagittal images of T1- and T2-weighted sequences with or without contrast enhancement were obtained with 1.5T or 3T systems. The presence and extent of spinal cord edema, perimedullary flow voids, and any epidural lesion that compressed the nerve root or thecal sac were assessed on imaging. Patients who needed imaging evaluation clinically after treatment underwent follow-up MR imaging.

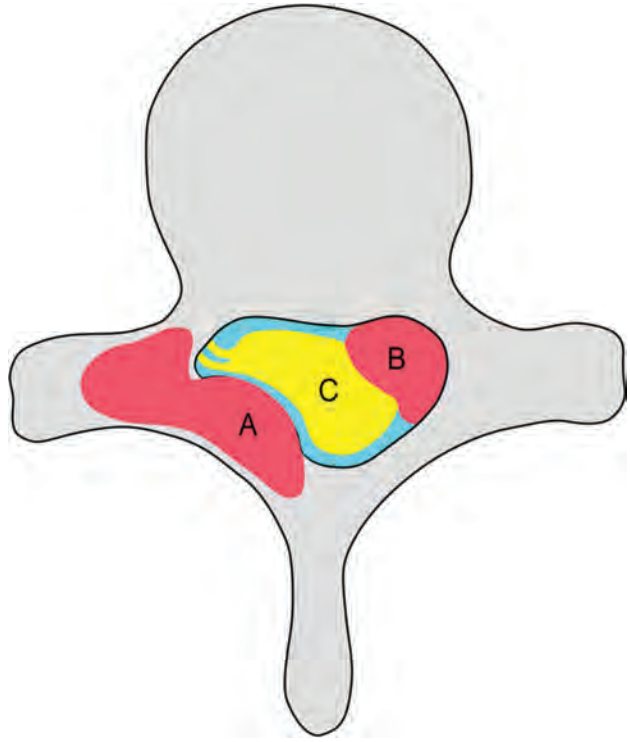


FIG 1. Schematic diagram of a spinal epidural arteriovenous fistula. An osseous SEDAVF shows bone involvement and compression of the nerve root or the spinal cord by bulging with cortical erosion (*A*). A nonosseous SEDAVF shows engorgement (fistulous sac or pouch) of the epidural vein (*B*) within the spinal canal compressing the spinal cord (*C*).

Spinal digital subtraction angiography (Artis zee; Siemens, Erlangen, Germany) was used for localization and characterization of the SEDAVFs. All possible feeders—including vertebral arteries, thyrocervical and costocervical trunks in the cervical lesions, and lateral and medial sacral arteries from internal iliac arteries in the lumbosacral lesions—and all segmental arteries were scrutinized. 3D rotational angiography further clarified the location of and relationship among the arterial feeders, fistulous sac (pouch), arterialized veins, and the point where intradural reflux occurred.

Treatment

Endovascular embolization was considered primarily when the vascular approach to the fistula appeared angiographically possible. There were 2 modes of treatment in which transarterial embolization (TAE) was used for nonosseous SEDAVFs; and transvenous embolization (TVE), for osseous SEDAVFs. The treatment strategy of nonosseous SEDAVFs was basically similar to that for SDAVFs: to occlude the shunt and the proximal drainage vein. The range of the embolization target was the point at which intradural reflux occurred, to obliterate retrograde venous drainage into the radicular vein leading to congestive venous myelopathy. The glue, a mixture of *N*-butyl-2-cyanoacrylate (Histoacryl®; B. Braun, Melsungen, Germany) and iodized oil (Lipiodol®; Guerbet, Roissy, France), was used at the wedged position via a transarterial approach.[16] The TVE was used when multiple arterial feeders converged on the fistulous sac in the osseous defect of the vertebral column and the fistulous sac could

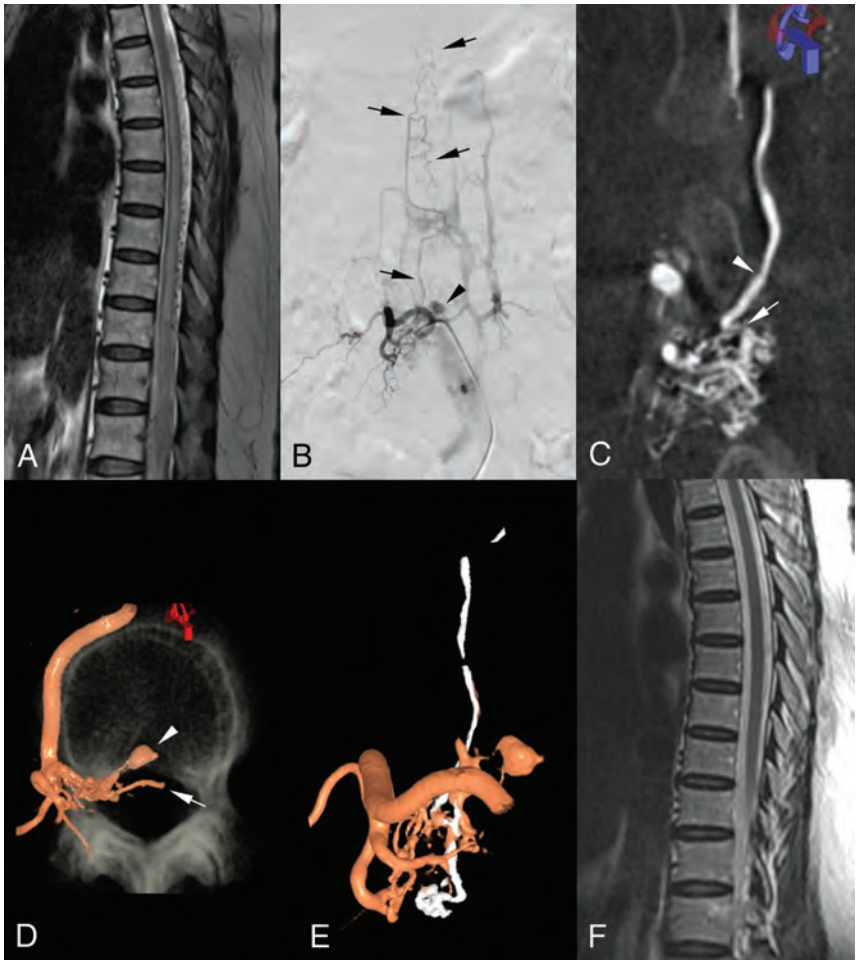


FIG 2. Spinal epidural arteriovenous fistula (perimedullary venous drainage) in a 63-year-old woman (case 4). *A*, Sagittal T2-weighted MR imaging reveals diffuse spinal cord edema up to the T5 level and multiple signal voids in the posterior aspect of the cord. *B*, Right lumbar arteriogram shows an arteriovenous fistula at the L2 level with multiple fine feeding arteries and early venous drainage to the epidural and paravertebral venous plexuses and to intradural veins (*arrows*). A small venous sac is visible around the fistula (*arrowhead*). *C*, 3D rotational angiogram shows the anatomy in detail where the intradural reflux originated (*arrow*). Note the focal narrowing of the vein where it penetrates the dura (*arrowhead*). *D*, Volume-rendering image demonstrates the epidural location of the fistula and the venous pouch (*arrowhead*). An intradural course of the radicular vein is also visible (*arrow*). *E*, Postembolization 3D angiogram confirms the presence of radiopaque glue in the fistula and along the intradural vein (*white areas*). *F*, Follow-up MR imaging 2 months later shows that the diffuse cord edema and the perimedullary vessels have disappeared.

be approached via a transvenous route through the epidural vein. Various types of detachable and pushable coils were used for TVE.

"Complete occlusion" was defined as disappearance of the fistula or shunted vein and an obliteration of the intradural venous drainage by disconnecting the fistulous sac with retrograde radicular venous drainage in the SEDAVF with congestive venous myelopathy (Fig 2). "Partial occlusion" was defined as having residual retrograde flow into the retrograde intradural vein, especially in SEDAVFs with congestive venous myelopathy or delayed opacification of the fistula or shunted veins on the venous phase with prominent contrast stasis in SEDAVFs with radiculopathy without retrograde intradural venous drainage.

Outcome and Statistical Consideration

The patients' neurologic symptoms and functional abilities were recorded with the use of the ALS and mRS scores on every follow-up. We evaluated the outcomes by comparing the scores at initial status with those at the most recent follow-up. We used STATA, Version 13.0 (StataCorp, College Station, Texas) to perform statistical analysis. All data were organized as the mean and range of continuous variables and as the number and percentage of categoric variables. Statistical significance ($P < .05$) among the scores was determined with a Student *t* test or the Wilcoxon rank sum test (if the assumption of normality was violated on the Shapiro-Wilk normality test).

RESULTS

Baseline Characteristics

A general overview of the patients' information is presented in the On-line Table. Of the 13 patients, 10 were male and 3 were female; the mean age was 49.6 years (range, 16–67 years). The mean duration from symptom onset to diagnosis was 3.6 months (range, 1 week to 24 months). Bilateral lower extremity weakness and sphincter dysfunction were typical manifestations of congestive myelopathy, which was noted in 7 patients with intradural venous drainage. Of the 6 patients without intradural reflux, 4 patients had radiculopathy or compressive myelopathy, which manifested mostly with pain, weakness, and paresthesia along the particular dermatome. The other 2 patients had only pain without neurologic symptoms. Three patients had a history of trauma (ie, case 1: T12 compression fracture due to a fall 3 years before; case 7: L1 compression fracture due to fall 3 years before; case 13: a fall 2.5 years before without vertebral fracture).

Imaging Diagnosis

Initial MR imaging, which was performed in all 13 patients, showed abnormal findings. Both spinal cord edema and perimedullary flow voids were noted in 7 patients with congestive myelopathy. Of the 6 patients without intradural reflux, 1 patient demonstrated focal signal change in the spinal cord, caused by the mass effect of the epidural vascular lesion. The other 5 patients demonstrated only epidural vascular lesions without spinal cord involvement. Among 5 patients who had an osseous type of SEDAVF, 4 patients underwent CT, which showed bone involvement at the fistula level.

2D and/or 3D spinal DSA confirmed the presence of epidural arteriovenous fistulas in all 13 patients. The lesions were categorized into 2 groups according to the existence of intradural reflux. Seven patients showed intradural drainage through the radicu-

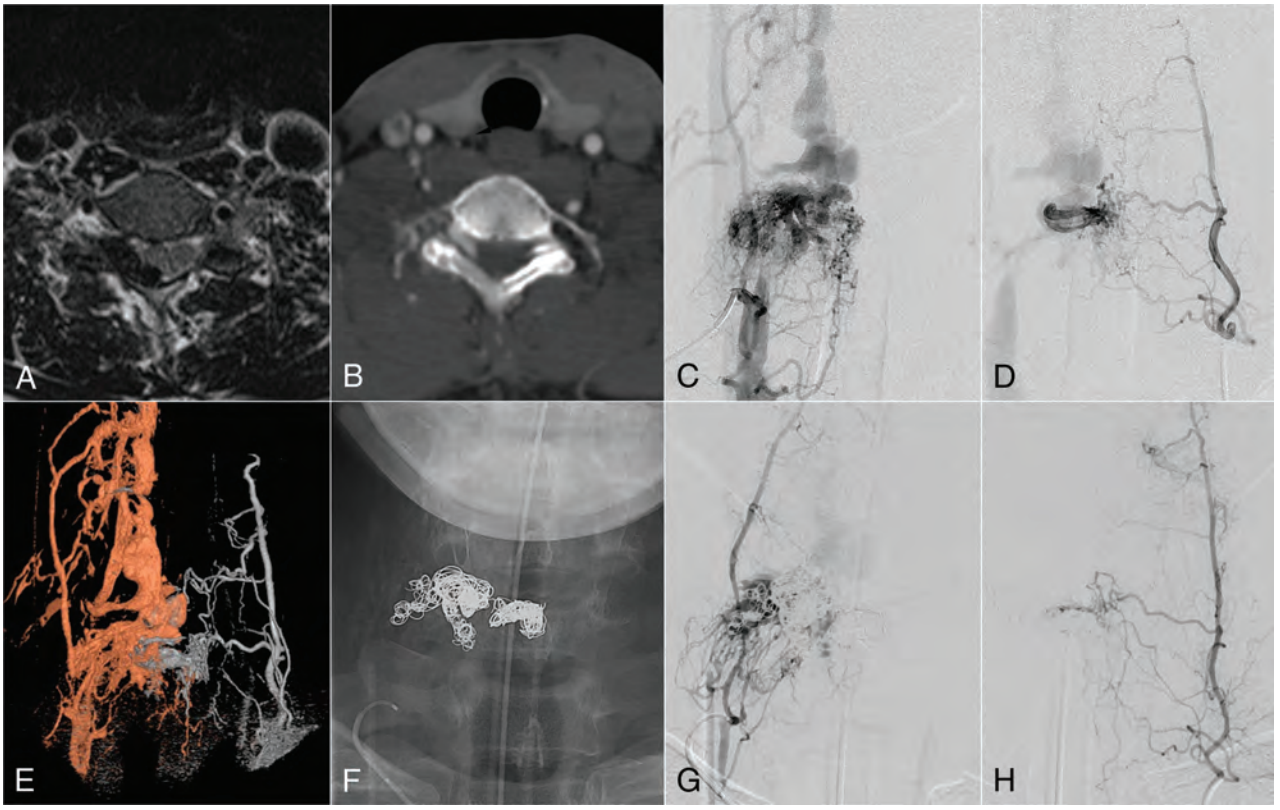


FIG 3. Osseous spinal epidural arteriovenous fistula in a 57-year-old man (case 9). *A*, T2-weighted axial MR image reveals a well-defined lesion with dark signal intensity in the right epidural space at the C6 level. *B*, CT scan at the corresponding level shows adjacent bone destruction of the right lamina and spinous process of the C6 vertebra that resulted from the well-enhanced epidural lesion. *C* (*right*) and *D* (*left*), Deep cervical arteriograms. A large arteriovenous fistula with multiple arterial feeders is visible at the C6 level. *E*, 3D fusion image demonstrates that the feeders from both sides converge on a focal region—the epidural venous plexus and internal jugular veins—from which venous flow drains exclusively via extradural veins. *F*, After 2 sessions of transvenous coil embolization, the fistula flow has almost completely disappeared (*G* and *H*).

lomedullary and perimedullary veins. Of those 7, only 3 patients had solely intradural drainage without any extradural component. The distribution of the fistulas in the intradural reflux group were thoracic ($n = 2$), lumbar ($n = 3$), and sacral ($n = 2$), while that of the pure extradural group was cervical ($n = 3$) and lumbar ($n = 3$). Adjacent bone involvement of the SEDAVF was demonstrated in 5 patients on 3D angiography.

Treatment

Among the 13 patients with SEDAVFs, 10 patients received endovascular embolization, 1 patient underwent an operation after the endovascular treatment was aborted, and 2 patients refused treatment due to tolerable symptoms (On-line Table). Endovascular embolization comprised TAE in 7 patients (Fig 3) and TVE in 3 patients (Fig 4). TAE was selected in all patients with nonosseous SEDAVFs. The embolic material used in TAE was glue in all patients and additional coils in 1 patient. Transvenous coil embolization was performed in 3 patients with osseous SEDAVFs. Complete obliteration of the fistula or intradural drainage vein was achieved in all 6 patients with congestive myelopathy. Among the patients without intradural venous reflux, complete occlusion was possible in 1 patient, and partial occlusion, in 3 patients because complete obliteration was not usually required. Surgical disconnection of the intradural drainage vein was successfully performed in the patient in whom the endovascular approach

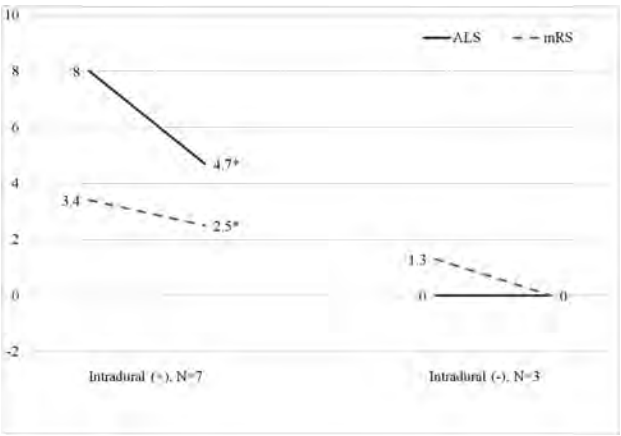


FIG 4. Outcomes in patients according to each scoring system and the presence of intradural reflux. The mean scores of the initial and follow-up evaluations are shown. *Asterisks* indicate statistically significant improvement at follow-up.

failed due to complex vascular anatomy. There was no periprocedural complication in all treated cases.

Outcomes

Among 11 patients who underwent either endovascular treatment ($n = 10$) or an operation ($n = 1$), the mean duration of follow-up after treatment was 23 months (range, 1–63 months;

median, 19 months). The other 2 patients without treatment were lost to follow-up. The mean ALS and mRS scores were significantly decreased at follow-up in the intradural reflux group (8.0 versus 4.7, $P = .01$; 3.4 versus 2.5, $P = .016$, respectively) (Fig 4). On the other hand, the group without intradural reflux showed a tendency toward decreasing scores without significance. The overall scores were higher in the patients with intradural reflux in all the grading systems. The scores of ALS were zero in the patients without intradural reflux at both initial assessment and follow-up. Symptoms and functional status improved, or at least did not change, after treatment in all except 1 patient (case 5). In this patient, mild lower back pain developed 4 months after treatment despite improvement in other symptoms.

DISCUSSION

Endovascular treatment is currently the favored method of SEDAVF treatment; the use of endovascular treatment has been reported in approximately two-thirds of pertinent studies and case reports.[2] The rates of complete or near-complete occlusion of the SEDAVF with endovascular treatment ranged from 73.3% to 94.4%, and the rates of symptom improvement ranged from 62.5% to 91%.[1-3] Our study also demonstrated an excellent occlusion rate (100%) and good treatment outcomes (91%). Intradural venous drainage, noted in 53.8%[7] of the 13 patients in this study, was seen with variable frequency in previous studies (27.0%[2] to 74.5%[4]). In this study, ALS scores were zero in all patients without intradural venous drainage except for 1 patient who had mild leg weakness caused by compressive myelopathy.

Although endovascular treatment produced good outcomes in most patients in this study, 1 patient (case 7) could not recover from the congestive myelopathy symptoms, even after complete occlusion of the shunt and improved cord edema at 1-month follow-up MR imaging. Although further follow-up is required, this intractable case may be attributed to a 24-month delay in diagnosis and the spinal cord possibly being irreversibly damaged as in patients with SDAVFs.[6] In the patient who finally underwent an operation, the endovascular approach was ruled out by the complicated vascular anatomy in which the microcatheter tip could not reach the appropriate position to occlude the shunt. An operation was a good alternative in this patient, and symptoms improved as in patients with SDAVFs.[17]

The vascular structures surrounding the lesion must be analyzed thoroughly to determine appropriate treatment plans. We routinely obtain a 3D rotational spinal angiogram to better understand the relationships among the feeders, fistulas, and draining veins.[13] Accurate localization of regurgitation from the epidural venous plexus to the intradural vein is especially crucial to treat the lesion with intradural reflux because during TAE, the liquid embolic material must penetrate the intradural vein to cure the disease. The microcatheter tip should be placed at the appropriate point from which the preferential flow goes through the intradural vein before embolic material is injected to achieve this goal. We used glue for TAE because it is easy to control, injection time is relatively short, and it has better visibility than Onyx (Covidien, Irvine, California).[8,16] The use of Onyx for SEDAVF embolization has been limited by difficulty in achieving venous penetration.[16]

This study included 5 cases (38.4%) of osseous SEDAVFs, even though only a few cases of SEDAVFs with bone involvement have been reported previously.[18-20] The mechanisms of formation of the bone lesion and its clinical significance are uncertain. When planning a treatment option, identification of the osseous type can be important. Because Osseous SEDAVFs tend to be a high-flow shunt with large epidural venous spaces, a transvenous approach may be possible.[8] In our cases, 3 underwent TVE, 1 underwent an operation, and the other one was under observation because the symptoms were tolerable. Although trauma was suggested as a cause of SEDAVFs, especially in the osseous type,[19,20] only 1 patient in our study (case 1) was confirmed to have a history of trauma in which a T12 compression fracture was caused by a fall 3 years before the initial presentation. Because the patient's symptoms developed after the trauma and the fistula occurred at the same level, we assumed that the SEDAVF could have been the result of the trauma, despite the long interval. A history of previous trauma was also noted in 2 other patients (cases 7 and 13) with the nonosseous type of SEDAVFs in our study.

A limitation of this study could be that the number of patients was very small because of the rarity of SEDAVFs. Diagnosing the patients only through imaging findings without pathologic confirmation could be another limitation. Follow-up spinal MR imaging or DSA was not performed in all patients because we did not do routine imaging studies as long as symptoms improved. Long-term follow-up will be required because not all patients were followed up for a sufficient time.

CONCLUSIONS

SEDAVFs were found in 13 (11.7%) patients of the all spinal vascular malformations in our center. The presenting patterns were diverse, from tolerable symptoms ($n = 2$) and radiculopathy ($n = 5$) to congestive venous myelopathy ($n = 7$). There is overlap of symptoms in 1 patient (No. 11). The intradural venous reflux was identified in all 7 patients with congestive venous myelopathy. Osseous SEDAVFs were observed in 5 patients in whom the fistulas were successfully controlled by TVE ($n = 3$) except 2 patients who underwent an operation ($n = 1$) and only clinical observation ($n = 1$). Nonosseous SEDAVFs were observed in 8 patients (intradural reflux in 5) in whom obliteration of the fistula including intradural venous drainage was achieved by TAE, except 1 patient who refused treatment due to tolerable symptoms and wanted clinical observation. The SEDAVFs were successfully treated in 11 patients by endovascular embolization ($n = 10$) and an operation ($n = 1$). After 19 months of median follow-up, most patients showed symptom improvement after treatment.

REFERENCES

1. Burkhardt JK, Safaee MM, Clark AJ, et al. **Sacral epidural arteriovenous fistulas: imitators of spinal dural arteriovenous fistulas with different pathologic anatomy—report of three cases and review of the literature.** *Acta Neurochir (Wien)* 2017;159:1087–92 CrossRef Medline
2. Huang W, Gross BA, Du R. **Spinal extradural arteriovenous fistulas: clinical article.** *J Neurosurg Spine* 2013;19:582–90 CrossRef Medline
3. Nasr DM, Brinjikji W, Clarke MJ, et al. **Clinical presentation and treatment outcomes of spinal epidural arteriovenous fistulas.** *J Neurosurg Spine* 2017;26:613–20 CrossRef Medline
4. Kiyosue H, Matsumaru Y, Niimi Y, et al. **Angiographic and clinical**

characteristics of thoracolumbar spinal epidural and dural arteriovenous fistulas. *Stroke* 2017;48:3215–22 CrossRef Medline
5. Ramanathan D, Levitt MR, Sekhar LN, et al. **Management of spinal epidural arteriovenous fistulas: interventional techniques and results.** *J Neurointerv Surg* 2014;6:144–49 CrossRef Medline
6. Park JE, Koo HW, Liu H, et al. **Clinical characteristics and treatment outcomes of spinal arteriovenous malformations.** *Clin Neuroradiol* 2018;28:39–46 CrossRef Medline
7. Jung SC, Song Y, Cho SH, et al. **Endovascular management of aneurysms associated with spinal arteriovenous malformations.** *J Neurointerv Surg* 2018;10:198–203 CrossRef Medline
8. Chul Suh D, Gon Choi C, Bo Sung K, et al. **Spinal osseous epidural arteriovenous fistula with multiple small arterial feeders converging to a round fistular nidus as a target of venous approach.** *AJNR Am J Neuroradiol* 2004;25:69–73 Medline
9. Lim SM, Choi IS. **Spinal epidural arteriovenous fistula: a unique pathway into the perimedullary vein—a case report.** *Interv Neuroradiol* 2009;15:466–69 CrossRef Medline
10. Krings T, Mull M, Bostroem A, et al. **Spinal epidural arteriovenous fistula with perimedullary drainage: case report and pathomechanical considerations.** *J Neurosurg Spine* 2006;5:353–58 CrossRef Medline
11. Lee CS, Pyun HW, Chae EY, et al. **Reversible aggravation of neurological deficits after steroid medication in patients with venous congestive myelopathy caused by spinal arteriovenous malformation.** *Interv Neuroradiol* 2009;15:325–29 CrossRef Medline
12. Brinjikji W, Yin R, Nasr DM, et al. **Spinal epidural arteriovenous fistulas.** *J Neurointerv Surg* 2016;8:1305–10 CrossRef Medline
13. Suh DC, Kim HS, Baek HJ, et al. **Angioarchitecture of spinal dural arteriovenous fistula: evaluation with 3D rotational angiography.** *Neurointervention* 2012;7:10–16 CrossRef Medline
14. Aminoff MJ, Logue V. **The prognosis of patients with spinal vascular malformations.** *Brain* 1974;97:211–18 CrossRef Medline
15. van Swieten JC, Koudstaal PJ, Visser MC, et al. **Interobserver agreement for the assessment of handicap in stroke patients.** *Stroke* 1988;19:604–07 CrossRef Medline
16. Suh DC, Cho SH, Park JE, et al. **Induced-wedge technique to improve liquid embolic agent penetration into spinal dural arteriovenous fistula.** *World Neurosurg* 2016;96:309–15 CrossRef Medline
17. Adrianto Y, Yang KH, Koo HW, et al. **Concomitant origin of the anterior or posterior spinal artery with the feeder of a spinal dural arteriovenous fistula (SDAVF).** *J Neurointerv Surg* 2017;9:405–10 CrossRef Medline
18. Suh DC, Kim JH, Lee MS, et al. **Penetration difference of N-butyl2-cyanoacrylate into the nidus in the embolisation of brain arteriovenous malformation.** *Interv Neuroradiol* 1998;4:63–74
19. Jin YJ, Chung SK, Kwon OK, et al. **Spinal intraosseous arteriovenous fistula in the fractured vertebral body.** *AJNR Am J Neuroradiol* 2010;31:688–90 CrossRef Medline
20. Ou CH, Wang HK, Yang TH, et al. **Spinal intraosseous epidural arteriovenous fistula with perimedullary drainage obliterated with Onyx embolization: case report.** *J Neurosurg Spine* 2015;23:250–53 CrossRef Medline

Bone Subtraction Iodine Imaging Using Area Detector CT for Evaluation of Skull Base Invasion by Nasopharyngeal Carcinoma

T. Hiyama, H. Kuno, K. Sekiya, S. Tsushima, O. Sakai, M. Kusumoto, and T. Kobayashi

ABSTRACT

BACKGROUND AND PURPOSE: Conventional CT has generally lower detectability of bone marrow invasion than MR imaging due to lower tissue contrast. The purpose of this study was to compare the diagnostic performance of conventional CT alone or in combination with bone subtraction iodine imaging using area detector CT for the evaluation of skull base invasion in patients with nasopharyngeal carcinoma.

MATERIALS AND METHODS: Forty-four consecutive patients who underwent contrast-enhanced CT using 320–row area detector CT and contrast-enhanced MR imaging for nasopharyngeal carcinoma staging between April 2012 and November 2017 were enrolled in this retrospective study. Bone subtraction iodine images were generated by subtracting pre- and postcontrast volume scans using a high-resolution deformable registration algorithm. Two blinded observers evaluated skull base invasion at multiple sites (sphenoid body, clivus, bilateral base of the pterygoid process, and petrous bone) using conventional CT images alone or in combination with bone subtraction iodine images. Examination of MR and CT images by an experienced neuroradiologist was the reference standard for evaluating sensitivity, specificity, and area under the receiver operating characteristic curve.

RESULTS: Twenty-six patients (59%) showed skull base invasion at 84 sites on the reference standard. Conventional CT plus bone subtraction iodine images showed higher sensitivity (92.9% versus 78.6%, $P = .02$) and specificity (95.6% versus 86.1%, $P = .01$) than conventional CT images alone for evaluating skull base invasion. The area under the receiver operating characteristic curve for conventional CT plus bone subtraction iodine (0.98) was significantly larger ($P < .001$) than the area under the receiver operating characteristic curve for conventional CT alone (0.90).

CONCLUSIONS: Conventional CT plus bone subtraction iodine performs more closely to the accuracy of combining CT and MR imaging compared with conventional CT alone.

ABBREVIATIONS: AUC = area under the receiver operating characteristic curve; BSI = bone subtraction iodine; CCT = conventional CT; NPC = nasopharyngeal carcinoma

Nasopharyngeal carcinoma (NPC), arising from the epithelial lining of the nasopharynx, has a propensity to invade the skull base just above the nasopharynx. Skull base invasion indicates a tumor stage of at least T3 according to the eighth edition of the American Joint Committee on Cancer staging system.[1] NPC is treated mainly by radiation therapy, with or without chemotherapy. Intensity-modulated radiation therapy has gradually become the standard treatment for NPC because of its superior tumor target conformity and significantly improved localization.[2,3] As the spatial accuracy of dose distribution improves, a more precise evaluation of the spatial extent of NPC is required.

MR imaging offers excellent soft-tissue contrast and is useful for the diagnosis, staging, and evaluation of tumor extent.[4-7] It is

Received June 20, 2018; accepted after revision October 22.

From the Department of Diagnostic Radiology (T.H., H.K., K.S., M.K., T.K.), National Cancer Center Hospital East, Kashiwa, Chiba, Japan; Canon Medical Systems Corporation (S.T.), Otawara, Tochigi, Japan; Departments of Radiology (O.S.), Otolaryngology–Head and Neck Surgery (O.S.), and Radiation Oncology (O.S.), Boston Medical Center, Boston University School of Medicine, Boston, Massachusetts; and Department of Diagnostic Radiology (M.K.), National Cancer Center Hospital, Chuo-ku, Tokyo, Japan.

This study was supported by a grant from the Japanese Ministry of Education, Culture, Sports, Science and Technology (Grant-in-Aid for Young Scientists [B] KAKEN; No. 26861033) and Canon Medical Systems.

Paper previously presented, in part, at Annual Meeting of the American Society of Neuroradiology and the Foundation of the ASNR Symposium, June 2–7, 2018; Vancouver, British Columbia, Canada.

Please address correspondence to Hirofumi Kuno, MD, PhD, Department of Diagnostic Radiology, National Cancer Center Hospital East, 6–5-one Kashiwanoha, Kashiwa, Chiba 277-8577, Japan; e-mail: hkuno@east.ncc.go.jp

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

Indicates article with supplemental on-line appendix.

Indicates article with supplemental on-line photos.

http://dx.doi.org/10.3174/ajnr.A5906

well-established that MR imaging is better than CT at identifying invasion of the skull base.[8-11] CT can detect skull base involvement manifesting as lytic or sclerotic lesions but has lower sensitivity than MR imaging due to lower tissue contrast and lower detectability of bone marrow invasion. However, CT is often the primary imaging technique for radiation therapy planning and dose computation and may be the only option for patients with contraindications to MR imaging such as intracranial or orbital metallic foreign bodies and MR imaging–unsafe pacemaker devices, or for patients who cannot tolerate MR imaging because of claustrophobia.

Currently, digital subtraction angiography-like bone subtraction iodine (BSI) imaging with 160-mm-wide area detector CT is used in some musculoskeletal and neurologic applications.[12-14] According to Gondim Teixeira et al,[12] BSI imaging with area detector CT showed high performance for the visual identification of bone marrow enhancement adjacent to lytic bone lesions. This technique reduces spatial mismatch by using volume scanning with wide area detector CT and a high-resolution deformable registration algorithm, enabling identification of contrast enhancement in the bone marrow. We hypothesized that BSI imaging might be useful for detecting skull base invasion and accurately assessing the extent of bone invasion by the NPC.

Hence, the purpose of this retrospective study was to compare the diagnostic performance of conventional CT (CCT) alone and in combination with BSI imaging using area detector CT for the evaluation of skull base invasion in patients with NPC.

Table 1: Summary of patient characteristics

Characteristics	No.	%
Age (yr)		
Mean	60	
Range	18–79	
Sex		
Female	16	36.0%
Male	28	64.0%
Histopathology		
Nonkeratinizing carcinoma	29	65.9%
Differentiated	8	
Undifferentiated	19	
Unknown	2	
Keratinizing carcinoma	8	18.2%
Unknown	7	15.9%
TNM (7th AJCC)		
T1	8	18.2%
T2	10	22.7%
T3	4	9.1%
T4	22	50.0%
N0	3	6.8%
N1	16	36.4%
N2	13	29.5%
N3a	1	2.3%
N3b	11	25.0%
M0	41	93.2%
M1	3	6.8%
Subsite		
Posterior superior	12	27.3%
Lateral wall	32	72.7%

Note:—TNM indicates tumor, node, metastasis tumor stage; AJCC, American Joint Committee on Cancer.

MATERIALS AND METHODS

Study Population

Our institutional review board at National Cancer Center Hospital East approved this retrospective study. The requirement to obtain written informed consent was waived. Between April 2012 and November 2017, forty-seven consecutive newly diagnosed patients with histologically proved NPC underwent contrast-enhanced CT for cancer staging before treatment. Among them, 46 patients also underwent contrast-enhanced 3T MR imaging before treatment. Two of the 46 patients were excluded because of unavailable volume data needed for subtraction reconstruction. The remaining 44 patients (28 men, 16 women; age range, 18–79 years; mean age, 60 years) were enrolled in this study. Of the 44 patients, 29 (65.9%) had nonkeratinizing squamous cell carcinoma and 8 (18.2%) had keratinizing squamous cell carcinoma. The detailed clinical profiles are summarized in Table 1. Tumor stage was determined according to the seventh edition of the American Joint Committee on Cancer staging system.[15]

Digital Subtraction Angiography-Like BSI Imaging Techniques

All CT studies were performed using a 320–detector row CT system with a detector width of 160 mm (Aquilion ONE Vision; Canon Medical Systems, Otawara, Japan), with the following parameters: 120 kV, 120 effective mAs (milliampere-second), 0.5-second rotation time, 160-mm collimation. The average CT dose index was 25.2 mGy. Patients received a 100-mL injection of 300 mg I/mL of iodinated contrast medium (iopamidol; Teva Takeda Yakuhin, Koka, Japan) at a rate of 2.5 mL/s into an antecubital vein through a 22-ga cannula. Scans were started 7 (mask volume) and 70 seconds (postcontrast volume) after the start of the injection in an intermittent acquisition mode with no table feed. During volume scanning, the patient's head was fixed to the bed of the CT with a headband and neck collar (Stifneck Select; Laerdal, Stavanger, Norway) (On-line Fig 1).

Two volume datasets (7- and 70-second CT) were reconstructed using a high-resolution deformable registration algorithm. The mask volume was subtracted from the postcontrast volume using the ^{SURE}Subtraction application (Canon Medical Systems). The high-resolution deformable registration algorithm is described in detail in the On-line Appendix and On-line Fig 2.[16] Axial and coronal CCT images with soft-tissue (window level 65, window width 330) and bone (window level 650, window width 3000) windows, and axial and coronal BSI images (window level 40, window width 180) were reconstructed with a 2-mm slice thickness, 512 × 512 matrix, and 16-cm FOV. The CCT and BSI images were reconstructed to yield precisely matching slices.

Image Interpretation

Two independent radiologists (with 12 years [H.K.] and 24 years [T.K.] of experience in oncologic diagnostic radiology), blinded to the patients' clinical histories and to the images from the other technique, independently analyzed the images. The CCT images and combined CCT-plus-BSI images were evaluated using a 5-point scale for calculating the area under the receiver operating characteristic curve (AUC) and weighted κ statistics. For CCT images, findings for skull base invasion were considered negative when the tumor was separated from the skull base (score 1, defi-

nitely negative) or was only in contact with the skull base (score 2, probably negative). Findings of skull base invasion were considered present when erosion or sclerosis (score 3, possibly positive), erosion and sclerosis (score 4, probably positive), or invasion into the bone marrow (score 5, definitely positive) was present. For combined analysis of CCT and BSI images, we scored skull base invasion using the BSI image in addition to the CCT image. If the BSI image showed definitive bone enhancement continuous with the tumor, we scored it as grade 5 (definitely positive); the absence of enhancement was scored as grade 2 (probably negative). If the bone marrow enhancement was equivocal or impossible to estimate, we gave scores of 2, three (possibly positive), or 4 (probably positive) based on the confidence level of enhancement on BSI images and findings on CCT. By means of a workstation (Shade Quest View R; Yokogawa Electric, Tokyo, Japan), the images were presented in random order in 2 sessions, initially with CCT alone (soft-tissue and bone window), followed 4 weeks later by a combination of CCT and BSI images. The readers were able to generate fusion images of the bone window and BSI images. Skull base invasion was evaluated at 6 sites on the skull base (sphenoid body, clivus, bilateral pterygoid process, and bilateral petrous apex). The final score for the skull base invasion was determined by a consensus between the 2 readers.

Standard of Reference

Owing to the difficulty in obtaining histopathologic confirmation of skull base invasion, a combination of MR imaging and CT features observed by an experienced neuroradiologist (T.H., 12 years of experience in neuro-/head and neck radiology) who was not involved in image interpretation in this study but made reports clinically in some cases was used as the reference standard for evaluating sensitivity, specificity, and the AUC for skull base invasion. The MR imaging protocol is discussed in the On-line Appendix. All patients underwent MR imaging and CT examinations within a median of 5 days (range, 0–23 days).

MR Imaging and CT Diagnostic Criteria for Skull Base Invasion

Findings of skull base invasion were considered positive in cases in which the skull base had signals similar to those of the adjacent tumor on all pre- and postcontrast T1-weighted images in the cortical bone and bone marrow space on the MR image (low signal intensity on a T1-weighted image and contrast enhancement on the postcontrast T1-weighted image with fat suppression).[8,9,17-21] We performed evaluations mainly on the basis of the precontrast 3D T1-weighted Dixon in-phase image and the postcontrast 3D T1-weighted Dixon water image. Other sequences were also used for identifying tumor location. CCT was used for the evaluation of bone cortex because CT is superior to MR imaging for the detection of bone cortex invasion. When CT and MR image findings were mismatched for bone marrow invasion, the diagnosis obtained on the basis of MR imaging was used.

Statistical Analysis

For estimating the sensitivity and specificity of both modalities for the detection of skull base invasion, we considered diagnostic confidence scores of >3 as positive diagnoses and generated contingency tables separately for each site on the skull base. For comparisons of sensitivity and specificity, the McNemar test and generalized estimating equations were used to take into account correlations between multiple sites in the same patient.[22] The overall diagnostic performance was quantified by the area under the receiver operating characteristic curve using a 5-point scale. Interreader agreement between the independent evaluations of the 2 readers for the scores based on the CCT-alone images and the CCT-plus-BSI images was estimated by weighted κ statistics (weight = 2). This was to account for the fact that differences of >1 point between the scores of the 2 readers were more important than smaller differences.[23] Commercial software (STATA, Version 12.1; StataCorp, College Station, Texas) was used. $P < .05$ was considered a significant difference.

RESULTS

Skull Base Invasion According to the Reference Standard Findings

All CCT and BSI images generated from the 2-volume datasets were considered of diagnostic image quality. In total, we examined 264 sites (44 in the sphenoid body, 44 in the clivus, 88 in the pterygoid process, and 88 in the petrous apex) from 44 patients. Of these 264 sites, skull base invasion was present at 84 sites (31.8%) according to the reference standard findings. A total of 26 patients (59.1%) showed skull base invasion. Bone sclerosis was observed in 21 (47.7%) of the 44 patients, and gross skull base invasion into the bone marrow was observed in 18 (40.9%) of 44 patients. Skull base invasion was present at 22/44 sphenoid body sites (50.0%), 20/44 clivus sites (45.5%), 10/44 right base of the pterygoid process sites (22.7%), 8/44 left base of the pterygoid process sites (18.2%), 12/44 right petrous bone sites (27.3%), and 12/44 left petrous bone sites (27.3%).

Differentiation of Imaging Findings between CCT Alone and CCT-Plus-BSI Imaging

The results of the image interpretation and diagnostic performance overall and at each site on the skull base for the CCT-alone and CCT-plus-BSI imaging are summarized in Table 2. Among the 84 skull base invasions, 66 (78.6%) were correctly detected by CCT alone and 18 (21.4%) produced false-negative findings (4 of the sphenoid body, 9 of the clivus, 1 base of the pterygoid process, and 4 of the petrous apex). With use of both CCT and BSI images, however, the number of false-negative findings decreased from 18 to 6 because the BSI images demonstrated iodine enhancement at the skull base without obvious cortical destruction. Figure 1 shows a representative case of a false-negative finding on CCT alone. On CCT images, extensive destruction or sclerosis of the skull base was unclear, but the BSI images clearly demonstrated iodine distribution within soft tissue and bone (both cortex and bone marrow space) continuous with the tumor. Seventy-four true-positive sites on CCT-plus-BSI had an intensity similar to that of the tumor on T2-weighted images. The remaining 4 cases showed hyperintensity on T2-weighted images compared with the primary tumor. Six (7.1%) false-negative diagnoses of 84 sites remained even after examination of BSI images, due to failure to detect visible iodine enhancement in the bone marrow continuous with the tumor. Among these 6 false-negative diagnoses, 4

Table 2: Comparison between CCT images alone and CCT-plus-BSI images of skull base invasion[a]

Parameter	TP[b]	TN[b]	FN[b]	FP[b]	Sensitivity (%)	*P* Value	Specificity (%)	*P* Value	PPV (%)	NPV (%)
All sites										
CCT alone	66	155	18	25	79 (68–87)	.016[c]	86 (80–90)	.010[c]	73	90
CCT-plus-BSI	78	172	6	8	93 (85–97)		96 (91–98)		91	97
Sphenoid body										
CCT alone	18	14	4	8	82 (60–95)	.625	64 (41–83)	.039[d]	69	78
CCT-plus-BSI	20	21	2	1	91 (71–99)		95 (77–100)		95	91
Clivus										
CCT alone	11	22	9	2	55 (32–77)	.031[d]	92 (73–99)	1.000	85	71
CCT-plus-BSI	17	22	3	2	85 (62–97)		92 (73–99)		89	88
Base of the pterygoid process										
CCT alone	17	60	1	10	94 (73–100)	.157	86 (75–93)	.012[c]	63	98
CCT-plus-BSI	18	68	0	2	100 (81–100)		97 (90–100)		90	100
Petrous apex										
CCT alone	20	59	4	5	83 (63–95)	.221	92 (83–97)	.306	80	94
CCT-plus-BSI	23	61	1	3	96 (79–100)		95 (87–99)		88	98

Note:—FN indicates false-negative findings; FP, false-positive findings; NPV, negative predictive value; PPV, positive predictive value; TN, true-negative findings; TP, true-positive findings.

[a] Numbers in parentheses are 95% confidence intervals.

[b] Data are numbers of findings.

[c] $P < .05$, according to the generalized estimating equations that accounted for the multiple observations within patients.

[d] $P < .05$, as determined with the McNemar test.

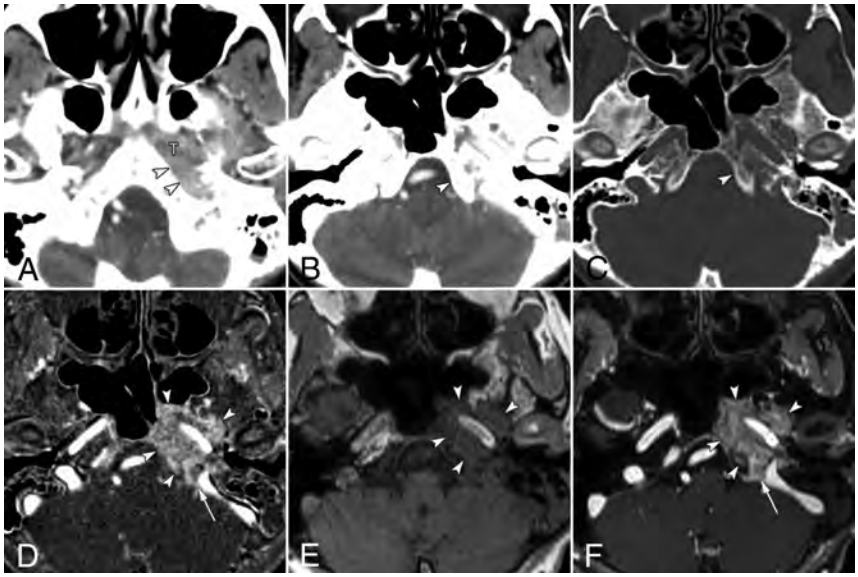


FIG 1. False-negative findings for skull base invasion by conventional CT images alone in a 74-year-old woman with nasopharyngeal carcinoma. *A*, Axial contrast-enhanced CCT image (soft-tissue window) shows nasopharyngeal tumor (T) spread into the lateral soft tissue around the foramen lacerum (*arrowheads*). CCT images at the skull base level (*B*, soft-tissue window; *C*, bone window) show no destruction of the skull base at the clivus (*arrowhead*). *D*, Bone subtraction iodine image shows remarkable skull base invasion into the bone marrow space such as the clivus, petrous apex, and sphenoid bone (*arrowheads*) with intracranial extension at the jugular foramen (*arrow*). A corresponding slice on a T1-weighted image (*E*) and fat-suppressed T1-weighted image after gadolinium administration (*F*) show tumor invading the clivus (*arrowheads*) and spread into the jugular foramen (*arrow*).

sites in 1 patient were not detected because the tumor showed a weak enhancement with density similar to that of normal bone marrow.

A total of 25 (13.9%) false-positive findings (8 of the sphenoid body, 2 of the clivus, 10 of the pterygoid process, and 5 of the petrous apex) of 180 sites without skull base invasion were observed on CCT-alone images, whereas 8 (4.4%) false-positive findings (1 sphenoid body, 2 of the clivus, 2 of the pterygoid process, and 3 of the petrous apex) were observed on CCT-plus-BSI images. The number of sites with false-positive findings decreased from 25 to 8 on CCT-plus-BSI imaging because the BSI images could detect an absence of corresponding enhancement in the sclerotic regions indicated on the CCT images. Figure 2 shows a case with a false-positive finding on CCT alone with sclerotic change. Two of 8 false-positive sites (1 of the clivus and 1 of the petrous apex) seen on CCT-plus-BSI were probably observed due to enhancement of normal red bone marrow in a young patient in her 30s, and 1 site (sphenoid body) was observed due to misidentifying noise as contrast enhancement. The other 5 false-positive findings seen on CCT-plus-BSI were caused by equivocal enhancement in the BSI images and were scored mainly on the basis of erosion or sclerosis observed on the CCT images.

Diagnostic Performance of CCT Alone and CCT-Plus-BSI Imaging

In an analysis of all sites taken together, the sensitivity of CCT-plus-BSI (93%; 95% CI, 85%–97%) for the detection of skull base invasion was significantly higher ($P = .02$) than that of CCT alone (79%; 95% CI, 68%–87%). If we split the analysis by site, for the clivus the sensitivity of CCT-plus-BSI was significantly higher ($P = .03$) (85%; 95% CI, 62%–97%) than that of CCT alone (55%; 95% CI, 32%–77%). For the other sites, the sensitivity for CCT-plus-BSI was higher than that for CCT alone, but none of the differences were significant: sphenoid body ($P = .63$; 91% for CCT-plus-BSI versus 82% for CCT-alone), pterygoid process ($P = .16$; 100% versus 94%), and petrous apex ($P = .22$; 96% versus 83%).

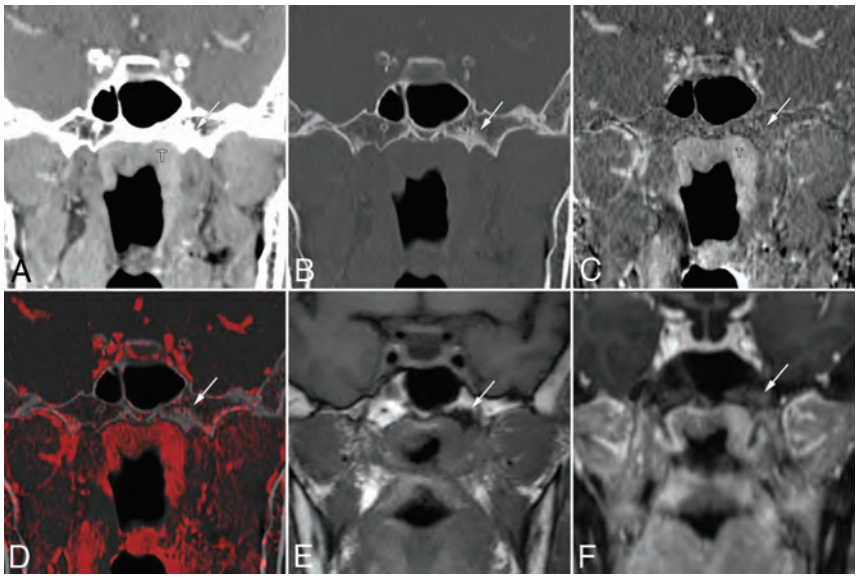


FIG 2. False-positive findings at the base of the left pterygoid process on conventional CT images alone in a 63-year-old woman with nasopharyngeal carcinoma. Coronal contrast-enhanced CCT images (*A*, soft-tissue window; *B*, bone window) show nasopharyngeal tumor (T) and bone sclerosis at the left base of the pterygoid process (*arrow*). The case was given a score of 3 based on CCT images alone. Bone subtraction iodine image (*C*) and color fusion image (*D*) clearly show no enhancement in the sclerotic area (*arrow*). *E*, A corresponding slice on the T1-weighted image shows low signal intensity due to sclerosis (*arrow*). *F*, Fat-suppressed T1-weighted images after gadolinium administration show contrast enhancement of the tumor mass and poor enhancement of the pterygoid process (*arrow*). This was considered a case of clinical T2 category disease without skull base invasion and was treated by chemoradiotherapy.

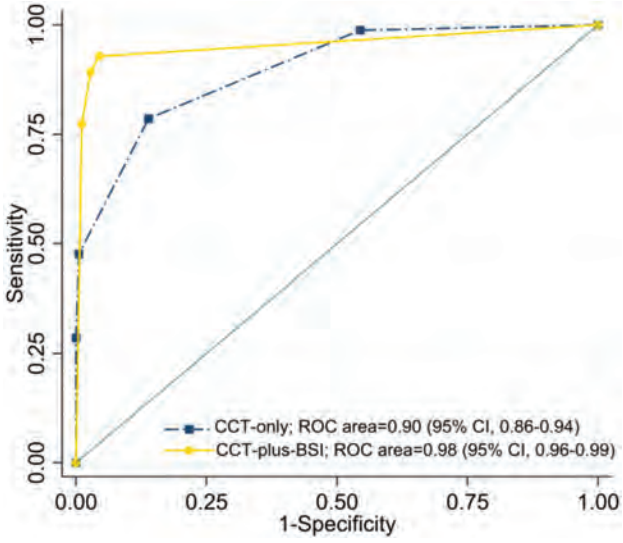


FIG 3. Receiver operating characteristic curves and corresponding areas under the curve for the prediction of skull base invasion. The AUC for CCT-plus-BSI imaging was significantly larger (AUC = 0.98 [$P < .001$]) than that for CCT imaging alone (AUC = 0.90).

The specificity of CCT-plus-BSI (96%; 95% CI, 91%–98%) for the detection of skull base invasion was significantly higher ($P = .01$) than that of CCT alone (86%; 95% CI, 80%–90%) in the analysis of all sites taken together. When we considered each site separately, the specificity of sphenoid body and pterygoid process detection was significantly higher ($P = .04$ and $P = .012$, respectively) in the case of CCT-plus-BSI (sphenoid body: 95%; 95% CI, 77%–100%; pterygoid process: 97%; 95% CI, 90%–100%) than that of CCT alone (sphenoid body: 64%; 95% CI, 41%–83%; pterygoid process: 86%; 95% CI, 75%–93%). The specificity of CCT-plus-BSI was the same or higher than that of CCT alone at the other sites as well, but none of the differences were significant (clivus: $P = 1.00$, 92% for CCT-plus-BSI versus 92% for CCT-alone; petrous apex: $P = .31$, 95% versus 92%).

The receiver operating characteristic curves for skull base invasion and the corresponding AUCs for diagnosis based on CCT-plus-BSI and CCT-alone are shown in Fig 3. The AUC for CCT-plus-BSI was significantly larger (AUC = 0.98, $P < .001$) than the AUC for CCT alone (AUC = 0.90).

Interobserver Reproducibility

The interreader agreement of the 2 readers for CCT alone and for CCT-plus-BSI was assessed with quadratic weighted statistics, with the following results: agreement of 0.82 (95% CI, 0.53–1.00) and 0.86 (95% CI, 0.56–1.00), respectively, for the sphenoid body; agreement of 0.84 (95% CI, 0.55–1.00) and 0.83 (95% CI, 0.54–1.00), respectively, for the clivus; agreement of 0.85 (95% CI, 0.57–1.00) and 0.86 (95% CI, 0.58–1.00), respectively, for the right base of the pterygoid process; agreement of 0.82 (95% CI, 0.55–1.00) and 0.81 (95% CI, 0.53–1.00), respectively, for the left base of the pterygoid process; agreement of 0.81 (95% CI, 0.52–1.00) and 0.89 (95% CI, 0.59–1.00), respectively, for the right petrous apex; agreement of 0.85 (95% CI, 0.57–1.00) and 0.86 (95% CI, 0.58–1.00), respectively, for the left petrous apex.

DISCUSSION

The results of our study demonstrate that the combined analysis of CCT and BSI images showed a better diagnostic performance than the analysis of CCT images alone in the evaluation of skull base invasion in patients with NPC. These results suggest that CCT-plus-BSI images potentially allow an assessment closely resembling that provided by combining CT and contrast-enhanced MR imaging. High sensitivity could be achieved because CCT-plus-BSI improves the detection of tumor spread within the skull base (cortical bone and bone marrow), especially in the clivus. Another factor contributing to the improved specificity was that CCT-plus-BSI reduced the overestimation caused by sclerotic parts such as the sphenoid body and pterygoid process, which may reflect the bone response to tumor proximity. The differences in the diagnostic performance of each site may be due to the amount of bone marrow, the thickness of the bone affected by pneumatization, and the frequency of sclerosis. These results may be particularly useful for improving the accuracy for radiation therapy treatment planning and primary tumor staging. Further-

more, CCT-plus-BSI imaging offers a benefit for patients with contraindications to MR imaging due to metallic foreign bodies, MR imaging–unsafe pacemaker devices, or claustrophobia.

CT is an excellent technique for the evaluation of bone details, particularly for the evaluation of cortical bone invasion. However, it is still challenging to assess tumor extent in the bone marrow using CCT imaging without contrast-enhanced MR imaging because of low contrast resolution and beam-hardening artifacts produced by the bone cortex. MR imaging is superior to CT in the detection of skull base involvement.[8-11] Several studies have used MR imaging to examine abnormalities of the skull base in patients with NPC.[17-21,24] Lu et al[20] reported worse prognosis when ≥2 sites of skull base invasion were found on MR imaging. Other groups found that the prognosis varied depending on the sites of skull base invasion.[17,21,24] Therefore, from the perspective of prognosis, comprehensive assessment is needed, including the sites and spatial extent within the skull base. BSI images have the potential to detect enhancement in the bone marrow and may complement CCT in assessing bone invasion.

Shatzkes et al[25] reported that sclerotic findings in the pterygoid process are detected on CT in 60% of cases of untreated NPC. These findings may indicate present or imminent skull base invasion and may be used for tumor staging and the formulation of intensity-modulated radiotherapy fields. These sclerotic changes may be a sensitive indicator of skull base invasion, but they are not specific because sclerosis can also be caused by inflammatory or reactive changes.[26,27] Recently, there have been several reports of the use of dual-energy CT for the diagnosis of bone marrow edema in patients with bone fractures[28-30] and for the detection of bone marrow involvement in patients with multiple myeloma.[31] However, with current dual-energy CT systems and material decomposition algorithm, dual-energy CT technology cannot clearly distinguish bone from iodinated contrast.[32,33] Thus, calcium (bone) remains visible on the iodine map. BSI images generated by area detector CT allow bone removal (both cortical and trabecular) without affecting the visualization of contrast enhancement in a nonlytic bone background.[12-14] BSI imaging may be helpful for the diagnosis of bone invasion with or without erosive changes in cortical bone.

There are limitations to our study. The first is the absence of histopathologic verification of the findings of skull base invasion because the difficulty in accessing the skull base precludes surgical confirmation and almost all patients were treated with chemoradiotherapy. According to a study of MR imaging findings for laryngeal cartilage invasion, contrast enhancement is also caused by reactive inflammation, edema, and fibrosis in close proximity to the tumor.[34] Therefore, there is no clear consensus on whether the contrast enhancement on BSI images was caused by tumor invasion or inflammation. These issues need to be examined at other primary sites such as the mandibular bone or maxillary bone in future studies. The second potential limitation is that skull base invasion was assessed only visually, without quantitative measurement of iodine concentrations. Therefore, potentially confounding effects of interreader error, including display settings, may be present. Moreover, the degree by which tumor and bone marrow enhancement influence the quality and diagnostic performance of BSI images remains unclear and needs to be evaluated in future studies.

CONCLUSIONS

The combined analysis of CCT and BSI images generated from 320–row area detector CT has the potential to improve diagnostic performance in the evaluation of skull base invasion by NPC. BSI images can provide additional contrast resolution for bone assessment. Although there is no direct comparison between the underlying tumor histopathology and the results of BSI imaging, CCT-plus-BSI imaging may be useful for accurate staging of the tumor and may be helpful for patients with contraindications for MR imaging.

Disclosures: Takashi Hiyama—*RELATED*: *Consulting Fee or Honorarium*: Canon Medical Systems.* Hirofumi Kuno—*RELATED*: *Grant*: Grant-in-Aid for Young Scientists (B) KAKEN (No. 26861033); *Consulting Fee or Honorarium*: Canon Medical Systems*; *UNRELATED*: *Grants/Grants Pending*: Grant-in-Aid for Young Scientists KAKEN (No. 18K15573). So Tsushima—*UNRELATED: Employment*: Canon Medical Systems. Osamu Sakai—*UNRELATED*: *Consultancy:* Boston Imaging Core Lab. Masahiko Kusumoto—*RELATED*: *Consulting Fee or Honorarium*: Canon Medical Systems.* Tatsushi Kobayashi—*RELATED*: *Consulting Fee or Honorarium*: Canon Medical Systems*; *UNRELATED: Grants/Grants Pending*: Japan Agency for Medical Research and Development, Grant No. 17ck0106343h0001.* *Money paid to the institution.

REFERENCES

1. Amin MB, Edge SB, Greene FL, et al. *AJCC Cancer Staging Manual.* 8th ed. New York: Springer-Verlag; 2017
2. Peng G, Wang T, Yang KY, et al. **A prospective, randomized study comparing outcomes and toxicities of intensity-modulated radiotherapy vs. conventional two-dimensional radiotherapy for the treatment of nasopharyngeal carcinoma.** *Radiother Oncol* 2012;104:286–93 CrossRef Medline
3. Lee AW, Ng WT, Chan LL, et al. **Evolution of treatment for nasopharyngeal cancer: success and setback in the intensity-modulated radiotherapy era.** *Radiother Oncol* 2014;110:377–84 CrossRef Medline
4. Mukherji SK, Pillsbury HR, Castillo M. **Imaging squamous cell carcinomas of the upper aerodigestive tract: what clinicians need to know.** *Radiology* 1997;205:629–46 CrossRef Medline
5. Dubrulle F, Souillard R, Hermans R. **Extension patterns of nasopharyngeal carcinoma.** *Eur Radiol* 2007;17:2622–30 CrossRef Medline
6. King AD, Bhatia KS. **Magnetic resonance imaging staging of nasopharyngeal carcinoma in the head and neck.** *World J Radiol* 2010;2:159–65 CrossRef Medline
7. Abdel Razek AA, King A. **MRI and CT of nasopharyngeal carcinoma.** *AJR Am J Roentgenol* 2012;198:11–18 CrossRef Medline
8. Chong VF, Fan YF. **Skull base erosion in nasopharyngeal carcinoma: detection by CT and MRI.** *Clin Radiol* 1996;51:625–31 CrossRef Medline
9. Zhang SX, Han PH, Zhang GQ, et al. **Comparison of SPECT/CT, MRI and CT in diagnosis of skull base bone invasion in nasopharyngeal carcinoma.** *Biomed Mater Eng* 2014;24:1117–24 CrossRef Medline
10. Ng SH, Chang TC, Ko SF, et al. **Nasopharyngeal carcinoma: MRI and CT assessment.** *Neuroradiology* 1997;39:741–46 CrossRef Medline
11. Liao XB, Mao YP, Liu LZ, et al. **How does magnetic resonance imaging influence staging according to AJCC staging system for nasopharyngeal carcinoma compared with computed tomography?** *Int J Radiat Oncol Biol Phys* 2008;72:1368–77 CrossRef Medline
12. Gondim Teixeira PA, Hossu G, Lecocq S, et al. **Bone marrow edema pattern identification in patients with lytic bone lesions using digital subtraction angiography-like bone subtraction on large-area detector computed tomography.** *Invest Radiol* 2014;49:156–64 CrossRef Medline
13. Gondim Teixeira PA, Gervaise A, Louis M, et al. **Musculoskeletal wide detector CT: principles, techniques and applications in clini-**

cal practice and research. *Eur J Radiol* 2015;84:892–900 CrossRef Medline
14. Kuno H, Sekiya K, Chapman MN, et al. **Miscellaneous and emerging applications of dual-energy computed tomography for the evaluation of intracranial pathology.** *Neuroimaging Clin N Am* 2017;27: 411–27 CrossRef Medline
15. Edge SE, Byrd DR, Compton CC, et al, eds. *AJCC Cancer Staging Manual*. 7th ed. New York: Springer-Verlag; 2009
16. Crum WR, Hill DL, Hawkes DJ. **Information theoretic similarity measures in non-rigid registration.** *Inf Process Med Imaging* 2003;18: 378–87 Medline
17. Chen L, Liu LZ, Mao YP, et al. **Grading of MRI-detected skull-base invasion in nasopharyngeal carcinoma and its prognostic value.** *Head Neck* 2011;33:1309–14 CrossRef Medline
18. Li YZ, Cai PQ, Xie CM, et al. **Nasopharyngeal cancer: impact of skull base invasion on patients prognosis and its potential implications on TNM staging.** *Eur J Radiol* 2013;82:e107–11 CrossRef Medline
19. Nishioka T, Shirato H, Kagei K, et al. **Skull-base invasion of nasopharyngeal carcinoma: magnetic resonance imaging findings and therapeutic implications.** *Int J Radiat Oncol Biol Phys* 2000;47:395–400 CrossRef Medline
20. Lu JC, Wei Q, Zhang YQ, et al. **Influence of MRI abnormality in skull base bone on prognosis of nasopharyngeal carcinoma.** *Cancer Radiother* 2004;8:230–33 CrossRef Medline
21. Wang H, Zhang ZQ, Huang LL, et al. **MR imaging prediction of local control of nasopharyngeal carcinoma treated with radiation therapy and chemotherapy.** *Br J Radiol* 2014;87:20130657 CrossRef Medline
22. Genders TS, Spronk S, Stijnen T, et al. **Methods for calculating sensitivity and specificity of clustered data: a tutorial.** *Radiology* 2012; 265:910–16 CrossRef Medline
23. Sim J, Wright CC. **The kappa statistic in reliability studies: use, interpretation, and sample size requirements.** *Phys Ther* 2005;85: 257–68 Medline
24. Roh JL, Sung MW, Kim KH, et al. **Nasopharyngeal carcinoma with skull base invasion: a necessity of staging subdivision.** *Am J Otolaryngol* 2004;25:26–32 CrossRef Medline
25. Shatzkes DR, Meltzer DE, Lee JA, et al. **Sclerosis of the pterygoid process in untreated patients with nasopharyngeal carcinoma.** *Radiology* 2006;239:181–86 CrossRef Medline
26. Becker M, Zbären P, Delavelle J, et al. **Neoplastic invasion of the laryngeal cartilage: reassessment of criteria for diagnosis at CT.** *Radiology* 1997;203:521–32 CrossRef Medline
27. Muñoz A, Ramos A, Ferrando J, et al. **Laryngeal carcinoma: sclerotic appearance of the cricoid and arytenoid cartilage—CT-pathologic correlation.** *Radiology* 1993;189:433–37 CrossRef Medline
28. Kaup M, Wichmann JL, Scholtz JE, et al. **Dual-energy CT-based display of bone marrow edema in osteoporotic vertebral compression fractures: impact on diagnostic accuracy of radiologists with varying levels of experience in correlation to MR imaging.** *Radiology* 2016;280:510–19 CrossRef Medline
29. Pache G, Krauss B, Strohm P, et al. **Dual-energy CT virtual noncalcium technique: detecting posttraumatic bone marrow lesions—feasibility study.** *Radiology* 2010;256:617–24 CrossRef Medline
30. Kellock TT, Nicolaou S, Kim SSY, et al. **Detection of bone marrow edema in nondisplaced hip fractures: utility of a virtual noncalcium dual-energy CT application.** *Radiology* 2017;284:798–805 CrossRef Medline
31. Thomas C, Schabel C, Krauss B, et al. **Dual-energy CT: virtual calcium subtraction for assessment of bone marrow involvement of the spine in multiple myeloma.** *AJR Am J Roentgenol* 2015;204: W324–31 CrossRef Medline
32. Gupta R, Phan CM, Leidecker C, et al. **Evaluation of dual-energy CT for differentiating intracerebral hemorrhage from iodinated contrast material staining.** *Radiology* 2010;257:205–11 CrossRef Medline
33. Kuno H, Onaya H, Iwata R, et al. **Evaluation of cartilage invasion by laryngeal and hypopharyngeal squamous cell carcinoma with dual-energy CT.** *Radiology* 2012;265:488–96 CrossRef Medline
34. Becker M, Zbären P, Casselman JW, et al. **Neoplastic invasion of laryngeal cartilage: reassessment of criteria for diagnosis at MR imaging.** *Radiology* 2008;249:551–59 CrossRef Medline

Solitary Parathyroid Adenoma Localization in Technetium Tc99m Sestamibi SPECT and Multiphase Multidetector 4D CT

T.H. Vu, D. Schellingerhout, N. Guha-Thakurta, J. Sun, W. Wei, S.C. Kappadth, N. Perrier, E.E. Kim, E. Rohren, H.H. Chuang, and F.C. Wong

ABSTRACT

BACKGROUND AND PURPOSE: Minimally invasive parathyroid surgery relies critically on image guidance, but data comparing the efficacy of various imaging modalities are scarce. Our aim was to perform a blinded comparison of the localizing capability of technetium Tc99m sestamibi SPECT, multiphase multidetector 4D CT, and the combination of these 2 modalities (technetium Tc99m sestamibi SPECT + multiphase multidetector 4D CT).

MATERIALS AND METHODS: We reviewed the records of 31 (6 men, 25 women; median age, 56 years) consecutive patients diagnosed with biochemically confirmed primary hyperparathyroidism between November 2009 and March 2010 who underwent preoperative technetium Tc99m sestamibi SPECT and multiphase multidetector 4D CT performed on the same scanner with pathologic confirmation by resection of a single parathyroid adenoma. Accuracy was determined separately for localization to the correct side and quadrant using surgical localization as the standard of reference.

RESULTS: Surgical resection identified 14 left and 17 right parathyroid adenomas and 2 left inferior, 12 left superior, 11 right inferior, and 6 right superior parathyroid adenomas. For left/right localization, technetium Tc99m sestamibi SPECT achieved an accuracy of 93.5% (29 of 31), multiphase multidetector 4D CT achieved 96.8% accuracy (30 of 31), and technetium Tc99m sestamibi SPECT + multiphase multidetector 4D CT achieved 96.8% accuracy (30 of 31). For quadrant localization, technetium Tc99m sestamibi SPECT accuracy was 67.7% (21 of 31), multiphase multidetector 4D CT accuracy was 87.1% (27 of 31), and technetium Tc99m sestamibi SPECT + multiphase multidetector 4D CT accuracy was 93.5% (29 of 31). Reader diagnostic confidence was consistently ranked lowest for technetium Tc99m sestamibi SPECT and highest for technetium Tc99m sestamibi SPECT + multiphase multidetector 4D CT.

CONCLUSIONS: For left/right localization of parathyroid adenomas, all modalities performed equivalently. For quadrant localization, technetium Tc99m sestamibi SPECT + multiphase multidetector 4D CT is superior to technetium Tc99m sestamibi SPECT.

ABBREVIATIONS: 4DCT = multiphase multidetector 4D CT; MIBI = technetium Tc99m sestamibi

Primary hyperparathyroidism is a common endocrine disorder caused most often by a solitary parathyroid adenoma.[1] The definitive cure for this disorder is surgical resection. The surgical approach has shifted from standard bilateral cervical exploration to modern minimally invasive parathyroidectomy.[2] For minimally invasive parathyroidectomy to be successful, accurate preoperative localization of a single offending parathyroid gland and exclusion of possible multigland disease or 4-gland hyperplasia are critical and useful. Several noninvasive preoperative localization modalities are available, including parathyroid scintigraphy using technetium Tc99m sestamibi (MIBI)[3,4]; ultrasonography[5,6]; CT, specifically multiphase multidetector 4D CT (4DCT)[7]; MR imaging[8]; and recently, ^{11}C-methionine positron-emission tomography/CT[9] and [^{18}F]-choline positron-emission tomography/CT.[10] The emergence of hybrid SPECT camera technology and advances in CT technology have allowed MIBI SPECT and 4DCT to be successfully used for the preoperative localization of abnormal parathyroid glands.[11-16]

The requirements for minimally invasive neck surgery have led to greater demands from imaging. Imaging needs to confidently localize the adenoma to a side, localize the adenoma to a quadrant or precise anatomic location, characterize the embryo-

Received May 20, 2018; accepted after revision October 14.

From the Departments of Diagnostic Radiology (T.H.V., D.S., N.G.-T.), Biostatistics (J.S., W.W.), Imaging Physics (S.C.K.), Surgical Oncology (N.P.), and Nuclear Medicine (H.H.C., F.C.W.), The University of Texas MD Anderson Cancer Center, Houston, Texas; Department of Radiological Sciences (E.E.K.), University of California at Irvine, Orange, California; and Baylor College of Medicine (E.R.), Houston, Texas.

Thinh H. Vu and Dawid Schellingerhout contributed equally to this work.

Please address correspondence to Thinh H. Vu, MD, Department of Diagnostic Imaging, Unit 1482, The University of Texas MD Anderson Cancer Center, 1400 Pressler St, FCT 16.5036, Houston, TX 77030; e-mail: thinh.vu@mdanderson.org

http://dx.doi.org/10.3174/ajnr.A5901

logic origin of the parathyroid lesion, be sensitive to ectopic disease, and preferably be able to identify potential multigland disease. Some traditional imaging studies may not meet these modern requirements.

The purpose of the present study was to do a blinded comparison of the localizing capability of MIBI SPECT, 4DCT, and the combination of these 2 modalities using surgical localization as a criterion standard. We evaluated the ability of each technique to localize a parathyroid adenoma to a side and a quadrant, with particular note of reader confidence in the imaging findings to direct surgery.

MATERIALS AND METHODS

Study Population

This was a retrospective study of consecutive patients who were diagnosed with primary hyperparathyroidism between November 2009 and March 2010. All patients had biochemically confirmed primary hyperparathyroidism (a serum calcium level of >10.4 mg/dL [2.6 mmol/L] and a serum parathyroid hormone level of >65 pg/mL [6.84 pmol/L]); underwent preoperative MIBI SPECT and 4DCT performed on same scanner; and had pathologic confirmation by resection of a single hypercellular parathyroid gland. Patients who had undergone a prior operation or had multigland resections were excluded. We obtained a waiver of consent from our institutional review board (MD Anderson Cancer Center, Houston, Texas) for this Health Insurance Portability and Accountability Act–compliant study.

Image Acquisition

MIBI SPECT. All patients underwent a dual-phase MIBI study after intravenous injection of 20–25 mCi (740–925 MBq) of technetium Tc99m methoxyisobutyl isonitrile. Anterior and posterior planar images of the head, neck, and thorax were acquired at 20 and 90 minutes, and SPECT was acquired at 30–60 minutes after intravenous injection of technetium Tc99m methoxyisobutyl isonitrile. Images were acquired using a high-resolution, low-energy, parallel-hole collimator and a large FOV, dual-headed gamma camera with a jointly mounted 16-slice CT scanner (Symbia T16; Siemens, Erlangen, Germany). SPECT was acquired for 128 frames over a full 360° arc at 22 seconds per frame on a noncircular orbit mapped to the body contour. The gamma camera photopeak window was centered at 140 keV with a 15% window and an adjacent 15% scatter window at lower energy. The mean count per frame was 196,000 with a range of 136,000–256,000. A noncontrast low-dose CT scan (130 kV[peak], 90 mAs, pitch = 1.2, CareDose4D, nominal volume CT dose index = 9.7 mGy) was obtained immediately after SPECT acquisition for attenuation correction. Image reconstruction was performed using 3D-ordered subset expectation maximization (OSEM) (Flash3D; Siemens) with 8 iterations and 16 subsets and a 5-mm Gaussian postreconstruction filter to yield a 128 × 128 × 128 matrix with 4.8 × 4.8 × 4.8 mm voxels. CT-based attenuation correction, energy-window–based scatter correction, and collimator resolution modeling were used during SPECT reconstructions.

4DCT. The 4DCT study was performed on the same table immediately after the MIBI SPECT study using a 16-row multidetector CT scanner of the SPECT system (Symbia T16; Siemens). Helical scans were obtained at 220 mA and 130 kVp. The scans were obtained with 1.25-mm collimation, table speed of 13.75 mm/s, pitch of 1.375, and gantry rotation time of 1 second. Initially, unenhanced scans were obtained from the carina to the mandible. An 18-ga intravenous cannula placed in the antecubital vein was used to inject 120 mL of nonionic contrast material, ioversol (Optiray 320; Mallinckrodt, St. Louis, Missouri), at 4.0 mL/s with a power injector. Multiphase scanning was performed with the second, third, and fourth scans (programmed to the same collimation, table feed, and duration as the first scan), which were obtained at 25 (arterial), 55 (venous), and 85 (delay) seconds, respectively, after the beginning of the administration of contrast.

Surgical Localization

The criterion standard for final anatomic localization of the parathyroid adenoma was an operation, as recorded in the operative notes.

Imaging Performance Analysis

Retrospective image analysis was performed on PACS workstations in configurations used for routine clinical image interpretation by 2 independent and blinded teams of readers. For each patient in the cohort, the MIBI SPECT was read independently using noncontrast CT and MIBI SPECT data only by 2 experienced nuclear physicians (E.E.K. and H.H.C.). The 4DCT data were read using both nonenhanced and contrast-enhanced CT data, as independently reviewed by 2 experienced neuroradiologists (D.S. and N.G.-T.). The combined MIBI SPECT and 4DCT (MIBI SPECT + 4DCT) data (all imaging information) were reviewed by a reader team consisting of an experienced neuroradiologist (T.H.V.) and a nuclear physician (E.R.). The team interpreted de-identified image sets knowing only the diagnosis of hyperparathyroidism, and readers were asked to localize the parathyroid adenoma. The localization to side (left or right) and quadrant (relative to the midpoint of the thyroid as dividing the upper and lower quadrants) and the embryologic origin of the abnormal parathyroid gland (superior or inferior) were recorded, as well as surgical localization based on surgical anatomy (Fig 1).[17] For each localization, reader confidence (certain, equivocal, or uncertain on a 3-point scale) was also recorded. The reader team had no knowledge of the imaging data outside their assigned area. The images were interpreted as a consensus reading without individual readings to measure interrater agreement.

After all imaging analysis was performed, 1 researcher (F.C.W.) consulted the surgical and pathologic records to determine adenoma localization, pathologic characteristics, and weight.

Statistical Analysis

Error matrices were constructed for overall true-versus-false localization for each of the readings (MIBI SPECT, 4DCT, and MIBI SPECT + 4DCT) against surgical localization. Error matrices were then constructed by level of confidence and analyzed similarly. A 95% confidence interval for accuracy was calculated using the Clopper-Pearson exact method. Diagnostic accuracy was compared among MIBI SPECT, 4DCT, and MIBI SPECT + 4DCT and determined by the McNemar test. All tests were 2-sided, and P values ≤ .05 were considered statistically signifi-

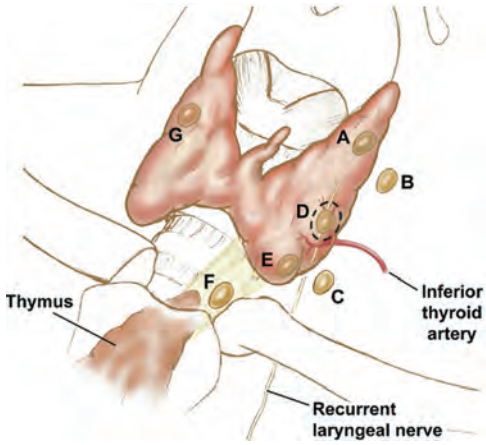


FIG 1. Surgical classification of parathyroid adenoma locations, anterior view. *A*, Superior, in proximity to the posterior surface of the thyroid parenchyma. *B*, Superior, fallen posteriorly into the tracheoesophageal groove and no longer in contact with the posterior surface of the thyroid tissue. *C*, Superior, fallen posteriorly into the tracheoesophageal groove and no longer in contact with the posterior surface of the thyroid tissue at the inferior pole close to the clavicles. *D*, Superior or inferior, in the midregion of the posterior surface of the thyroid parenchyma near the junction of the recurrent laryngeal nerve and the inferior thyroidal artery. *E*, Inferior, in the region inferior to the thyroid gland, lying in the anteroposterior plane of the thyroid and anterior to the trachea. *F*, Inferior, descended into the thyrothymic ligament or superior thymus and possibly appearing to be "ectopic" or in the mediastinum. *G*, Intrathyroidal.

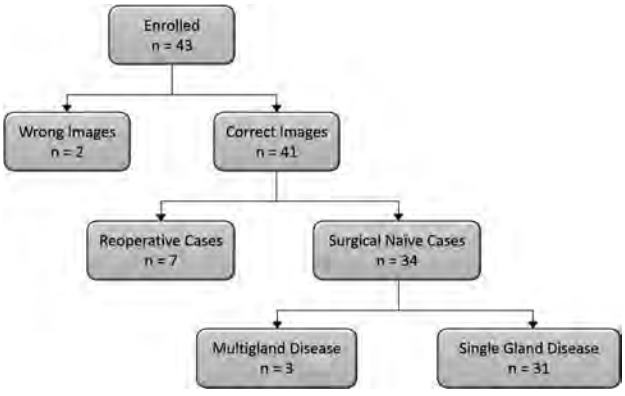


FIG 2. Patient selection flowchart. Thirty-one patients were included in the final analysis.

cant. Statistical analysis was performed using SAS, Version 9.4 (SAS Institute, Cary, North Carolina) and R version 2.3 (R Development Core Team, http://www.r-project.org).

RESULTS

Patient selection is summarized in Fig 2. Forty-three patients were included in our initial study group. Seven were excluded because of a prior operation, and 2 were excluded owing to incomplete imaging data, leaving 34 patients. Three patients had multiple parathyroid adenomas resected and were excluded. The remaining 31 patients had a single abnormal parathyroid gland resected.

The study group consisted of 6 male and 25 female patients with a median age of 56 years (range, 26–78 years). Patients had mean corrected calcium values of 10.6 mmol/L (median, 10.6 mmol/L; range, 9.5–11.8 mmol/L) against a reference range of 8.4–10.2 mmol/L at our laboratory. Patients had mean parathyroid hormone levels of 111 pmol/L (median, 105 pmol/L; range, 56–209 pmol/L) against a reference range of 9–80 pmol/L at our laboratory. The resected parathyroid glands had weights ranging from 0.07 to 6.08 g; the mean weight was 0.77 g and the median weight was 0.40 g.

Table 1: Diagnostic accuracy of left/right localization of parathyroid adenomas in the 31 patients in our study group

Imaging Modality	Left No. (%)	Right No. (%)	Total No. (%)	Accuracy (95% CI)
MIBI SPECT				93.5 (78.6–99.2)
Left	13 (41.9)	1 (3.2)	14 (45.2)	
Right	1 (3.2)	16 (51.6)	17 (54.8)	
4DCT				96.8 (83.3–99.9)
Left	14 (45.2)	1 (3.2)	15 (48.4)	
Right	0 (0.0)	16 (51.6)	16 (51.6)	
MIBI SPECT + 4DCT				96.8 (83.3–99.9)
Left	14 (45.2)	1 (3.2)	15 (48.4)	
Right	0 (0.0)	16 (51.6)	16 (51.6)	
Total	14 (45.2)	17 (54.8)		

Table 2: Diagnostic accuracy of embryologic origin of the abnormal parathyroid gland in the 31 patients in our study group

Imaging Modality	Superior No. (%)	Inferior No. (%)	Total No. (%)	Accuracy (95% CI)
MIBI SPECT				74.2 (55.4–88.1)
Superior gland	10 (32.3)	0 (0)	10 (32.3)	
Inferior gland	8 (25.8)	13 (41.9)	21 (67.7)	
Total	18 (58.1)	13 (41.9)		
4DCT				90.3 (74.3–98.0)
Superior gland	15 (48.4)	0 (0)	15 (48.6)	
Inferior gland	3 (9.7)	13 (41.9)	16 (51.6)	
Total	18 (58.1)	13 (41.9)		
MIBI SPECT + 4DCT				96.8 (83.3–99.9)
Superior gland	17 (54.8)	0 (0)	17 (54.8)	
Inferior gland	1 (3.2)	13 (41.9)	14 (45.2)	
Total	18 (58.1)	13 (41.9)		

Surgical Localization

Surgical resection localized parathyroid adenomas to the following locations: by side, 14 left and 17 right; by quadrant, 2 left inferior, 12 left superior, 11 right inferior, and 6 right superior; and by surgical classification (Fig 1), 9 type A, 9 type B, 3 type C, 1 type D, and 9 type E.

Overall Imaging Performance

Tables 1–4 summarize the error matrices for left- or right-sided localization, embryologic origin of the abnormal parathyroid gland (superior or inferior), quadrant localization (both left or right and upper or lower), and surgical classification of MIBI SPECT, 4DCT, and MIBI SPECT + 4DCT readings, along with accuracies.

MIBI SPECT achieved an accuracy of 93.5% (29 of 31 patients) in correctly localizing the parathyroid adenomas on the left or right side. Embryologic origin accuracy was 74.2% (23 of 31 patients). Quadrant localization accuracy was 67.7% (21 of 31 patients). Surgical classification accuracy was 54.8% (17 of 31 patients).

4DCT achieved an accuracy of 96.8% (30 of 31 patients) in correctly localizing parathyroid adenomas on the left or right side.

Table 3: Diagnostic accuracy of quadrant localization of parathyroid adenomas in the 31 patients in our study group

Imaging Modality	LI No. (%)	LS No. (%)	RI No. (%)	RS No. (%)	Total No. (%)	Accuracy (95% CI)
MIBI SPECT						67.7 (48.6–83.3)
LI	2 (6.5)	2 (6.5)	1 (3.2)	0 (0)	5 (16.1)	
LS	0 (0)	9 (29.0)	0 (0)	0 (0)	9 (29.0)	
RI	0 (0)	0 (0)	10 (32.3)	6 (19.4)	16 (51.6)	
RS	0 (0)	1 (3.2)	0 (0)	0 (0)	1 (3.2)	
4DCT						87.1 (70.2–96.4)
LI	2 (6.5)	1 (3.2)	1 (3.2)	0 (0)	4 (12.9)	
LS	0 (0)	11 (35.5)	0 (0)	0 (0)	11 (35.5)	
RI	0 (0)	0 (0)	10 (32.3)	2 (6.5)	12 (38.7)	
RS	0 (0)	0 (0)	0 (0)	4 (12.9)	4 (12.9)	
MIBI SPECT + 4DCT						93.5 (78.6–99.2)
LI	2 (6.5)	0 (0)	1 (3.2)	0 (0)	3 (9.7)	
LS	0 (0)	12 (38.7)	0 (0)	0 (0)	12 (38.7)	
RI	0 (0)	0 (0)	10 (32.3)	1 (3.2)	11 (35.5)	
RS	0 (0)	0 (0)	0 (0)	5 (16.1)	5 (16.1)	
Total	2 (6.5)	12 (38.7)	11 (35.5)	6 (19.4)		

Note:—LI indicates left inferior; LS, left superior; RI, right inferior; RS, right superior.

Table 4: Diagnostic accuracy of surgical classification (Fig 1) of parathyroid adenomas in the 31 patients in our study group

Imaging Modality	A No. (%)	B No. (%)	C No. (%)	D No. (%)	E No. (%)	Total No. (%)	Accuracy (95% CI)
MIBI SPECT							54.8 (36.0–72.7)
A	3 (9.7)	1 (3.2)	0 (0)	0 (0)	0 (0)	4 (12.9)	
B	2 (6.5)	3 (9.7)	0 (0)	0 (0)	0 (0)	5 (16.1)	
C	2 (6.5)	2 (6.5)	3 (9.7)	1 (3.2)	0 (0)	8 (25.8)	
D	0 (0)	1 (3.2)	0 (0)	0 (0)	0 (0)	1 (3.2)	
E	2 (6.5)	2 (6.5)	0 (0)	0 (0)	8 (25.8)	12 (38.7)	
F	0 (0)	0 (0)	0 (0)	0 (0)	1 (3.2)	1 (3.2)	
4DCT							61.3 (42.2–78.2)
A	7 (22.6)	4 (12.9)	0 (0)	0 (0)	0 (0)	11 (35.5)	
B	1 (3.2)	2 (6.5)	0 (0)	0 (0)	0 (0)	3 (9.7)	
C	0 (0)	1 (3.2)	3 (9.7)	0 (0)	0 (0)	4 (12.9)	
D	0 (0)	0 (0)	0 (0)	1 (3.2)	3 (9.7)	4 (12.9)	
E	1 (3.2)	2 (6.5)	0 (0)	0 (0)	6 (19.4)	9 (29.0)	
F	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	
MIBI SPECT + 4DCT							74.2 (55.4–88.1)
A	7 (22.6)	3 (9.7)	0 (0)	0 (0)	0 (0)	10 (32.3)	
B	0 (0)	4 (12.9)	0 (0)	0 (0)	0 (0)	4 (12.9)	
C	0 (0)	2 (6.5)	3 (9.7)	1 (3.2)	0 (0)	6 (19.4)	
D	2 (6.5)	0 (0)	0 (0)	0 (0)	0 (0)	2 (6.5)	
E	0 (0)	0 (0)	0 (0)	0 (0)	9 (29)	9 (29.0)	
F	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	
Total	9 (29)	9 (29)	3 (9.7)	1 (3.2)	9 (29.0)		

Embryologic origin accuracy was 90.3% (28 of 31 patients). Quadrant localization accuracy was 87.1% (27 of 31 patients). Surgical classification accuracy was 61.3% (19 of 31 patients).

MIBI SPECT + 4DCT achieved an accuracy of 96.8% (30 of 31 patients) in correctly localizing the parathyroid adenomas on the left or right side. Embryologic origin accuracy was 96.8% (30 of 31 patients). Quadrant localization accuracy was 93.5% (29 of 31 patients). Surgical classification accuracy was 74.2% (23 of 31 patients).

We found that most errors in MIBI SPECT were due to superior glands being misidentified as inferior glands (Table 2), whereas this error was much less frequently observed in the 4DCT and MIBI SPECT + 4DCT readings.

Imaging Performance by Confidence Level

Table 5 shows reader team confidence in each localization technique. Confidence levels were generally higher for 4DCT and MIBI SPECT + 4DCT readings than they were for MIBI SPECT readings. Confidence in left/right localization was generally much higher than for quadrant localization. The more detailed level of surgical classification reduced confidence levels for MIBI SPECT to a greater degree than it did for 4DCT and MIBI SPECT +

Table 5: Level of confidence response from 3 reading groups for lateralization, upper/lower quadrant localization, and localization by surgical classification

Response/Imaging Modality	Certain (%)	Equivocal (%)	Uncertain (%)
Lateralization			
MIBI SPECT	27 (87.1)	1 (3.2)	3 (9.7)
4DCT	30 (96.8)	1 (3.2)	0 (0.0)
MIBI SPECT + 4DCT	29 (93.5)	2 (6.5)	0 (0.0)
Upper/lower			
MIBI SPECT	22 (71.0)	6 (19.4)	3 (9.7)
4DCT	30 (96.8)	1 (3.2)	0 (0.0)
MIBI SPECT + 4DCT	31 (100.0)	0 (0.0)	0 (0.0)
Surgical localization			
MIBI SPECT	15 (48.4)	15 (48.4)	1 (3.2)
4DCT	28 (90.3)	3 (9.7)	0 (0.0)
MIBI SPECT + 4DCT	29 (93.5)	1 (3.2)	1 (3.2)

Table 6: McNemar test of paired imaging modalities for diagnostic accuracy of left/right localization of parathyroid adenomas in the 31 patients in our study group

Imaging Modality	Correct	Wrong	Total	*P* Value
	4DCT vs			
MIBI SPECT				
Correct	28	1	29	.56
Wrong	2	0	2	
Total	30	1	31	
	MIBI SPECT + 4DCT vs			
4DCT				
Correct	29	1	30	1.00
Wrong	1	0	1	
Total	30	1	31	
	MIBI SPECT + 4DCT vs			
MIBI SPECT				
Correct	28	1	29	.56
Wrong	2	0	2	
Total	30	1	31	

Table 7: McNemar test of paired imaging modalities for diagnostic accuracy of embryologic origin of the abnormal parathyroid gland in the 31 patients in our study group

Imaging Modality	Correct	Wrong	Total	*P* Value
	4DCT vs			
MIBI SPECT				
Correct	22	1	23	.06
Wrong	6	2	8	
Total	28	3	31	
	MIBI SPECT + 4DCT vs			
4DCT				
Correct	28	0	28	.16
Wrong	2	1	3	
Total	30	1	31	
	MIBI SPECT + 4DCT vs			
MIBI SPECT				
Correct	23	0	23	.008
Wrong	7	1	8	
Total	30	1	31	

Table 8: McNemar test of paired imaging modalities for diagnostic accuracy of quadrant localization of the abnormal parathyroid gland in the 31 patients in our study group

Imaging Modality	Correct	Wrong	Total	*P* Value
	4DCT vs			
MIBI SPECT				
Correct	19	2	21	.06
Wrong	8	2	10	
Total	27	4	31	
	MIBI SPECT + 4DCT vs			
4DCT				
Correct	26	1	27	.32
Wrong	3	1	4	
Total	29	2	31	
	MIBI SPECT + 4DCT vs			
MIBI SPECT				
Correct	20	1	21	.01
Wrong	9	1	10	
Total	29	2	31	

4DCT, for which confidence was fairly consistently high regardless of the level of detail required.

Comparison of Imaging Modalities

Tables 6-9 summarize McNemar test results comparing the diagnostic accuracy of left- or right-sided localization, embryologic origin of the abnormal parathyroid gland (superior or inferior), quadrant localization (both left/right and upper/lower), and surgical classification among MIBI SPECT, 4DCT, and MIBI SPECT + 4DCT.

Table 9: McNemar test of paired imaging modalities for diagnostic accuracy of surgical localization of the abnormal parathyroid gland in the 31 patients in our study group

Imaging Modality	Correct	Wrong	Total	*P* Value
	4DCT vs			
MIBI SPECT				
Correct	11	6	17	.06
Wrong	8	6	14	
Total	19	12	31	
	MIBI SPECT + 4DCT vs			
4DCT				
Correct	16	3	19	0.21
Wrong	7	5	12	
Total	23	8	31	
	MIBI SPECT + 4DCT vs			
MIBI SPECT				
Correct	15	2	17	.06
Wrong	8	6	14	
Total	23	8	31	

McNemar test results comparing the diagnostic accuracy of left- or right-sided localization revealed no statistically significant differences, indicating substantially similar diagnostic performance.

McNemar test results comparing the diagnostic accuracy of embryologic origin revealed that MIBI SPECT + 4DCT was significantly better than MIBI SPECT ($P = .008$) and 4DCT was better than MIBI SPECT on the borderline of statistical significance ($P = .06$).

McNemar test results comparing the diagnostic accuracy of quadrant localization revealed that MIBI SPECT + 4DCT was significantly better than MIBI SPECT ($P = .01$) and 4DCT was better than MIBI SPECT on the borderline of statistical significance ($P = .06$).

McNemar test results comparing the diagnostic accuracy of surgical localization revealed that MIBI SPECT + 4DCT was better than MIBI SPECT on the borderline of statistical significance ($P = .06$).

Examples of correct and incorrect localization are shown in Figs 3 and 4. Figure 3 shows a large right tracheoesophageal groove parathyroid adenoma that was correctly localized by 3 modalities. Figure 4 shows a small left paraesophageal parathyroid adenoma that had correct left/right and quadrant localization by 4DCT and MIBI SPECT + 4DCT and correct left/right localization and incorrect quadrant localization by MIBI SPECT.

DISCUSSION

Our results showed that overall, MIBI SPECT was less accurate than 4DCT, which, in turn, was less accurate than combined MIBI SPECT + 4DCT. All imaging methods performed equally well in correctly localizing the parathyroid adenoma to the left or right side, and most of the inconsistency between MIBI SPECT and 4DCT was a result of errors in classification of embryologic origin. MIBI SPECT + 4DCT showed clear superiority over MIBI SPECT for embryologic and quadrant localization. Although our results were only marginally statistically significant, 4DCT may be better than MIBI SPECT for embryologic and quadrant localization and

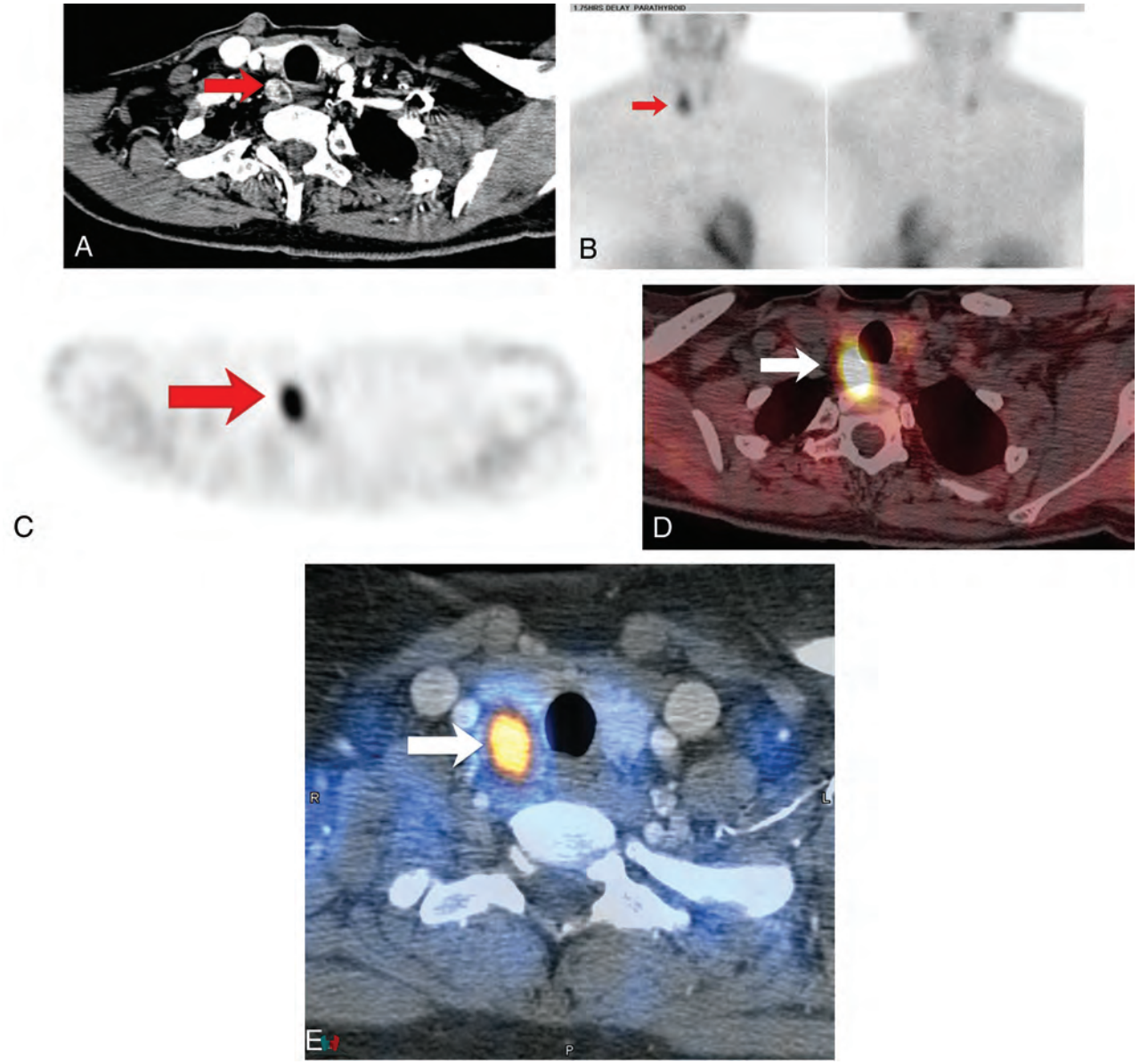


FIG 3. *A*, The axial arterial phase of a multiphase multidetector 4D CT image shows a right tracheoesophageal parathyroid adenoma (type C; Fig 1). *B*, Anterior and posterior delay planar scintigraphy shows retention of the radiotracer on the right side. *C*, An axial SPECT image shows retention of the radiotracer in the lower neck. *D*, An axial SPECT image fused to a noncontrast CT image localizes the retention of the radiotracer to the right tracheoesophageal groove. *E*, An axial SPECT image fused to the axial arterial phase of a multiphase multidetector 4D CT image localizes the retention of the radiotracer with concomitant early enhancement to the right tracheoesophageal groove. *Red and white arrows* show the parathyroid adenoma.

MIBI SPECT + 4DCT may be better than MIBI SPECT for surgical localization. The confidence of readers' localizations was generally higher with 4DCT and MIBI SPECT + 4DCT than it was with MIBI SPECT. The differential in both imaging performance and reader confidence decreased as the required level of detail in anatomic localization increased.

A recent meta-analysis of 24 published MIBI SPECT studies evaluating 1276 patients, in which patients with mixed diagnoses, secondary hyperparathyroidism, previous surgeries, and multigland disease were excluded, showed that MIBI SPECT had an estimated pooled sensitivity of 86%.[18] In a recent retrospective study of 65 patients with primary hyperparathyroidism, MIBI SPECT accurately localized the offending parathyroid gland with an average weight of 0.82 g in 80% of cases.[19] Compared with these previous studies, our study showed improved accuracy of MIBI SPECT for left/right localization and diminished accuracy of MIBI SPECT for quadrant localization. This could be explained by the poor image resolution of SPECT and lack of anatomic detail in low-dose CT for discriminating parathyroid adenomas from the thyroid gland. The ability to make this distinction would aid in quadrant localization of the adenoma as well as surgical classification.

Several investigators have shown that low-dose CT added to SPECT or MIBI SPECT can improve the diagnostic value for accurate localization of parathyroid adenomas.[12,20] Our study showed a substantial improvement in parathyroid adenoma localization accuracy with MIBI SPECT + 4DCT, or "diagnostic CT" added to SPECT, compared with MIBI SPECT alone.

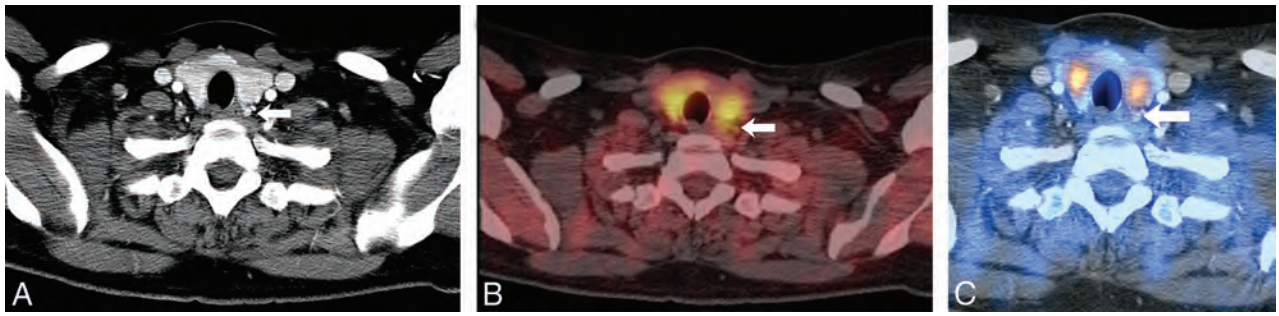


FIG 4. A case in which 4DCT could diagnose a parathyroid adenoma in the face of a sestamibi study with negative findings. *A*, The axial arterial phase of a multiphase multidetector 4D CT image shows a small early-enhancing left paraesophageal parathyroid adenoma (*arrow*). *B*, An axial MIBI SPECT image fused to a noncontrast CT image reveals no retention of the radiotracer in the left central compartment of a surgically proven parathyroid adenoma, which was effectively demonstrated on 4DCT (*arrow*). *C*, An axial SPECT image fused to the axial arterial phase of a multiphase multidetector 4D CT image shows a small early-enhancing left paraesophageal parathyroid adenoma (*arrow*).

In a retrospective study of 143 patients with primary hyperparathyroidism and no previous neck surgery, 4DCT had a sensitivity of 93.7% for left/right localization and 86.6% for quadrant localization.[15] Compared with that study, we report slightly improved accuracy of 4DCT for left/right localization and similar accuracy for quadrant localization. A recent meta-analysis of thirty-four 4DCT studies evaluating 2563 patients showed a pooled sensitivity of 81% for left/right localization and 73% for quadrant localization.[21] Compared with that study, we report improved accuracy of 4DCT for left/right localization and quadrant localization. This may be because our 4DCT imaging protocol used 4 phases, similar to the protocol described by Hunter et al,[15] whereas the meta-analysis represented various CT protocols across studies, ranging from 1 to 4 phases. In addition, the meta-analysis included studies of patients with previous operations and multigland disease, which, in our experience, are likely to be harder to interpret.

Heiba et al[22] compared neck pinhole dual-tracer and dual-phase sestamibi with and without SPECT for localization of 153 parathyroid adenomas, and they found that pinhole dual-tracer SPECT had significantly higher sensitivity (93%) than dual-phase SPECT (68%), SPECT with only pinhole delay (39%), pinhole dual-tracer CT alone (25%), and dual-phase CT alone (18%). In a prospective study comparing dual-isotope subtraction pinhole scintigraphy, dual-phase SPECT, 4DCT, and sonography in 91 patients with 97 parathyroid adenomas, Krakauer et al[23] found that the sensitivity of dual-isotope subtraction pinhole scintigraphy was 93%. The sensitivity of dual-phase SPECT (65%), 4DCT (58%), and sonography (57%) was significantly lower. In our study, we confirmed that dual-phase MIBI SPECT has lower localization accuracy compared with dual-isotope subtraction scintigraphy. In our study, 4DCT and MIBI SPECT + 4DCT yielded a sensitivity similar to that of dual-isotope subtraction scintigraphy.

The use of combined CT imaging modalities raises concerns about radiation exposure. The mean effective dose for 22.5 mCi of sestamibi is estimated at 7.5 mSv. The mean effective dose for 4DCT is 17.9 mSv (volume CT dose index = 20.8 mGy). Our effective dose for sestamibi was 7.5 mSv, and our dose with 4DCT was 17.9 mSv. These doses are similar to those in a recent article evaluating the total effective radiation doses associated with 4DCT (20.2 ± 2.8 mSv) and sestamibi (5.6 ± 0.24 mSv) used for parathyroid adenoma localization before surgery.[24] The total effective dose of MIBI SPECT + 4DCT is estimated at 27.7 mSv, which includes contributions from the low-dose noncontrast CT for SPECT attenuation correction. Although there are no good data on cancer incidence after low radiation exposure, conservatively, the net radiation exposure increases a patient's annual cancer risk by 0.019% and lifetime cancer risk by 0.52% compared with the baseline cancer incidence.[25,26] We believe that this is a favorable risk/benefit ratio given the very small cancer risk balanced against the serious health consequences of persistent hyperparathyroidism, especially in an older patient population.

The main strength of our study is the fully blinded, unbiased evaluation of different imaging modalities, all of which were acquired and viewed on identical platforms, thus eliminating any concerns about image coregistration or image interpretation in different environments. We also used a highly reliable reference standard and defined the exact location of each lesion by cross-referencing the histopathology reports and the operating notes with the imaging studies. In addition, all imaging scans were reviewed by expert radiologists or nuclear medicine physicians with considerable clinical experience in the interpretation of parathyroid localization studies. We directly compared MIBI SPECT and 4DCT as well as MIBI SPECT with diagnostic 4DCT to evaluate the incremental diagnostic value for parathyroid adenoma detection.

Limitations of this study include its retrospective nature and the small sample size. We included only patients with single-gland disease because of the limited sample size of multigland disease and the considerable differences between single- and multigland disease both biologically and in image interpretation. None of our imaging tests could perfectly localize lesions in the surgical classification scheme by Perrier et al,[17] but this classification is based on features (such as the location of the recurrent laryngeal nerve) that cannot yet be perceived by imaging and is thus of greater utility in the intraoperative environment.

CONCLUSIONS

Our results suggest that although all imaging modalities are useful for localizing parathyroid adenomas, MIBI SPECT + 4DCT is superior and 4DCT is marginally better than MIBI SPECT for determining embryologic and quadrant-level localization, particularly when detailed anatomic information is required, such as that needed to direct minimally invasive (targeted) surgery. This finding may strengthen the role of MIBI SPECT + 4DCT and

4DCT as first-line imaging techniques for preoperative localization. The confidence observed in localization, particularly useful to surgeons who need actionable information to plan minimally invasive surgery, was also higher for 4DCT than for MIBI SPECT. Combined interpretation of MIBI SPECT + 4DCT showed a still further increase in diagnostic accuracy and confidence beyond 4DCT alone, suggesting that 4DCT and MIBI SPECT are complementary modalities, supplying nonoverlapping information that might be useful, especially in difficult diagnostic cases.

ACKNOWLEDGMENTS

The authors thank Erica Goodoff for her assistance in editing.

Disclosures: W. Wei—*UNRELATED*: *Grant*: National Institutes of Health/National Cancer Institute, *Comments*: P30 CA016672.* S. Cheenu Kappadath—*UNRELATED*: *Consultancy*: BTG International Inc; *Grants*: GE Healthcare, BTG International Inc.* Hubert H. Chuang—*UNRELATED*: *Consultancy*: SAGE Consulting. *Money paid to the institution.

REFERENCES

1. Rosen CJ; American Society for Bone and Mineral Research. *Primer on the Metabolic Bone Diseases and Disorders of Mineral Metabolism.* Ames: Wiley-Blackwell; 2013
2. Mihai R, Barczynski M, Iacobone M, et al. **Surgical strategy for sporadic primary hyperparathyroidism an evidence-based approach to surgical strategy, patient selection, surgical access, and reoperations.** *Langenbecks Arch Surg* 2009;394:785–98 CrossRef Medline
3. Thomas DL, Bartel T, Menda Y, et al. **Single photon emission computed tomography (SPECT) should be routinely performed for the detection of parathyroid abnormalities utilizing technetium-99m sestamibi parathyroid scintigraphy.** *Clin Nucl Med* 2009;34:651–55 CrossRef Medline
4. Lindqvist V, Jacobsson H, Chandanos E, et al. **Preoperative 99Tc(m)-sestamibi scintigraphy with SPECT localizes most pathologic parathyroid glands.** *Langenbecks Arch Surg* 2009;394:811–15 CrossRef Medline
5. Smith RB, Evasovich M, Girod DA, et al. **Ultrasound for localization in primary hyperparathyroidism.** *Otolaryngol Head Neck Surg* 2013;149:366–71 CrossRef Medline
6. Levy JM, Kandil E, Yau LC, et al. **Can ultrasound be used as the primary screening modality for the localization of parathyroid disease prior to surgery for primary hyperparathyroidism? A review of 440 cases.** *ORL J Otorhinolaryngol Relat Spec* 2011;73:116–20 CrossRef Medline
7. Rodgers SE, Hunter GJ, Hamberg LM, et al. **Improved preoperative planning for directed parathyroidectomy with 4-dimensional computed tomography**. *Surgery* 2006;140:932–40; discussion 940–41 CrossRef Medline
8. Kabala JE. **Computed tomography and magnetic resonance imaging in diseases of the thyroid and parathyroid.** *Eur J Radiol* 2008;66:480–92 CrossRef Medline
9. Weber T, Maier-Funk C, Ohlhauser D, et al. **Accurate preoperative localization of parathyroid adenomas with C-11 methionine PET/CT.** *Ann Surg* 2013;257:1124–28 CrossRef Medline
10. Orevi M, Freedman N, Mishani E, et al. **Localization of parathyroid adenoma by ^{11}C-choline PET/CT: preliminary results.** *Clin Nucl Med* 2014;39:1033–38 CrossRef Medline
11. Krausz Y, Bettman L, Guralnik L, et al. **Technetium-99m-MIBI SPECT/CT in primary hyperparathyroidism.** *World J Surg* 2006;30:76–83 CrossRef Medline
12. Lavely WC, Goetze S, Friedman KP, et al. **Comparison of SPECT/CT, SPECT, and planar imaging with single- and dual-phase (99m)Tc-sestamibi parathyroid scintigraphy.** *J Nucl Med* 2007;48:1084–89 CrossRef Medline
13. Vaz A, Griffiths M. **Parathyroid imaging and localization using SPECT/CT: initial results.** *J Nucl Med Technol* 2011;39:195–200 CrossRef Medline
14. Starker LF, Mahajan A, Björklund P, et al. **4D parathyroid CT as the initial localization study for patients with de novo primary hyperparathyroidism.** *Ann Surg Oncol* 2011;18:1723–28 CrossRef Medline
15. Hunter GJ, Schellingerhout D, Vu TH, et al. **Accuracy of four-dimensional CT for the localization of abnormal parathyroid glands in patients with primary hyperparathyroidism.** *Radiology* 2012;264:789–95 CrossRef Medline
16. Kelly HR, Hamberg LM, Hunter GJ. **4D-CT for preoperative localization of abnormal parathyroid glands in patients with hyperparathyroidism: accuracy and ability to stratify patients by unilateral versus bilateral disease in surgery-naive and reexploration patients.** *AJNR Am J Neuroradiol* 2014;35:176–81 CrossRef Medline
17. Perrier ND, Edeiken B, Nunez R, et al. **A novel nomenclature to classify parathyroid adenomas.** *World J Surg* 2009;33:412–16 CrossRef Medline
18. Wong KK, Fig LM, Gross MD, et al. **Parathyroid adenoma localization with 99mTc-sestamibi SPECT/CT: a meta-analysis.** *Nucl Med Commun* 2015;36:363–75 CrossRef Medline
19. Keidar Z, Solomonov E, Karry R, et al. **Preoperative [99mTc]MIBI SPECT/CT interpretation criteria for localization of parathyroid adenomas: correlation with surgical findings.** *Mol Imaging Biol* 2017;19:265–70 CrossRef Medline
20. Neumann DR, Obuchowski NA, Difilippo FP. **Preoperative 123I/99mTc-sestamibi subtraction SPECT and SPECT/CT in primary hyperparathyroidism.** *J Nucl Med* 2008;49:2012–17 CrossRef Medline
21. Kluijfhout WP, Pasternak JD, Beninato T, et al. **Diagnostic performance of computed tomography for parathyroid adenoma localization: a systematic review and meta-analysis.** *Eur J Radiol* 2017;88:117–28 CrossRef Medline
22. Heiba SI, Jiang M, Rivera J, et al. **Direct comparison of neck pinhole dual-tracer and dual-phase MIBI accuracies with and without SPECT/CT for parathyroid adenoma detection and localization.** *Clin Nucl Med* 2015;40:476–82 CrossRef Medline
23. Krakauer M, Wieslander B, Myschetzky PS, et al. **A prospective comparative study of parathyroid dual-phase scintigraphy, dual-isotope subtraction scintigraphy, 4D-CT, and ultrasonography in primary hyperparathyroidism.** *Clin Nucl Med* 2016;41:93–100 CrossRef Medline
24. Moosvi SR, Smith S, Hathorn J, et al. **Evaluation of the radiation dose exposure and associated cancer risks in patients having preoperative parathyroid localization.** *Ann R Coll Surg Engl* 2017;99:363–68 CrossRef Medline
25. Ray P, Vu T, Romero M, et al. **Limiting the risks of radiation exposure in diagnostic imaging.** *Surgery* 2014;156:1297–99 CrossRef Medline
26. Hoang JK, Reiman RE, Nguyen GB, et al. **Lifetime attributable risk of cancer from radiation exposure during parathyroid imaging: comparison of 4D CT and parathyroid scintigraphy.** *AJR Am J Rroentgenol* 2015;204:W579–85 CrossRef Medline

Parapharyngeal Space Venous Malformation: An Imaging Mimic of Pleomorphic Adenoma

C.M. Tomblinson, G.P. Fletcher, T.K. Lidner, C.P. Wood, S.M. Weindling, and J.M. Hoxworth

ABSTRACT

SUMMARY: Venous malformations in the parapharyngeal space are rare and may be challenging to diagnose with imaging secondary to multiple overlapping features with pleomorphic adenoma, which is much more commonly found in this region. While both lesions are T1 isointense and T2 hyperintense relative to skeletal muscle and demonstrate contrast enhancement, more uniform T2 hyperintensity and progressive contrast pooling on delayed postcontrast T1WI may allow the radiologist to include venous malformation in the differential diagnosis. This is important because it has the potential to alter management from surgical resection to observation. The primary aim of this study was to review the imaging appearance of parapharyngeal venous malformations through a retrospective case series.

ABBREVIATIONS: PPS = parapharyngeal space; VM = venous malformation

Venous malformations (VMs) are typically considered unencapsulated, circumscribed, or trans-spatial lesions comprising dysplastic serpentine sinusoids that may contain phleboliths. Although "cavernous hemangioma" is a colloquial term that remains in use, it has been increasingly recognized that many such lesions are actually low-flow VMs as opposed to hemangiomas, with the latter being a true neoplasm.[1-4] In 1982, Mulliken and Glowacki[5] proposed a biologic classification scheme that categorized vascular anomalies on the basis of the rate of cell turnover, histology, natural history, and examination findings. The current International Society for the Study of Vascular Anomalies classification system is grounded in this sentinel work, which groups vascular anomalies into 2 categories: tumors and malformations.[4] The most common vascular tumor is infantile hemangioma, while VM is the most common vascular malformation.

Parapharyngeal space (PPS) masses are uncommon, composing only 0.5% of all head and neck tumors.[6] Of these, VMs represent <1% of all PPS masses.[7] Salivary gland tumors, most commonly pleomorphic adenomas, account for 40%–50% of all PPS tumors, arising either primarily from minor salivary rests or secondarily encroaching on the PPS from the deep parotid lobe.[7-9]

Both VM and pleomorphic adenoma may appear as a well-circumscribed, mildly lobulated enhancing mass that is isodense to muscle on CT and markedly T2 hyperintense on MR imaging.[1,3,10,11] As a result, distinguishing these 2 entities when they occur in the PPS can be challenging, and this distinction is clinically important because management for salivary gland tumor and VM may differ. To date, very few case reports have published the imaging features of patients with primary PPS VMs, which have been frequently described as "hemangiomas," reflecting the controversial ongoing use of older nomenclature.[12-16] The primary aim of this study was to review the imaging appearance of PPS VMs via a retrospective case series and to seek specific imaging features that may allow differentiation from benign salivary gland tumors such as pleomorphic adenoma.

Case Series

This retrospective study was approved by the institutional review board at the Mayo Clinic with a waiver of informed consent. The pathology data base was electronically queried from 2000 to 2017 using a keyword search. Patients with pathologically confirmed PPS VMs were identified, and those with pretreatment CT and/or MR imaging were included. Demographic and clinical data were obtained from the electronic medical record. The imaging appearance of these lesions, including CT density, MR imaging signal characteristics, enhancement pattern, the presence or absence of calcification, the relationship to adjacent structures, size, and

Received June 22, 2018; accepted after revision September 4.

From the Departments of Radiology (C.M.T., G.P.F., J.M.H.) and Laboratory Medicine and Pathology (T.K.L.), Mayo Clinic, Phoenix, Arizona; Department of Radiology (C.P.W.), Mayo Clinic, Rochester, Minnesota; and Department of Radiology (S.M.W.), Mayo Clinic, Jacksonville, Florida.

Paper previously presented at: Annual Meeting of the American Society of Head and Neck Radiology, September 16–20, 2017; Las Vegas, Nevada.

Please address correspondence to Joseph M. Hoxworth, MD, Mayo Clinic, Department of Radiology, 5777 E Mayo Blvd, Phoenix, AZ 85054; e-mail: hoxworth.joseph@mayo.edu

Indicates article with supplemental on-line table.

http://dx.doi.org/10.3174/ajnr.A5859

stability (On-line Table), was characterized through a detailed review by a board-certified neuroradiologist.

Six patients met the inclusion criteria, 5 women and 1 man (34–76 years of age; mean, 59 years). All lesions were identified incidentally on imaging studies obtained for an unrelated clinical problem, which included headache, tongue numbness (contralateral to the side of the VM), multiple sclerosis, cervical spondylosis, and staging work-up for sinonasal carcinoma. Following initial detection, all patients were evaluated by a head and neck surgeon, and none of the masses were palpable externally. Typically following a bloody acellular aspirate, diagnosis was established with CT-guided core biopsy in 4 patients and with surgical resection in 2 patients. Three of the lesions were diagnosed as PPS VMs at the time of tissue sampling, while the remaining 3 were initially characterized as cavernous hemangiomas but later confirmed to be VMs at the time of study inclusion.

All VMs were located in the prestyloid PPS.[17,18] Four were right-sided, 2 were left-sided, and only 2 contacted the deep lobe of the parotid gland. Maximum dimensions ranged from 1.3 to 2.2 cm, while volume (ROI Volume tool, OsiriX Imaging Software; Version 9.0, http://www.osirix-viewer.com) ranged from 0.8 to 5.5 cm^3 (mean, 2.3 $\pm$ 1.7 cm^3).

All MR imaging scans were of good technical quality with no need for repeat imaging. On MR imaging, all VMs were well-defined and, relative to skeletal muscle, appeared markedly hyperintense on T2WI and isointense-to-minimally hyperintense on precontrast T1WI (Fig 1). Postgadolinium T1WI was available for review in all 6 patients and revealed highly variable enhancement patterns (Fig 2). Heterogeneously diffuse enhancement and stippled enhancement were each seen in 2 VMs, while single lesions demonstrated homogeneous diffuse enhancement and focal central enhancement, respectively. Four of 6 VMs demonstrated a qualitatively greater volume of lesion enhancement on the more delayed of the 2 consecutive postgadolinium sequences (Fig 2*F*). All postgadolinium T1WI used fat suppression, with the axial plane obtained first and immediately followed by the coronal acquisition.

On noncontrast CT, all of the PPS VMs were isodense to skeletal muscle and none contained phleboliths or other calcifications. Only 1 patient had postcontrast CT images that demonstrated enhancement of the central portion of the lesion (Fig 2*D*).

Sonography, available for review in 2 patients, demonstrated well-defined masses that were minimally hyperechoic compared with the pterygoid muscles (Fig 3). Identifiable vessels within the masses exhibited venous waveforms. Intraoral sonography, which was pursued in both patients to evaluate candidacy for transoral resection, allowed better proximity to the PPS to improve image quality.

Most patients had no follow-up imaging in our institution after biopsy or surgical resection. In 1 patient, the VM was stable at 8 months following biopsy. In a second patient, the lesion was 60% smaller by volume 5 years following core biopsy.

All MR imaging and CT studies were initially interpreted by a

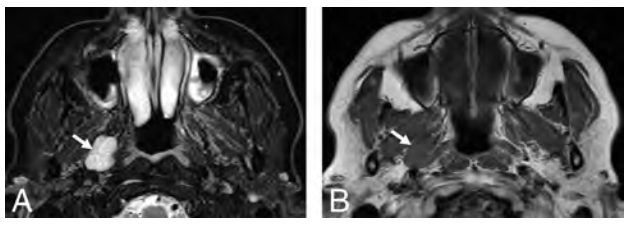


FIG 1. Typical MR imaging signal characteristics of a VM (*white arrow*) in the right PPS of a 56-year-old woman. *A*, Axial fat-suppressed T2-weighted MR imaging depicts a well-circumscribed lobulated markedly hyperintense PPS mass. *B*, Axial T1-weighted MR imaging shows the mass as isointense to skeletal muscle.

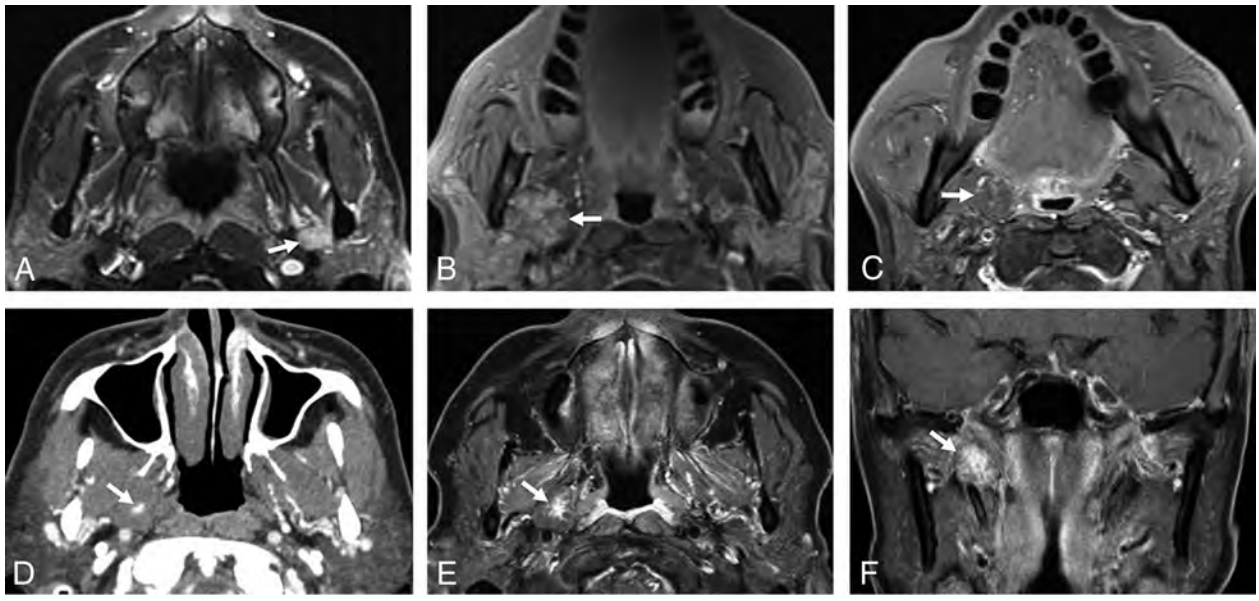


FIG 2. Variable enhancement patterns of PPS VMs (*white arrows*). *A*, Axial T1-weighted fat-suppressed postcontrast MR imaging in a 34-year-old man demonstrates diffuse homogeneous enhancement. *B*, Axial T1-weighted fat-suppressed postcontrast MR imaging from a 66-year-old woman illustrates a diffuse heterogeneous enhancement pattern. *C*, Axial T1-weighted fat-suppressed postcontrast MR imaging from a 61-year-old woman depicts stippled enhancement. Axial contrast-enhanced CT (*D*), axial T1-weighted fat-suppressed postcontrast MR imaging (*E*), and coronal T1-weighted fat-suppressed postcontrast MR imaging (*F*) from a 56-year-old woman. The axial CT and MR imaging demonstrate focal central enhancement, while the coronal image, which was acquired after additional time had elapsed following the intravenous contrast infusion, illustrates progressively more diffuse contrast enhancement.

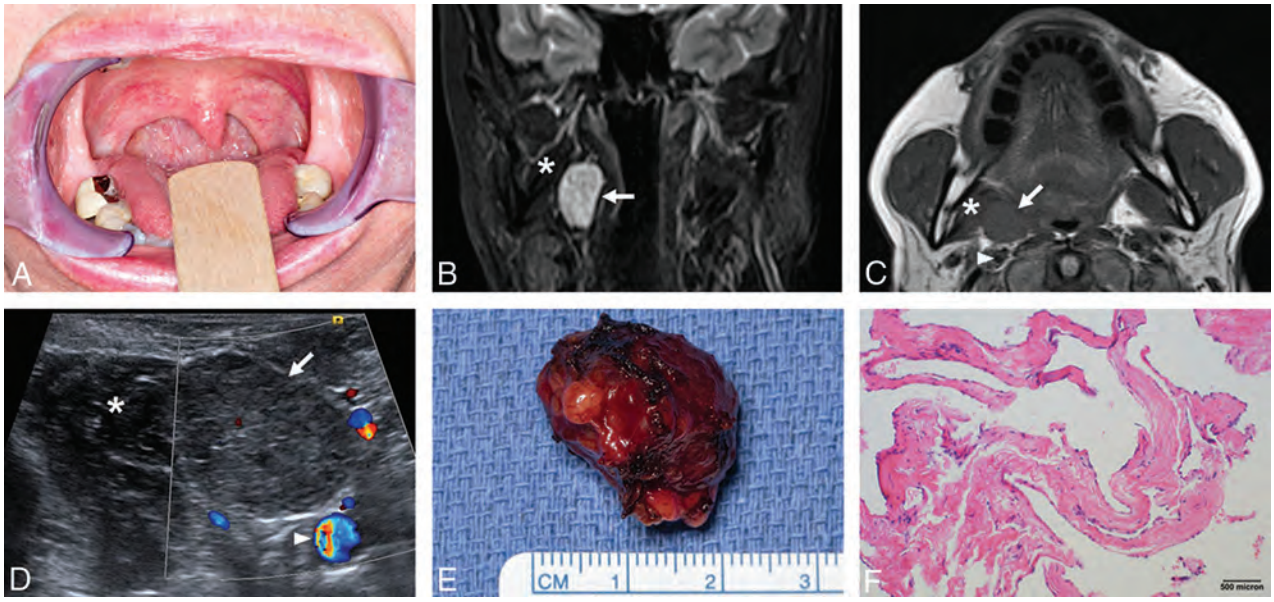


FIG 3. Clinicopathologic correlation for a right PPS VM in a 61-year-old woman. *A*, Intraoral photograph demonstrates subtle right pharyngeal fullness with mild leftward deviation of the uvula. *B*, Coronal fat-suppressed T2-weighted MR imaging demonstrates a homogeneously hyperintense right PPS mass. *C*, Axial T1-weighted MR imaging depicts a round, well-circumscribed mass that is minimally hyperintense relative to skeletal muscle with effacement of the right parapharyngeal fat. *D*, Intraoral sonography illustrates the close relationship to the medial pterygoid muscle and the internal carotid artery. The VM is slightly hyperechoic compared with skeletal muscle. *E*, Gross pathologic specimen demonstrates the lobulated shape of this VM, correlating with the sonographic findings. *F*, Hematoxylin-eosin-stained photomicrograph illustrates nonanastomosing dilated endothelial-lined spaces, and the endothelial cells lack atypia. These findings are consistent with VM. VM (*arrow*), medial pterygoid muscle (*asterisk*), internal carotid artery (*arrowhead*).

board-certified neuroradiologist. On the basis of review of the radiology reports, salivary gland tumor or, more specifically, pleomorphic adenoma was the most commonly suggested diagnosis before tissue sampling in 5 of 6 patients. One lesion was thought to represent a metastatic lymph node in an aggressive sinonasal carcinoma. The possibility of a VM was not suggested in the radiology report differential diagnosis before tissue diagnosis in any of our patients.

DISCUSSION

Evaluation of head and neck vascular anomalies is heavily dependent on MR imaging, which has excellent spatial resolution and soft-tissue contrast, allowing both lesion characterization and assessment of extent for pretreatment planning.[1-3,19,20] This dependence on MR Imaging is particularly true for PPS VMs, which may not be detectable by palpation or skin discoloration. Because of the rich slow-flowing venous blood supply, VMs are T2 hyperintense and T1 intermediate on MR imaging, and signal voids may be present in the case of phleboliths. VM enhancement is often patchy and delayed, demonstrating sequential filling of the lesion with slow washout. At our institution, postgadolinium T1WI sequences are first obtained in the axial plane approximately 90 seconds postinjection and take approximately 5 minutes to acquire. The subsequent coronal postgadolinium T1WI acquisition begins approximately 6–7 minutes postinjection. In 4 of 6 patients, the PPS VM demonstrated increased enhancement on the delayed coronal postgadolinium images, best characterized as contrast filling-in the mass to a greater degree.

Unenhanced CT may be used as an adjunctive tool to assess phleboliths in VM. In our series, none of the VMs contained phleboliths, but they have indeed been reported in PPS VMs.[12] As a result, there is a high likelihood that the absence of phleboliths in the current series represents a sampling bias, given the small size of our cohort. While phleboliths are considered pathognomonic for VMs, 1 study of head and neck VMs reported phleboliths in only 28.6%, though a rigorous, uniform imaging approach was not used for phlebolith identification.[21]

Although sonography is useful for the assessment of soft-tissue vascular anomalies,[22] lesions deep in the PPS may be difficult to adequately visualize. Intraoral endosonography (Fig 3*D*) may be performed using an endocavitary probe if the lesion is located medially in the PPS. This can be a useful adjunct for planning transoral surgery, which is an increasingly viable approach for PPS lesions.[23] At sonography, VMs appear as lobulated hypoechoic masses with small internal venous channels.[3] On spectral Doppler imaging, waveforms are venous, and a curvilinear hyperechoic focus with posterior shadowing may be identified if phleboliths are present.

Although vascular malformations are more likely to be transspatial and less well-defined than hemangiomas,[22] all of the PPS VMs in our series appeared as lobulated, well-circumscribed soft-tissue masses, which is concordant with other reports.[12-16] This imaging appearance can lead to diagnostic uncertainty, with PPS VMs being mistaken for the much more common pleomorphic adenoma because both can occur with or without a connection to the deep lobe of the parotid gland. Several imaging features may assist in differentiating PPS VM from pleomorphic adenoma. First, although PPS pleomorphic adenomas may calcify with greater frequency than those within the parotid gland (50% versus 15%), the calcification is more likely to be coarse or punctate,[11] which differs from the typical laminated appearance of a phlebolith.[21] Second, PPS VMs are more

commonly homogeneously T2 hyperintense compared with muscle, while pleomorphic adenomas in the PPS may exhibit T2 isointensity or hypointensity within their solid components secondary to hypercellularity with less myxoid stroma.[11,24] Last, two-thirds of PPS VMs in our series demonstrated progressive contrast pooling on routine delayed postcontrast MR imaging (ie, without performing dedicated dynamic contrast-enhanced imaging). Although pleomorphic adenomas can slowly accumulate and retain intravenous contrast,[24,25] the pattern of initial central or stippled enhancement with delayed filling-in seen with some VMs is less characteristic of pleomorphic adenomas.

In a systematic review of 1293 cases of PPS masses by Kuet et al,[7] the most common clinical presentations were cervical mass and intraoral swelling. In contrast to our series, none of the published reports that they reviewed were based on incidental findings from imaging. Patients are more likely to present with a detectable PPS mass when the lesion measures >2.5–3.0 cm,[26,27] so the small size of the VMs in our series is concordant with incidental detection. Certainly, large PPS VMs can become palpable and symptomatic.[12] Symptoms that should raise concern for malignancy include referred otalgia, facial pain, trismus, and cranial nerve involvement.[7,27] None of the patients in our series demonstrated any of these worrisome symptoms.

The primary limitations of the current case series are its small size and retrospective methodology, though the imaging features of these uncommon PPS VMs were relatively consistent throughout the cohort. Comparison of the current results with the few previously published reports of PPS VM imaging features is also limited because we must presume that previous reports describing PPS "hemangiomas" in adult patients were indeed referring to VMs.

CONCLUSIONS

Because of their rarity, PPS VMs are difficult to diagnose prospectively on MR imaging and CT and, when small, may be encountered incidentally on imaging studies obtained for other reasons. This lesion can present a diagnostic dilemma related to overlapping location and imaging characteristics of the much more common PPS pleomorphic adenoma. With heightened awareness of VMs occurring in the PPS and their typical imaging features, the radiologist may appropriately include VM in the differential diagnosis of a PPS mass.

REFERENCES

1. Baker LL, Dillon WP, Hieshima GB, et al. **Hemangiomas and vascular malformations of the head and neck: MR characterization.** *AJNR Am J Neuroradiol* 1993;14:307–14 Medline
2. Lowe LH, Marchant TC, Rivard DC, et al. **Vascular malformations: classification and terminology the radiologist needs to know.** *Semin Roentgenol* 2012;47:106–17 CrossRef Medline
3. Steinklein JM, Shatzkes DR. **Imaging of vascular lesions of the head and neck.** *Otolaryngol Clin North Am* 2018;51:55–76 CrossRef Medline
4. Wassef M, Blei F, Adams D, et al; ISSVA Board and Scientific Committee. **Vascular anomalies classification: recommendations from the International Society for the Study of Vascular Anomalies.** *Pediatrics* 2015;136:e203–14 CrossRef Medline
5. Mulliken JB, Glowacki J. **Hemangiomas and vascular malformations in infants and children: a classification based on endothelial characteristics.** *Plast Reconstr Surg* 1982;69:412–22 CrossRef Medline
6. Batsakis JG, Sneige N. **Parapharyngeal and retropharyngeal space diseases.** *Annals of Otology, Rhinology & Laryngology* 1989;98:320–21 CrossRef Medline
7. Kuet ML, Kasbekar AV, Masterson L, et al. **Management of tumors arising from the parapharyngeal space: a systematic review of 1,293 cases reported over 25 years.** *Laryngoscope* 2015;125:1372–81 CrossRef Medline
8. Khafif A, Segev Y, Kaplan DM, et al. **Surgical management of parapharyngeal space tumors: a 10-year review.** *Otolaryngol Head Neck Surg* 2005;132:401–06 CrossRef Medline
9. Hughes KV 3rd, Olsen KD, McCaffrey TV. **Parapharyngeal space neoplasms.** *Head Neck* 1995;17:124–30 CrossRef Medline
10. Ikeda K, Katoh T, Ha-Kawa SK, et al. **The usefulness of MR in establishing the diagnosis of parotid pleomorphic adenoma.** *AJNR Am J Neuroradiol* 1996;17:555–59 Medline
11. Kato H, Kanematsu M, Mizuta K, et al. **Imaging findings of parapharyngeal space pleomorphic adenoma in comparison with parotid gland pleomorphic adenoma.** *Jpn J Radiol* 2013;31:724–30 CrossRef Medline
12. Cho JH, Joo YH, Kim MS, et al. **Venous hemangioma of parapharyngeal space with calcification.** *Clin Exp Otorhinolaryngol* 2011;4:207–09 CrossRef Medline
13. Gennaro P, Chisci G, Gabriele G, et al. **Management of a bulky capillary hemangioma in the parapharyngeal space with minimally invasive surgery.** *J Craniofac Surg* 2014;25:e161–63 CrossRef Medline
14. Granell J, Alonso A, Garrido L, et al. **Transoral fully robotic dissection of a parapharyngeal hemangioma.** *J Craniofac Surg* 2016;27:1806–07 CrossRef Medline
15. Kale US, Ruckley RW, Edge CJ. **Cavernous haemangioma of the parapharyngeal space.** *Indian J Otolaryngol Head Neck Surg* 2006;58:77–80 Medline
16. Surasi DS, Meibom S, Grillone G, et al. **F-18 FDG PET/CT imaging of a parapharyngeal hemangioma.** *Clin Nucl Med* 2010;35:612–13 CrossRef Medline
17. Curtin HD. **Separation of the masticator space from the parapharyngeal space.** *Radiology* 1987;163:195–204 CrossRef Medline
18. Gupta A, Chazen JL, Phillips CD. **Imaging evaluation of the parapharyngeal space.** *Otolaryngol Clin North Am* 2012;45:1223–32 CrossRef Medline
19. Baer AH, Parmar HA, DiPietro MA, et al. **Hemangiomas and vascular malformations of the head and neck: a simplified approach.** *Neuroimaging Clin N Am* 2011;21:641–58, viii CrossRef Medline
20. Donnelly LF, Adams DM, Bisset GS 3rd. **Vascular malformations and hemangiomas: a practical approach in a multidisciplinary clinic.** *AJR Am J Roentgenol* 2000;174:597–608 CrossRef Medline
21. Eivazi B, Fasunla AJ, Guldner C, et al. **Phleboliths from venous malformations of the head and neck.** *Phlebology* 2013;28:86–92 CrossRef Medline
22. Paltiel HJ, Burrows PE, Kozakewich HP, et al. **Soft-tissue vascular anomalies: utility of US for diagnosis.** *Radiology* 2000;214:747–54 CrossRef Medline
23. Duek I, Sviri GE, Billan S, et al. **Minimally invasive surgery for resection of parapharyngeal space tumors.** *J Neurol Surg B Skull Base* 2018;79:250–56 CrossRef Medline
24. Motoori K, Yamamoto S, Ueda T, et al. **Inter- and intratumoral variability in magnetic resonance imaging of pleomorphic adenoma: an attempt to interpret the variable magnetic resonance findings.** *J Comput Assist Tomogr* 2004;28:233–46 CrossRef Medline
25. Lev MH, Khanduja K, Morris PP, et al. **Parotid pleomorphic adenomas: delayed CT enhancement.** *AJNR Am J Neuroradiol* 1998;19:1835–39 Medline
26. Som PM, Biller HF, Lawson W, et al. **Parapharyngeal space masses: an updated protocol based upon 104 cases.** *Radiology* 1984;153:149–56 CrossRef Medline
27. Carrau RL, Myers EN, Johnson JT. **Management of tumors arising in the parapharyngeal space.** *Laryngoscope* 1990;100:583–89 CrossRef Medline

MR Imaging–Based Radiomic Signatures of Distinct Molecular Subgroups of Medulloblastoma

M. Iv, M. Zhou, K. Shpanskaya, S. Perreault, Z. Wang, E. Tranvinh, B. Lanzman, S. Vajapeyam, N.A. Vitanza, P.G. Fisher, Y.J. Cho, S. Laughlin, V. Ramaswamy, M.D. Taylor, S.H. Cheshier, G.A. Grant, T. Young Poussaint, O. Gevaert, and K.W. Yeom

ABSTRACT

BACKGROUND AND PURPOSE: Distinct molecular subgroups of pediatric medulloblastoma confer important differences in prognosis and therapy. Currently, tissue sampling is the only method to obtain information for classification. Our goal was to develop and validate radiomic and machine learning approaches for predicting molecular subgroups of pediatric medulloblastoma.

MATERIALS AND METHODS: In this multi-institutional retrospective study, we evaluated MR imaging datasets of 109 pediatric patients with medulloblastoma from 3 children's hospitals from January 2001 to January 2014. A computational framework was developed to extract MR imaging–based radiomic features from tumor segmentations, and we tested 2 predictive models: a double 10-fold cross-validation using a combined dataset consisting of all 3 patient cohorts and a 3-dataset cross-validation, in which training was performed on 2 cohorts and testing was performed on the third independent cohort. We used the Wilcoxon rank sum test for feature selection with assessment of area under the receiver operating characteristic curve to evaluate model performance.

RESULTS: Of 590 MR imaging–derived radiomic features, including intensity-based histograms, tumor edge-sharpness, Gabor features, and local area integral invariant features, extracted from imaging-derived tumor segmentations, tumor edge-sharpness was most useful for predicting sonic hedgehog and group 4 tumors. Receiver operating characteristic analysis revealed superior performance of the double 10-fold cross-validation model for predicting sonic hedgehog, group 3, and group 4 tumors when using combined T1- and T2-weighted images (area under the curve = 0.79, 0.70, and 0.83, respectively). With the independent 3-dataset cross-validation strategy, select radiomic features were predictive of sonic hedgehog (area under the curve = 0.70–0.73) and group 4 (area under the curve = 0.76–0.80) medulloblastoma.

CONCLUSIONS: This study provides proof-of-concept results for the application of radiomic and machine learning approaches to a multi-institutional dataset for the prediction of medulloblastoma subgroups.

ABBREVIATIONS: AUC = area under the curve; LAII = local area integral invariant; MB = medulloblastoma; ROC = receiver operating characteristic; SHH = sonic hedgehog; SVM = support vector machines; WNT = wingless type

Medulloblastoma (MB) is the most common malignant brain tumor in children and a leading cause of cancer-related morbidity and mortality in this population.[1] Although once considered a single tumor type, recent molecular advances have identified at least 4 biologically distinct subgroups of MB (sonic hedgehog [SHH],

Received June 29, 2018; accepted after revision October 6.

From the Department of Radiology (M.I., M.Z., K.S., E.T., B.L., K.W.Y.), Department of Pediatrics (P.G.F.), Pediatric Neurology, and Department of Neurosurgery (G.A.G.), Pediatric Neurosurgery, Lucile Packard Children's Hospital, Stanford University, Palo Alto, California; Stanford Center for Biomedical Informatics (M.Z., O.G., Z.W.) and Department of Radiology (K.W.Y.), Artificial Intelligence in Medicine and Imaging, Stanford University, Stanford, California; Department of Pediatrics (S.P.), Pediatric Neurology, Centre Hospitalier Universitaire Sainte Justine, University of Montréal, Montreal, Quebec, Canada; Department of Radiology (S.V., T.Y.P.), Boston Children's Hospital, Harvard University, Boston, Massachusetts; Department Pediatrics Hematology-Oncology (N.A.V.), Seattle Children's Hospital, University of Washington, Seattle, Washington; Department of Pediatrics (Y.J.C.), Pediatric Neurology, Oregon Health & Science University, Portland, Oregon; Departments of Radiology, Neuro-Oncology, and Neurosurgery (S.L., V.R., M.D.T.), Hospital for Sick Children, Toronto, Ontario, Canada; and Department of Neurosurgery (S.H.C.), Pediatric Neurosurgery, University of Utah, Salt Lake City, Utah.

Michael Iv and Mu Zhou are co-first authors and contributed equally to this article.

Kristen W. Yeom and Olivier Gevaert are co-senior and co-corresponding authors and contributed equally to this article.

Research reported in this publication was supported by the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health under Award Number R01EB020527. V.R. was supported by the collaborative ependymoma research network.

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

Please address correspondence to Kristen W. Yeom, MD, Lucile Packard Children's Hospital, Stanford University Medical Center, Department of Radiology, Pediatric MRI and CT, Room 0511, 725 Welch Rd, Palo Alto, CA 94304; e-mail: kyeom@stanford.edu.

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

Indicates article with supplemental on-line tables.

Indicates article with supplemental on-line photos.

http://dx.doi.org/10.3174/ajnr.A5899

Table 1: Patient demographics

Characteristic	Institutional Cohort		
	Stanford	Boston	Toronto
No. of patients	32	28	49
Age (mean) (yr)	10.14 ± 8.49	8.54 ± 4.52	7.53 ± 3.69
Male sex (No.) (%)	23 (72)	9 (32)	32 (65)
Molecular subgroup (No.)			
SHH	11	9	10
WNT	4	5	10
Group 3	7	5	12
Group 4	10	9	17
MRI availability (No.)			
3T	5	0	1
1.5T	27	28	48
T1-weighted	32	26	48
2D T1-weighted	28	25	7
3D T1-weighted	4	1	41
T2-weighted	30	27	29
T1- and T2-weighted	30	25	27

wingless-type [WNT], group 3, and group 4) with specific subgroups conferring important prognostic and therapeutic differences.[2-4] For example, patients with WNT-pathway-activated tumors have favorable outcomes with a nearly 90% 5-year survival rate, while patients with group 3 tumors have <50% overall survival.[5] These divergent prognostic outcomes have propelled the recognition of these 4 subgroups, reflected in the recent revision of the World Health Organization classification of MB.[6] These molecular subgroups now drive risk-stratification, clinical outcome modeling, and novel therapeutic development.[7,8]

Subtyping of tumors is frequently performed on tissues obtained from surgical resection but can also be performed from tissues obtained from a single biopsy. Even single biopsies of MB can yield accurate information for subtyping because of the presence of spatially homogeneous transcriptomes in MBs, in contrast to other tumor types such as high-grade gliomas.[9] However, surgical sampling is invasive and confers added risk to patients. In addition, despite the increasing clinical utility of MB subtyping, the translation of these genomic insights into clinical practice has been limited by extensive cost and a lack of access to sophisticated methods for accurate and expedient subgroup/subtype analyses.[3]

Radiomics is an emerging discipline that can link imaging features to tumor genotype and serves as a promising approach to identify surrogate biomarkers that can accurately reflect tumor genomics.[10] Radiomic strategies have been extensively investigated in multiple cancer types, including non-small cell lung cancer,[11] glioblastoma,[12-16] hepatocellular carcinoma,[17] prostate cancer,[18] and breast cancer.[19] However, few studies have applied radiomics to MB; in those that have, the focus has been on the qualitative characterization of these tumors on MR imaging.[20-26] Specifically, these studies have shown that tumor location and enhancement patterns differ across MB subgroups.[20-24,26] For example, group 3 and group 4 MBs often arise in the midline, SHH tumors occur most frequently in the cerebellar hemispheres, and WNT tumors occur in both the midline and the cerebellar peduncle/cerebellopontine angle cistern locations.[20,22-24] Moreover, absence of enhancement is predictive of group 4 tumors,[23] while extensive enhancement in non-WNT/SHH tumors is predictive of poorer overall and event-free survival.[21] While qualitative image features of MB subgroups can provide useful clinical insight, they are subject to interobserver variability and do not capture all the multidimensional data that are acquired by MR imaging.

To date, the use of a quantitative imaging approach for the predictive analysis of MB subgroups has not yet been well-developed. In this multi-institutional study, we aimed to develop and validate radiomic and machine-learning methods to identify computational MR image signatures that are predictive of distinct molecular subgroups of MB. The discovery and establishment of noninvasive and surrogate imaging markers of MB subgroups can provide clinicians with a window into the genomics of these tumors, which can ultimately be helpful for clinical prognostication and informing management.

MATERIALS AND METHODS

Patients

This multicenter retrospective study was approved by the institutional review board or research ethics board from each of the 3 participating academic institutions: Lucile Packard Children's Hospital (Stanford University, Palo Alto, California), Boston Children's Hospital (Boston, Massachusetts), and the Hospital for Sick Children (Toronto, Ontario, Canada). Because this was a retrospective study, informed consent was waived. Interinstitutional data agreement was obtained for data-sharing. All patients with de novo and histologically confirmed MBs were identified from the medical record data base of each institution from January 2001 to January 2014. These patients were further screened using the following inclusion criteria: availability of high-quality preoperative MR imaging as determined by experienced pediatric neuroradiologists, neurosurgeons, and neuro-oncologists and the availability of molecular subgroup information or the availability of tumor tissue for molecular subtyping. A total of 109 patients were included across the 3 institutions (Lucile Packard Children's Hospital, $n = 32$; Boston Children's Hospital, $n = 28$; the Hospital for Sick Children, $n = 49$), comprising 64 males and 45 females; mean age, 8.56 ± 5.75 years; range, 1–18 years (Table 1). Clinicopathologic information including age, sex, histology diagnosis, and molecular subgroups, if available, was obtained from the medical record.

Molecular Analysis

Four distinct MB molecular subgroups (WNT, SHH, group 3, and group 4) were identified on the basis of gene-expression profiling using a nanoString-based assay (http://www.nanostring.com) as previously described.[27] For most patients, molecular analysis was performed with formalin-fixed paraffin-embedded tissue that was obtained at the time of surgical diagnosis. A small number of patients had molecular subtyping based on frozen tissue.

MR Imaging Acquisition, Image Data Retrieval, and Image Segmentation

All patients from Lucile Packard Children's Hospital/Stanford University underwent brain MR imaging at 1.5T or 3T (Signa or

Discovery 750; GE Healthcare, Milwaukee, Wisconsin). MRIs were performed using the brain tumor protocol of the institution, which included 2D axial T2-weighted spin-echo (TR/TE, 2500–5600/80–110 ms; 4- to 5-mm slice thickness; 0- to 1.5-mm skip), 2D axial or sagittal precontrast T1-weighted spin-echo, and 2D axial gadolinium-enhanced T1-weighted spin-echo (TR/TE, 400–1000/8–21 ms; 2- to 5-mm slice thickness; 0- to 1.5-mm skip) sequences. Four patients had 3D T1-weighted spoiled gradient recalled-echo (TR/TE, 8/3 ms; 1-mm slice thickness; 0-mm skip) instead of 2D T1-weighted spin-echo imaging. Patients from Boston Children's Hospital all underwent 1.5T brain MR imaging (Signa; GE Healthcare). Sequences acquired included 2D axial T2-weighted spin-echo (TR/TE, 3000–5000/80–100 ms; 4- to 5-mm slice thickness; 0- to 1.5-mm skip), 2D axial precontrast T1-weighted spin-echo, and 2D axial gadolinium-enhanced T1-weighted spin-echo (TR/TE, 500–700/8–22 ms; 4- to 6-mm slice thickness; 0- to 1.5-mm skip). One patient had 3D T1-weighted spoiled gradient recalled-echo (TR/TE, 8/2 ms; 1.5-mm slice thickness; 0-mm skip) instead of 2D T1-weighted spin-echo imaging. Patients from the Hospital for Sick Children underwent brain MR imaging at 1.5T or 3T across various scanner vendors (Signa, GE Healthcare; Achieva, Philips Healthcare Best, the Netherlands; Avanto, Siemens, Erlangen, Germany). Sequences acquired included 2D axial T2-weighted fast spin-echo (TR/TE, 3000–6800/80–120 ms; 3- to 6-mm slice thickness; 0.5- to 2.5-mm skip) and, in contrast to the other 2 cohorts, 3D axial precontrast and 3D axial gadolinium-enhanced T1-weighted turbo or fast-field echo (TR/TE, 5–11/2–5 ms; 1- to 2-mm slice thickness; 0-mm skip) sequences were acquired in most patients. Seven patients had 2D T1-weighted fast spin-echo (TR/TE, 8/2 ms; 1.5-mm slice thickness; 0-mm skip) instead of 3D T1-weighted turbo or fast-field echo imaging. All MR imaging data were extracted from the PACS at each respective institution and were subsequently de-identified for compliance with the Health Insurance Portability and Accountability Act before any analyses occurred.

An experienced team of radiologists supervised by a board-certified neuroradiologist with ≥10 years of experience in pediatric neuro-oncology imaging (T.Y.P.) manually drew ROIs around the tumor margin via a quantitative imaging informatics platform, electronic Physician Annotation Device (ePAD, https://epad.stanford.edu). Radiologists annotated on each contiguous image slice where the tumor was present on T2-weighted and contrast-enhanced T1-weighted images. The presence of tumor enhancement was confirmed by comparing precontrast with postcontrast T1-weighted images. Areas of intrinsic T1-hyperintensity (T1 signal that was present on precontrast images and likely representative of blood) were excluded from the final tumor ROIs. Final and proper placement of ROIs was confirmed by another board-certified neuroradiologist with ≥10 years of experience in pediatric neuro-oncology imaging (K.W.Y.).

Radiomic Feature-Extraction Methodology

We developed a computational framework to capture a variety of phenotypic characteristics of tumor. A total of 590 MR imaging–based radiomic features were extracted from the ROIs on T2-weighted and contrast-enhanced T1-weighted MR images, respectively. The primary types of radiomic features included intensity-based histograms, tumor edge-sharpness, Gabor features, and local area integral invariant (LAII) (all features used are described in On-line Tables 1 and 2). The Daube on Histogram features were based on Daubechies wavelet decomposition. The Quantitative Image-Feature Engine[28] offers additional detailed definitions of the extracted radiomic features (On-line Table 1). The *z* score normalization was used on each feature to standardize the range of all image features.

Statistical Analysis

Statistical analysis was conducted with Python software (2.7.14, https://www.python.org/). A nonparametric Wilcoxon rank sum test was used for feature selection, and a support vector machine (SVM) classifier was used for prediction. Statistical significance levels were all 2-sided, with statistical significance set at $P < .05$. Receiver operating characteristic (ROC) curve analysis was used to perform prediction evaluation of each molecular subgroup of MB.

Feature Selection, Radiomics, and Machine Learning Approach

The feature-selection method was applied to select the most discriminative features within a 10-fold cross-validation evaluation strategy (see "Model Evaluation"). Specifically, we used the Wilcoxon rank sum test[29] on individual features and sorted them by the acquired *P* values. After cross-validation analysis, the top k (k = 5, 10, 15, 20, 30, 40, 50, 100, 200, and 300) features with smallest k *P* values were selected in the training set. We then assessed the predictive power of selected radiomic features on the validation set.

We applied the SVM classifier using a double 10-fold cross-validation strategy for testing the performance of the model in predicting the 4 main MB molecular subgroups. SVM tackles high-dimensional data classification by weighting features and the use of a Gaussian radial basis function kernel. During the training process, to avoid potential overfitting, we determined the optimal parameters of the SVM classifier and the optimal number of image features using an internal 10-fold cross-validation and tested them by a range of selected features (top 5, 10, 15, 20, 30, 40, 50, 100, 200, and 300 features). Next, the trained model with the best area under the receiver operating characteristic curve (AUC) value was used for testing unseen samples in an outer 10-fold cross-validation strategy to determine the test set performance.

Model Evaluation

Two validation schemes were incorporated to evaluate the predictive performance of extracted radiomic features. To determine the generalization accuracy of the predictive models, we first performed a double 10-fold cross-validation on a single dataset containing all 3 patient cohorts (Fig 1). To validate the model across different institutions, we next tested an evaluation strategy in which we trained the model using the combined dataset from 2 institutions; then, we tested the model on data from the third independent institution. This process was repeated 3 times with each institutional cohort serving once as the test set (Fig 1), allow-

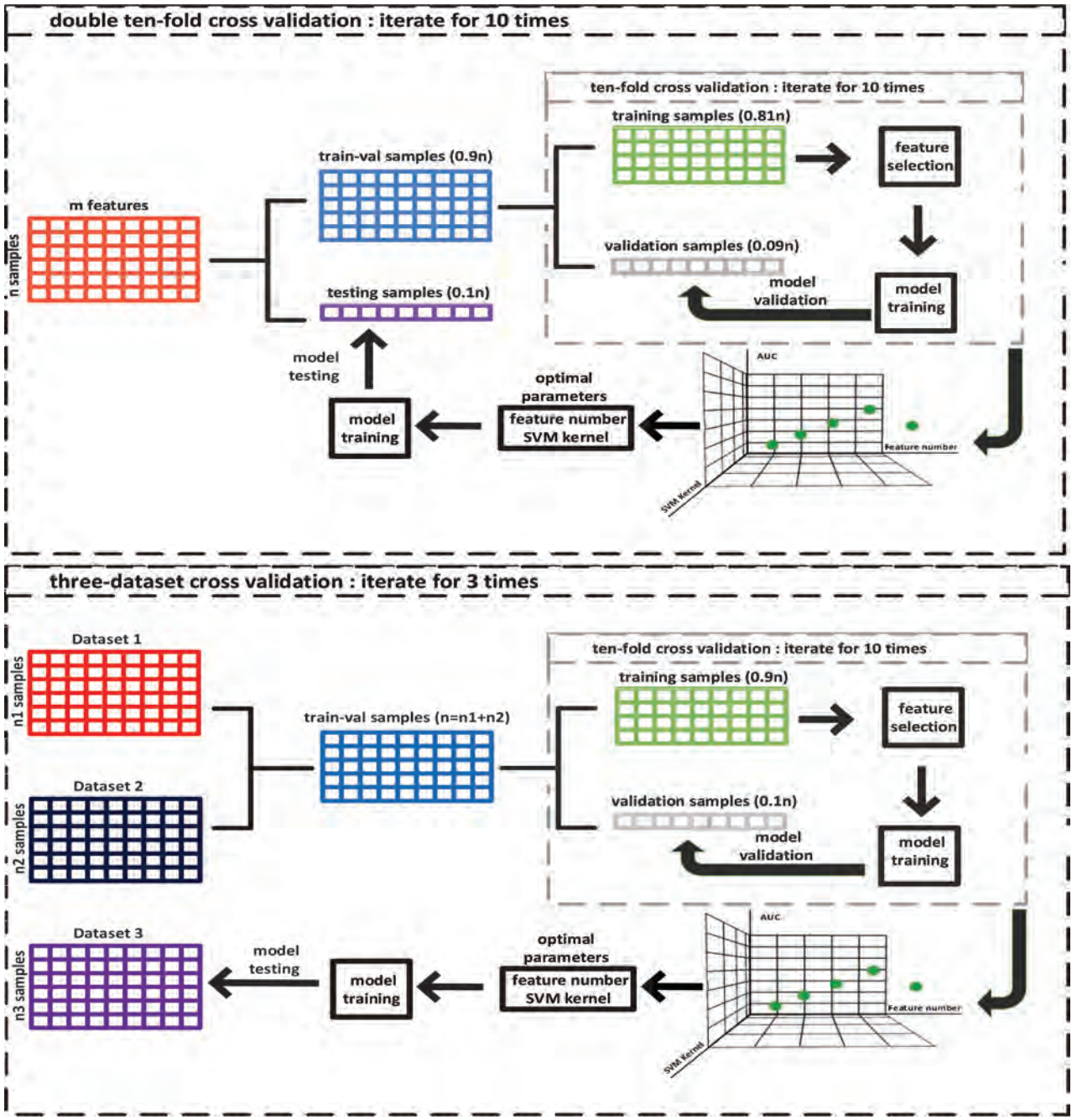


FIG 1. Illustration of 2 strategies used to evaluate the supervised machine learning models to predict the molecular subgroups of medulloblastoma. The upper and lower figures show details of double 10-fold cross-validation and 3-dataset cross-validation schemes, respectively.

ing us to evaluate truly predictive radiomic features across clinical sites with different vendors and imaging parameters. The overall model performance was assessed using the average of the 3 iterations and by determination of the AUC.

RESULTS

Model Evaluation

Table 2 summarizes the mean AUCs for prediction of the MB subgroups using the double 10-fold cross-validation and 3-dataset cross-validation strategies on solely T1-weighted, solely T2-weighted, and combined T1- and T2-weighted image datasets.

The double 10-fold cross-validation strategy, which combines all institutional cohorts into 1 dataset, showed that SVMs resulted in the best performance for predicting molecular subgroups. ROC analysis revealed superior performance of this model for predicting the SHH, group 3, and group 4 tumors, particularly when using extracted quantitative data from both T1- and T2-weighted images (AUC = 0.79, 0.70, and 0.83, respectively) (Fig 2). In contrast, the model was not strongly predictive of WNT tumors, despite using all the different image types (AUC = 0.45–0.63).

Fig 3 shows ROC curves of the best models for each MB subgroup obtained with our second validation scheme (independent 3-dataset cross-validation) (See On-line Fig 1 for all ROC curves). While not strongly predictive of group 3 tumors, the computa-

tional features extracted from T1-weighted images and the combined dataset from T1- and T2-weighted images were predictive of SHH (AUC = 0.73 and 0.70, respectively) and group 4 (AUC = 0.76 and 0.80, respectively) tumors (Table 2). In addition, while the mean AUC for predicting WNT tumors using T2-weighted images was good (0.72), there were institutional differences in performance (Stanford, AUC = 0.90; Boston, AUC = 0.49; Toronto, AUC = 0.76), suggesting that more training samples of the WNT group are needed to yield stable prediction outcomes.

Identification of Discriminative Radiomic Features

To identify discriminative radiomic features for predicting the 4 main molecular subgroups of MB (Fig 4) within our study population, we analyzed the results of selected features for all tested models. On-line Table 3 shows the best number of features for each institutional cohort and the number of overlapping features that was selected in all 3 cross-validation loops (see On-line Table 4 for a complete list of feature categories and the number of overlapped features in each category). We observed that the prediction of SHH is the most robust across all institutions because the optimal feature number for 3 cross-validation loops is the same (40 features), which represented a small subset of all 590 features (6.8%). Of all the features evaluated, there were 4 leading categories: lesion area, edge-sharpness, LAII, and histogram features (On-line Fig 2), with edge-sharpness features being the most important for predicting SHH and group 4.

Table 2: Predictive performance of 2 machine learning models for the identification of medulloblastoma molecular subgroups

MRI Dataset/Targeted Subgroup	AUC with Double 10-Fold Cross-Validation	AUC with 3-Dataset Cross-Validation
T1		
SHH	0.67	0.73
WNT	0.56	0.47
Group 3	0.40	0.54
Group 4	0.79	0.76
T2		
SHH	0.70	0.66
WNT	0.63	0.72
Group 3	0.51	0.57
Group 4	0.54	0.59
T1 + T2		
SHH	0.79	0.70
WNT	0.45	0.45
Group 3	0.70	0.39
Group 4	0.83	0.80

DISCUSSION

In this study, we developed and validated radiomic and machine learning approaches to identify individual categories of MR imaging–based radiomic features that predict distinct biologic subgroups of MB. Our first method using the double 10-fold cross-validation scheme allowed the prediction of SHH and group 4 tumors using combined information extracted from T1- and T2-weighted sequences, which are frequently acquired as part of the brain tumor MR imaging protocol of any institution. The second method using an independent 3-dataset cross-validation scheme showed the potential for applying our computational pipeline to datasets from outside institutions. In keeping with the results of our first method, this approach yielded a good predictive performance of SHH and group 4 tumors using combined T1 and T2 datasets. However, both models performed comparatively less robustly in predicting WNT and group 3 tumors, perhaps related to the lower amount of available imaging data for these specific subgroups and more molecular heterogeneity across group 3 tumors.

Several brain tumors are known to have spatial molecular[1,9,30,31] and imaging[13,22,23,25,26] heterogeneity. With regard to MB, the identification of 4 molecular subgroups in the past decade has deepened our understanding of the underlying biology of this tumor and the correlation of a specific tumor genotype with different clinical outcomes.[2,3,5] A recent study analyzing multiple biopsies within MB showed that a single biopsy can accurately and reliably subtype MB due to its spatially homogeneous transcriptomes, in contrast to the markedly heterogeneous genomic landscape of glioblastomas; however, actionable somatic mutations found in a single biopsy of MB were infrequently clonal across the entire tumor, which underscores the true molecular heterogeneity of this tumor.[9] In fact, Cavalli et al[32] have further identified 12 distinct subtypes within each of the 4 core MB subgroups, each with differing clinical presentations, prognosis, and copy-number mutations. Thus, the complex molecular heterogeneity of MB subgroups and the paucity of imaging data available for individual subgroups in this study may help to explain the performance of our models for predicting WNT and group 3 tumors. Additionally, prior studies have shown tumor location to be a unique factor, particularly for predicting WNT tumors,[22-24] and radiomic analysis of isolated tumor volume may be another explanation for our model performance. Thus, the incorporation of qualitative semantic features such as tumor location into our model may improve its performance, particularly for the prediction of WNT medulloblastomas, which is important clinically be-

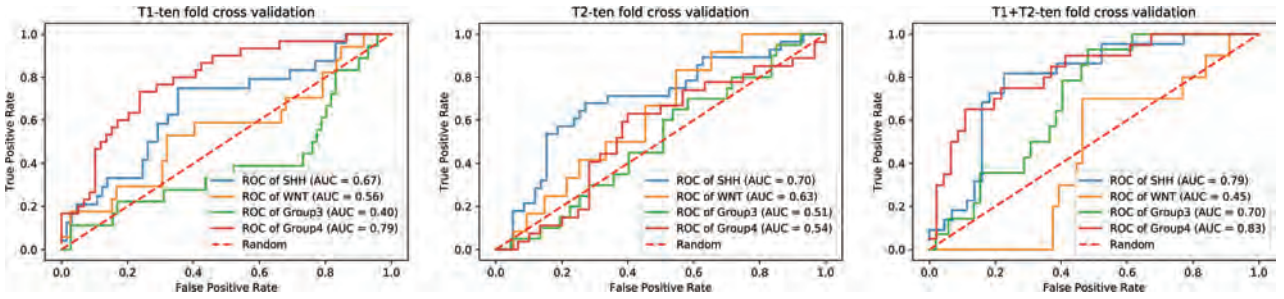


FIG 2. Receiver operating characteristic curves with a double 10-fold cross-validation scheme for support vector machine to predict the 4 main molecular subgroups of medulloblastoma with the use of computational MR imaging features.

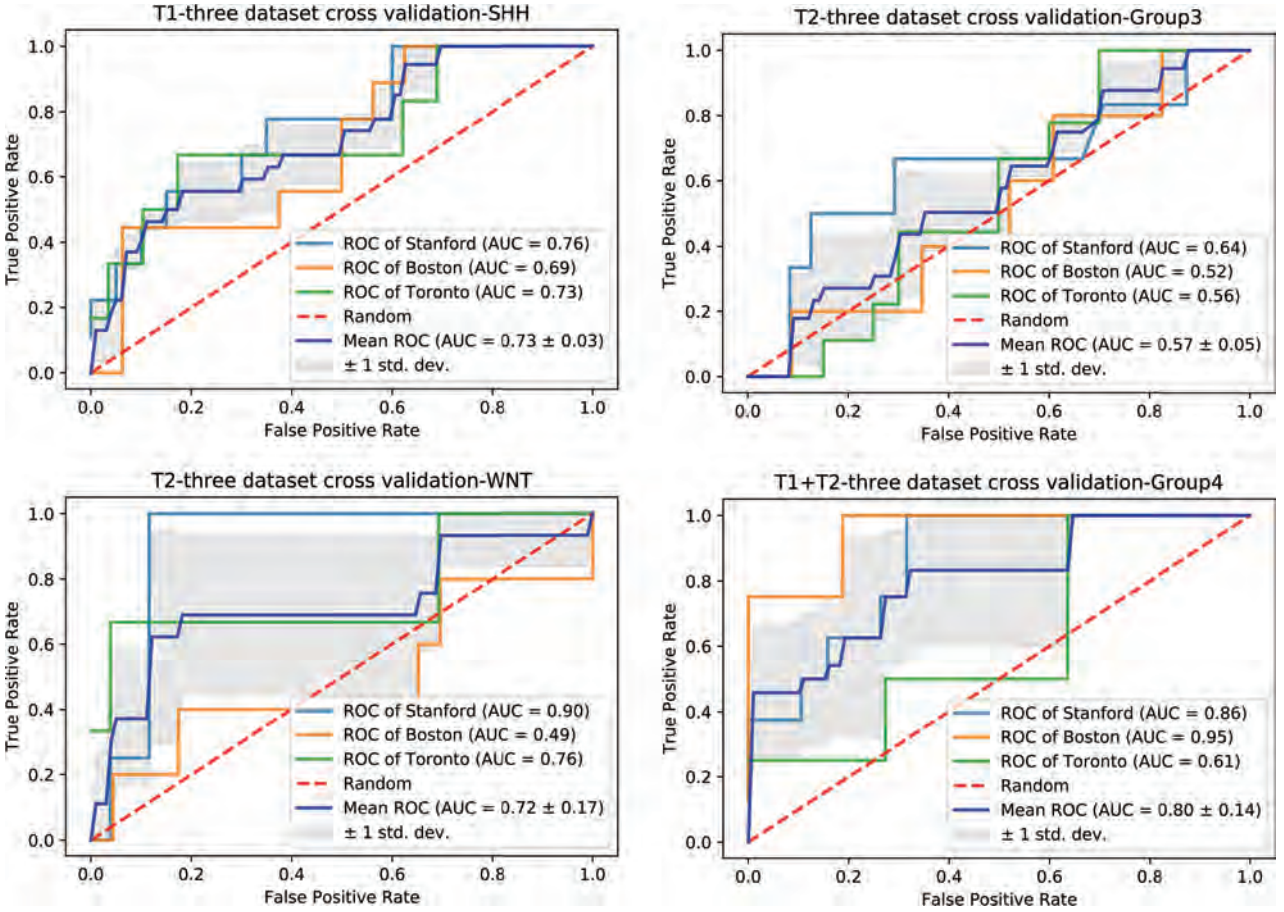


FIG 3. Receiver operating characteristic curves with the largest mean AUC values for 4 distinct molecular subgroups of medulloblastoma with a 3-dataset cross-validation scheme.

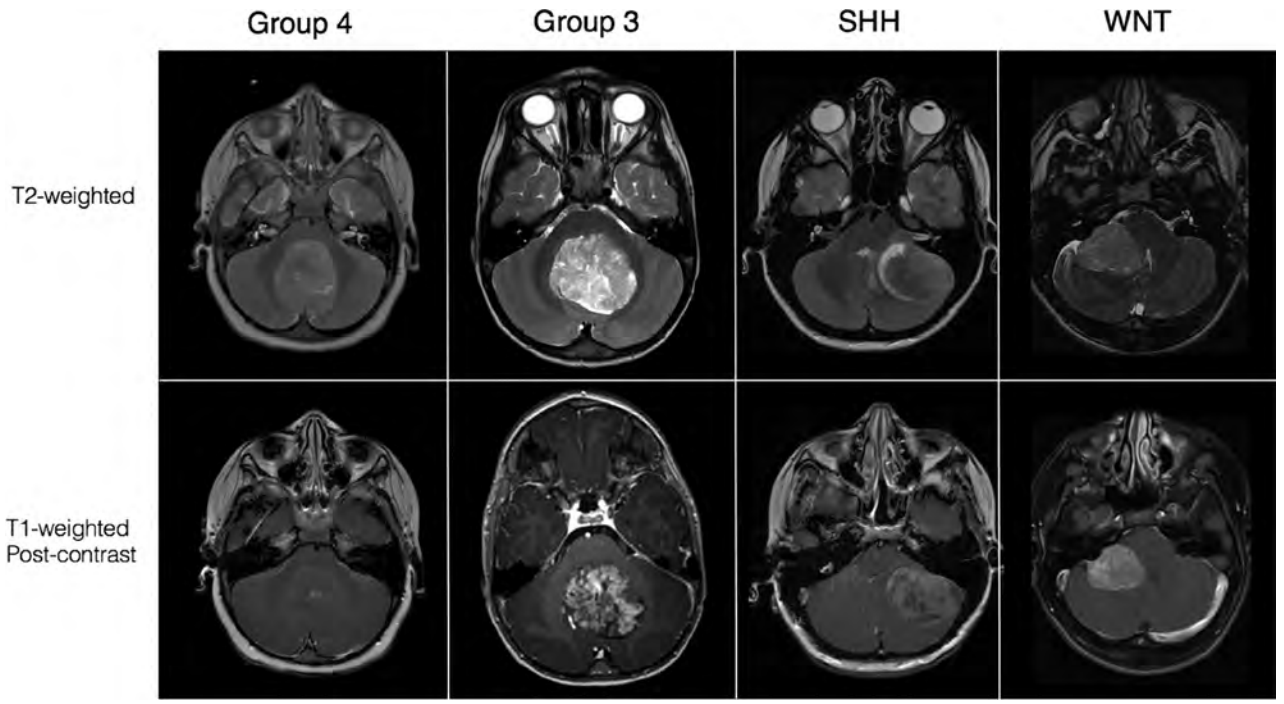


FIG 4. MR imaging appearance of the 4 core molecular subgroups of medulloblastoma on T2-weighted and contrast-enhanced T1-weighted images.

cause this subgroup is associated with the best prognosis and may not need the aggressive therapies used to treat other subgroups.[4]

Because MR imaging has the capacity to capture the structure and physiology of an entire tumor, it can be an invaluable tool for noninvasively evaluating tumoral genetic heterogeneity.[33] The spatial variations in genetic and molecular expression of MBs can manifest as different imaging phenotypes on MR imaging, with varying degrees of intratumoral enhancement, hemorrhage, and signal intensity on T1- and T2-weighted images.[16,21-23,25,26] Radiomic studies, most of which have focused on adult glioblastomas, have shown success in linking quantitative imaging features with key mutations as well as clinical outcomes.[12-14,34-36] For example, 1 study proposed that a distinct glioblastoma subtype, specifically, a rim-enhancing cluster found to upregulate the vascular endothelial growth factor receptor signaling pathway, is more likely to respond to upfront antiangiogenic therapy.[12] Many genetic factors (eg, *MYC*, *MYCN*, *OTX2*, *CDK6*, *SNCAIP*, and *ACVR1*) contribute to the 4 main MB subgroups and associated prognostic differences; even within the SHH and group 3 subgroups, more granular sub-subgroups have emerged with significant differences in the rate of metastases and 5-year survival.[32,37] Given molecular complexities that pose challenges in MB subclassification, there is an important future role for a high-performance, image-based biomarker that either predicts such unique molecular groups or subgroups of MB or provides a more robust tumor risk-stratification scheme for treatment decision-making independent of molecular grouping or subgrouping.

Furthermore, a rapid, low-risk, and inexpensive platform for classifying MB that is feasible with radiomic and machine learning algorithms can potentially enable more widespread tumor subtyping in clinical institutions that may have limited histopathologic and genomic resources. While immunohistochemistry markers (*GAB1*, β-catenin, filamin A, and *YAP1*) are currently used in some institutions to identify SHH and WNT tumors, identifying specific group 3 or group 4 tumors remains expensive and difficult without the application of gene expression or methylation profiling.[38,39] In this study, we show a relatively high performance for predicting group 4 MB that is feasible with a computational analysis scheme.

This study has limitations. Because this was a retrospective and multi-institutional study, there was heterogeneity in MR image data, including the use of different scanner vendors and imaging parameters. However, in clinical practice, different scanner vendors at different field strengths and different imaging protocols are used daily for tumor diagnosis and surveillance; thus, a predictive model that incorporates such technical variations results in a more practical clinical translation of radiogenomic strategies. A recent study showed that radiomic features varied considerably on T1-weighted images generated by different pulse sequences and parameters.[40] In this study, we chose to retain differences in imaging protocols from 3 different cohorts to assess the robustness of radiomic features extracted from multi-institutional data. To facilitate evaluation of classifiers, after feature extraction, we performed feature-level normalization (z score) across patients to help with the predictive performance of the machine learning classifier. Future studies may need to look into other strategies to overcome the heterogeneity of data, including normalizing the degree of T1-weighting (eg, normalizing images to the signal intensity of the tumor) and weighting the importance of specific features or image sequences (eg, T1 versus T2) as input in a predictive radiogenomic model.[40] In addition, while this study used T1 contrast-enhanced and T2-weighted images for feature discrimination and model development, incorporating additional image sequences, such as diffusion, permeability, or T2* perfusion, could further boost model performance. Despite these challenges, our study showed that radiomic strategies can be used to extract discriminating computational features and create a machine learning–based prediction model for pediatric MB subgroups.

CONCLUSIONS

We present proof-of-concept results for the application of radiomics and machine learning using multi-institutional data for the prediction of distinct MB molecular subgroups. High-throughput quantitative features were extracted from contrast-enhanced T1- and T2-weighted images and linked to 4 core subgroups of MB. Model performance for the prediction of SHH and group 4 was more robust than for WNT and group 3. Future investigations using a larger sample size for all subgroups, particularly WNT (because we had the least amount of WNT cases in this study), is needed to improve classifier training and evaluation during cross-validation. The use of other imaging sequences such as diffusion-weighted, permeability, and T2* perfusion imaging may also yield additional radiomic features and help to improve performance. Computational analyses of MR imaging offer a wealth of opportunities to noninvasively characterize tumors, which can have an important role in the clinical and treatment decision-making processes for pediatric MB.

Disclosures: Paul G. Fisher—*UNRELATED*: *Employment*: Elsevier Publishing, *Comments*: stipend of $22,500 per year for work as Associate Editor for the *Journal of Pediatrics*; *Grants/Grants Pending*: National Institutes of Health, *Comments*: salary support from the National Cancer Institute (Pediatric Brain Tumor Consortium) and from the National Human Genome Research Institute (Undiagnosed Diseases Consortium).* Vijay Ramaswamy—*RELATED*: *Grant*: collaborative ependymoma research network. Tina Young Poussaint—*UNRELATED*: *Grants/Grants Pending*: National Institutes of Health, Pediatric Brain Tumor Consortium Neuroimaging Center.* Olivier Gevaert—*RELATED*: *Grant*: National Institutes of Biomedical Imaging and Bioengineering, *Comments*: Research reported in this publication was supported by the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health under Award Number R01EB020527. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health*; *UNRELATED*: *Grants/Grants Pending*: National Institutes of Health, *Comments*: support from various grants from the National Institutes of Health, and foundation grant from the Innovation in Cancer Informatics fund.* *Money paid to the institution.

REFERENCES

1. Pollack IF, Jakacki RI. **Childhood brain tumors: epidemiology, current management and future directions.** *Nat Rev Neurol* 2011;7:495–506 CrossRef Medline
2. Kool M, Korshunov A, Remke M, et al. **Molecular subgroups of medulloblastoma: an international meta-analysis of transcriptome, genetic aberrations, and clinical data of WNT, SHH, group 3, and group 4 medulloblastomas.** *Acta Neuropathol* 2012;123:473–84 CrossRef Medline
3. Northcott PA, Korshunov A, Witt H, et al. **Medulloblastoma comprises four distinct molecular variants.** *J Clin Oncol* 2011;29:1408–14 CrossRef Medline
4. Taylor MD, Northcott PA, Korshunov A, et al. **Molecular subgroups**

of medulloblastoma: the current consensus. *Acta Neuropathol* 2012; 123:465–72 CrossRef Medline
5. Ramaswamy V, Remke M, Bouffet E, et al. **Risk stratification of childhood medulloblastoma in the molecular era: the current consensus.** *Acta Neuropathol* 2016;131:821–31 CrossRef Medline
6. Louis DN, Perry A, Reifenberger G, et al. **The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary.** *Acta Neuropathol* 2016;131:803–20 CrossRef Medline
7. Gajjar AJ, Robinson GW. **Medulloblastoma-translating discoveries from the bench to the bedside.** *Nat Rev Clin Oncol* 2014;11:714–22 CrossRef Medline
8. Archer TC, Mahoney EL, Pomeroy SL. **Medulloblastoma: molecular classification-based personal therapeutics.** *Neurotherapeutics* 2017; 14:265–73 CrossRef Medline
9. Morrissy AS, Cavalli FM, Remke M, et al. **Spatial heterogeneity in medulloblastoma.** *Nat Genet* 2017;49:780–88 CrossRef Medline
10. Zhou M, Scott J, Chaudhury B, et al. **Radiomics in brain tumor: image assessment, quantitative feature descriptors, and machine-learning approaches.** *AJNR Am J Neuroradiol* 2018;39:208–16 CrossRef Medline
11. Gevaert O, Xu J, Hoang CD, et al. **Non-small cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray data–methods and preliminary results.** *Radiology* 2012;264:387–96 CrossRef Medline
12. Itakura H, Achrol AS, Mitchell LA, et al. **Magnetic resonance image features identify glioblastoma phenotypic subtypes with distinct molecular pathway activities.** *Sci Transl Med* 2015;7:303ra138 CrossRef Medline
13. Gevaert O, Mitchell LA, Achrol AS, et al. **Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features.** *Radiology* 2014;273:168–74 CrossRef Medline
14. Kickingereder P, Bonekamp D, Nowosielski M, et al. **Radiogenomics of glioblastoma: machine learning-based classification of molecular characteristics by using multiparametric and multiregional MR imaging features.** *Radiology* 2016;281:907–18 CrossRef Medline
15. Jamshidi N, Diehn M, Bredel M, et al. **Illuminating radiogenomic characteristics of glioblastoma multiforme through integration of MR imaging, messenger RNA expression, and DNA copy number variation.** *Radiology* 2014;270:1–2 CrossRef Medline
16. Zhou M, Chaudhury B, Hall LO, et al. **Identifying spatial imaging biomarkers of glioblastoma multiforme for survival group prediction.** *J Magn Reson Imaging* 2017;46:115–23 CrossRef Medline
17. Banerjee S, Wang DS, Kim HJ, et al. **A computed tomography radiogenomic biomarker predicts microvascular invasion and clinical outcomes in hepatocellular carcinoma.** *Hepatology* 2015;62:792–800 CrossRef Medline
18. Stoyanova R, Takhar M, Tschudi Y, et al. **Prostate cancer radiomics and the promise of radiogenomics.** *Transl Cancer Res* 2016;5:432–47 CrossRef Medline
19. Grimm LJ. **Breast MRI radiogenomics: current status and research implications.** *J Magn Reson Imaging* 2016;43:1269–78 CrossRef Medline
20. Gibson P, Tong Y, Robinson G, et al. **Subtypes of medulloblastoma have distinct developmental origins.** *Nature* 2010;468:1095–99 CrossRef Medline
21. Lastowska M, Jurkiewicz E, Trubicka J, et al. **Contrast enhancement pattern predicts poor survival for patients with non-WNT/SHH medulloblastoma tumours.** *J Neurooncol* 2015;123:65–73 CrossRef Medline
22. Patay Z, DeSain LA, Hwang SN, et al. **MR imaging characteristics of wingless-type-subgroup pediatric medulloblastoma.** *AJNR Am J Neuroradiol* 2015;36:2386–93 CrossRef Medline
23. Perreault S, Ramaswamy V, Achrol AS, et al. **MRI surrogates for molecular subgroups of medulloblastoma.** *AJNR Am J Neuroradiol* 2014;35:1263–69 CrossRef Medline
24. Teo WY, Shen J, Su JM, et al. **Implications of tumor location on subtypes of medulloblastoma.** *Pediatr Blood Cancer* 2013;60: 1408–10 CrossRef Medline
25. Yeom KW, Mobley BC, Lober RM, et al. **Distinctive MRI features of pediatric medulloblastoma subtypes.** *AJR Am J Roentgenol* 2013; 200:895–903 CrossRef Medline
26. Zhao F, Li C, Zhou Q, et al. **Distinctive localization and MRI features correlate of molecular subgroups in adult medulloblastoma.** *J Neurooncol* 2017;135:353–60 CrossRef Medline
27. Northcott PA, Shih DJ, Remke M, et al. **Rapid, reliable, and reproducible molecular sub-grouping of clinical medulloblastoma samples.** *Acta Neuropathol* 2012;123:615–26 CrossRef Medline
28. Echegaray S, Bakr S, Rubin DL, et al. **Quantitative Image Feature Engine (QIFE): an open-source, modular engine for 3D quantitative feature extraction from volumetric medical images.** *J Digit Imaging* 2018;31:403–14 CrossRef Medline
29. Wilcoxin F. **Probability tables for individual comparisons by ranking methods.** *Biometrics* 1947;3:119–22 CrossRef Medline
30. Pajtler KW, Witt H, Sill M, et al. **Molecular classification of ependymal tumors across all CNS compartments, histopathological grades, and age groups.** *Cancer Cell* 2015;27:728–43 CrossRef Medline
31. Northcott PA, Buchhalter I, Morrissy AS, et al. **The whole-genome landscape of medulloblastoma subtypes.** *Nature* 2017;547:311–17 CrossRef Medline
32. Cavalli FM, Remke M, Rampasek L, et al. **Intertumoral heterogeneity within medulloblastoma subgroups.** *Cancer Cell* 2017;31:737–54.e6 CrossRef Medline
33. Gatenby RA, Grove O, Gillies RJ. **Quantitative imaging in cancer evolution and ecology.** *Radiology* 2013;269:8–15 CrossRef Medline
34. Kickingereder P, Neuberger U, Bonekamp D, et al. **Radiomic subtyping improves disease stratification beyond key molecular, clinical and standard imaging characteristics in patients with glioblastoma.** *Neuro Oncol* 2018;20:848–57 CrossRef Medline
35. Grossmann P, Narayan V, Chang K, et al. **Quantitative imaging biomarkers for risk stratification of patients with recurrent glioblastoma treated with bevacizumab.** *Neuro Oncol* 2017;19:1688–97 CrossRef Medline
36. Hu LS, Ning S, Eschbacher JM, et al. **Radiogenomics to characterize regional genetic heterogeneity in glioblastoma.** *Neuro Oncol* 2017; 19:128–37 CrossRef Medline
37. Northcott PA, Shih DJ, Peacock J, et al. **Subgroup-specific structural variation across 1,000 medulloblastoma genomes.** *Nature* 2012;488: 49–56 CrossRef Medline
38. Lastowska M, Trubicka J, Niemira M, et al. **Medulloblastoma with transitional features between group 3 and group 4 is associated with good prognosis.** *J Neurooncol* 2018;138:231–40 CrossRef Medline
39. Ellison DW, Dalton J, Kocak M, et al. **Medulloblastoma: clinicopathological correlates of SHH, WNT, and non-SHH/WNT molecular subgroups.** *Acta Neuropathol* 2011;121:381–96 CrossRef Medline
40. Ford J, Dogan N, Young L, et al. **Quantitative radiomics: impact of pulse sequence parameter selection on MRI-based textural features of the brain.** *Contrast Media & Molecular Imaging* 2018. https://doi.org/10.1155/2018/1729071. Accessed August 16, 2018

Early Diagnosis of Spastic Cerebral Palsy in Infants with Periventricular White Matter Injury Using Diffusion Tensor Imaging

H. Jiang, X. Li, C. Jin, M. Wang, C. Liu, K.C. Chan, and J. Yang

ABSTRACT

BACKGROUND AND PURPOSE: Periventricular white matter injury is the common cause of spastic cerebral palsy. However, the early diagnosis of spastic cerebral palsy still remains a challenge. Our aim was to investigate whether infants with periventricular white matter injury with bilateral spastic cerebral palsy have unique lesions different from those in infants without cerebral palsy and to evaluate the efficiency of DTI in the early diagnosis of spastic cerebral palsy.

MATERIALS AND METHODS: Infants with periventricular white matter injury and controls underwent MR imaging at 6–18 months of age. Fractional anisotropy was calculated from DTI. Cerebral palsy was diagnosed by 24–30 months of age. Subjects were divided into 3 groups: infants with periventricular white matter injury with bilateral spastic cerebral palsy, infants with periventricular white matter injury without cerebral palsy, and controls. Tract-Based Spatial Statistics and Automated Fiber Quantification were used to investigate intergroup differences. Receiver operating characteristic curves were used to assess the diagnostic accuracy of spastic cerebral palsy. Correlations between motor function scores and fractional anisotropy were evaluated along white matter tracts.

RESULTS: There were 20, 19, and 33 subjects in periventricular white matter injury with spastic cerebral palsy, periventricular white matter injury without cerebral palsy, and control groups, respectively. Decreased fractional anisotropy in the corticospinal tract was only observed in infants with periventricular white matter injury with spastic cerebral palsy, whereas decreased fractional anisotropy in the posterior thalamic radiation and genu and splenium of the corpus callosum was seen in both periventricular white matter injury subgroups. Fractional anisotropy in the corticospinal tract at the internal capsule level was effective in differentiating infants with periventricular white matter injury with spastic cerebral palsy from those without cerebral palsy by a threshold of 0.53, and it had strong correlations with motor function scores.

CONCLUSIONS: Corticospinal tract lesions play a crucial role in motor impairment related to spastic cerebral palsy in infants with periventricular white matter injury. Fractional anisotropy in the corticospinal tract at the internal capsule level could aid in the early diagnosis of spastic cerebral palsy with high diagnostic accuracy.

ABBREVIATIONS: CP = cerebral palsy; CST = corticospinal tract; CST-CP = CST at the cerebral peduncle level; CST-CR = CST at the corona radiata level; CST-IC = CST at the internal capsule level; FA = fractional anisotropy; GCC = genu of the corpus callosum; GMFCS = Gross Motor Function Classification System; PTR = posterior thalamic radiation; PWMI = periventricular white matter injury; SCC = splenium of the corpus callosum; SCP = spastic cerebral palsy

Periventricular white matter injury (PWMI) is a major form of white matter injury in both preterm and term infants and is the common cause of spastic cerebral palsy (SCP).[1-3] MR imaging is effective in identifying PWMI and demonstrates periventricular white matter signal abnormality and/or volume loss, enlargement of the lateral ventricles, and thinning of the corpus callosum.[3] On the basis of signs on MR imaging, PWMI can be classified into mild,

Received June 6, 2018; accepted October 30.

From the Department of Radiology (H.J., X.L., C.J., M.W., C.L., J.Y.), First Affiliated Hospital, and Department of Biomedical Engineering (H.J., J.Y.), Key Laboratory of Biomedical Information Engineering of the Ministry of Education, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, China; and Department of Ophthalmology and Radiology (K.C.C.), School of Medicine, New York University, New York, New York.

Haoxiang Jiang and Xianjun Li contributed equally to this work.

This study was supported by the National Key Research and Development Program of China (2016YFC0100300); the National Natural Science Foundation of China (81471631, 81771810 and 81171317); the 2011 New Century Excellent Talent Support Plan of the Ministry of Education, China (NCET-11-0438); the Fundamental Research Funds for the Central Universities (xjj2018265); and the Fundamental Research Funds of the First Affiliated Hospital of Xi'an Jiaotong University (2017QN-09).

Preliminary data from this research were previously presented at: Annual Meeting of the International Society for Magnetic Resonance in Medicine, April 22–27, 2017; Honolulu, Hawaii.

Please address correspondence to Jian Yang, PhD, Department of Radiology, First Affiliated Hospital, Xi'an Jiaotong University, Xi'an, Shaanxi, China; e-mail: cjr.yangjian@vip.163.com

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

Indicates article with supplemental on-line tables.

Indicates article with supplemental on-line photos.

http://dx.doi.org/10.3174/ajnr.A5914

moderate, and severe grades.[4] Severe PWMI develops frequently into cerebral palsy (CP).[4,5] However, the early diagnosis of SCP in infants with mild and moderate PWMI still remains a major challenge.

MR imaging and neurologic examinations are predictive tools for detecting the risk of CP.[6-8] Compared with neurologic examinations, MR imaging is preferred for revealing the anatomic position and severity of brain lesions.[6] However, conventional MR imaging has limitations in delineating white matter tracts precisely and has failed to differentiate individual tracts specifically.[9] DTI is effective in overcoming the above limitations and improving the diagnostic accuracy.[6] Patients with SCP typically present with motor deficits. DTI studies have revealed that the damage in the corticospinal tract (CST) is related to motor impairment.[10-12] Meanwhile, the sensory pathways or commissural tract lesions or both are also involved in the motor deficits in children with SCP.[13-15] However, the prerequisite white matter lesions accountable for SCP remain unclear.[16] The exploration of the differences in white matter alterations between infants with PWMI with and without CP may provide clues for the identification of the responsible white matter tracts related to motor impairment.

In this study, the analyses of DTI data with Tract-Based Spatial Statistics (TBSS; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/TBSS) and Automated Fiber Quantification (AFQ; https://pypi.org/project/AFQ-Browser/) were used to compare the differences among groups of infants with PWMI with bilateral SCP, infants with PWMI without CP, and controls. The aim was to investigate whether infants with PWMI with bilateral SCP have unique lesions different from lesions in those without CP and to evaluate the early diagnostic efficiency of DTI for differentiating infants with SCP from those with PWMI without CP.

MATERIALS AND METHODS

This was a retrospective cohort study approved by the institutional review board of the First Affiliated Hospital of Xi'an Jiaotong University.

Participants

Parents of the infants were informed of the potential risks of MR imaging. Written consent was obtained before the brain MR imaging examination.

Brain MR imaging was performed in infants 6–18 months of age, from July 2011 to January 2014. The reasons for MR imaging examination included suspected developmental delay, seizures, or other risks of cerebral disorders. The MR imaging scanner and sequences did not change throughout the study period. Next, these infants underwent follow-up neurologic examinations between 24 and 30 months of age, performed by a pediatric neurologist using several measurements, including motor milestones, primitive reflexes, postural reactions, muscle strength, and cognitive and psychomotor development assessment using the second edition of the Bayley Scales of Infant Development and the Gross Motor Function Classification System (GMFCS; https://research.cerebralpalsy.org.au/what-is-cerebral-palsy/severity-of-cerebral-palsy/gross-motor-function-classification-system/) assessment. Developmental delay was defined by a score of the Mental or Psychomotor Development Index of <85 according to the Bayley Scales of Infant Development assessment.

CP was diagnosed in infants on the basis of the follow-up examination results using the definition provided by the International Executive Committee in the United States.[17] The infants with CP were classified into GMFCS levels I–V, representing mild-to-severe motor dysfunction.[10] The GMFCS level zero represented infants without motor dysfunction in this study.

The inclusion criteria of infants with PWMI were as follows: 1) MR imaging performed between 6 and 18 months of age and 2) PWMI diagnosed by MR imaging. The severity of PWMI was graded as mild, moderate, or severe on the basis of the MR imaging signs as follows: 1) mild, abnormally high signal in the periventricular white matter on T2-weighted images with mild white matter reduction limited to the peritrigonal region; 2) moderate, abnormally high white matter signal on T2-weighted images, with moderate periventricular white matter decrease and irregular enlargement of the ventricles; 3) severe, large, or extensive cystic changes of periventricular white matter, with marked reduction of white matter and severe irregular enlargement of the ventricles (On-line Fig 1).[4] The exclusion criteria included severe PWMI, intrauterine congenital infection, and unsuccessful follow-up neurologic examinations. According to the follow-up outcomes, infants with mild and moderate PWMI were divided into those with PWMI with bilateral SCP and those without CP after the exclusion of unilateral and nonspastic CP.

The infants with MR imaging performed between 6 and 18 months of age were selected as controls after excluding the following conditions: 1) abnormalities on MR imaging, including PWMI; dilated perivascular spaces; arachnoid cyst; hydrocephalus; congenital malformations; hematencephalon; subarachnoid space enlargement; encephalomalacia and so forth; 2) history of nervous system infection; preterm birth; hypoxic-ischemic encephalopathy and congenital heart disease; 3) seizures; developmental delay; and hypertonia; and 4) lack of follow-up neurologic examinations.

MR Imaging Protocols

MR imaging was performed on a 3T scanner (Signa HDxt; GE Healthcare, Milwaukee, Wisconsin). All infants were required to sleep soundly and wear sponge earplugs for hearing protection. The infants who could not remain still were sedated with 10% chloral hydrate (25–50 mg/kg) to reduce motion artifacts. The potential risks of chloral hydrate were fully explained to the parents. The patient selection, monitoring, and management were performed in strict compliance with the guidelines for monitoring and management of pediatric patients before, during, and after sedation for diagnostic and therapeutic procedures.[18] Heart rate, transcutaneous oxygen saturation, and respiration rate were monitored throughout the procedure. All infants underwent the following MR imaging protocols: 3D T1-weighted imaging (TR = 10.2 ms, TE = 4.6 ms, slice thickness = 1 mm), T2-weighted imaging (TR = 6500 ms, TE = 124 ms, slice thickness = 4 mm), T2 FLAIR imaging (TR = 9600 ms, TE = 110 ms, TI = 2400 ms, slice thickness = 4 mm), and DTI. DTI was performed using a spin-echo echo-planar imaging sequence. The parameters for DTI were as follows: TR = 5500 ms, TE = 95 ms, flip angle = 90°, slice thickness = 4 mm without gap, FOV = 180 × 180 mm^2, matrix = 128 × 128, b-values = 0, 1000 s/mm^2 with 35 gradient directions.

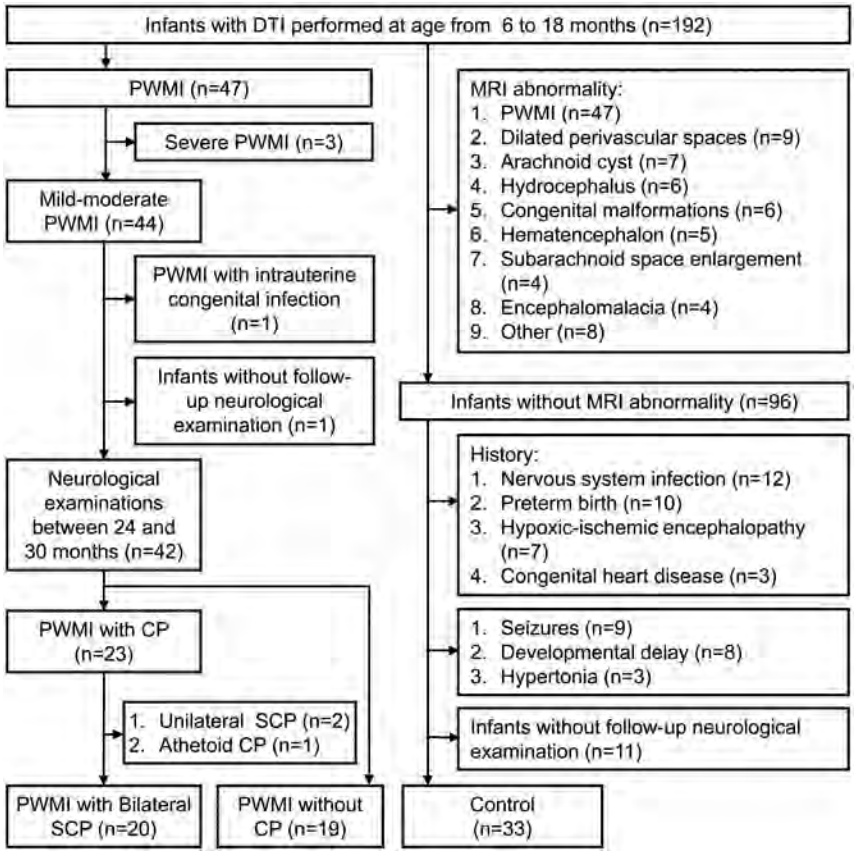


FIG 1. Flow chart of study participants.

Image Postprocessing and Data Analysis

Fractional anisotropy (FA) maps were obtained after brain extraction and eddy current correction using the FMRIB Software Library (FSL; http://www.fmrib.ox.ac.uk/fsl).[19] Linear and nonlinear image registrations were used for alignment of the FA maps of all subjects to a selected FA map. An averaged image of the coregistered FA maps was created as the target map. Then, FA maps of all subjects were registered to the target map. A mean FA map and mean FA skeleton were created. The aligned FA map of each subject was projected onto the mean FA skeleton (threshold = 0.2). Voxelwise statistical analysis was performed to assess the intergroup differences in FA.

AFQ software was used to quantitatively analyze white matter tracts.[20] Automatic fiber tract segmentation and cleaning were performed first. Then, FA values along the tracts were calculated by clipping each fiber to 100 equally spaced nodes. The CST, posterior thalamic radiation (PTR), genu of the corpus callosum (GCC), and splenium of corpus callosum (SCC) were evaluated. To characterize the properties of the CST in detail, we divided the CST into 3 parts according to the brain atlas[21]: CST at the cerebral peduncle level (CST-CP; nodes: 1–37), CST at the internal capsule level (CST-IC; nodes: 38–81), and CST at the corona radiata level (CST-CR; nodes: 82–100).

Statistical Analysis

Categoric variables were analyzed using the χ^2 test in SPSS software (Version 17.0; IBM, Armonk, New York). Continuous variables were analyzed across groups using analysis of variance with the least significant difference test. Sex ratios were compared across tracts by using the χ^2 test. Interrater reliability analyses for GMFCS were performed with the κ test. The variables of gestational age, birth weight, and age at MR imaging were adjusted during the intergroup comparisons and correlation analyses to remove their potential effects on the variations in DTI metrics. Intergroup comparisons in FA were performed by TBSS with the threshold-free cluster enhancement and family-wise error correction. In the AFQ analysis, a Mann-Whitney *U* test with a Bonferroni correction was used to assess the intergroup differences in FA. Receiver operating characteristic curves were used to assess the diagnostic performances of FA in different regions for differentiating infants with PWMI with bilateral SCP from those without CP. Correlations between FA values of the white matter tracts and motor function scores (GMFCS levels and Psychomotor Development Index) in infants with PWMI were analyzed by partial correlation analysis. *P* values $< .05$ were considered statistically significant for all tests.

RESULTS

Demographic and Clinical Information

Thirty-nine infants with PWMI and 33 controls were finally enrolled on the basis of the inclusion and exclusion criteria (Fig 1). According to the follow-up diagnosis, the infants with PWMI were divided into infants with bilateral SCP ($n = 20$) and those without CP ($n = 19$) groups. The time between the MR imaging examination and diagnosis of SCP was 14.8 ± 3.2 months.

Among infants with PWMI with bilateral SCP, 7 (35%), 6 (30%), 4 (20%), and 3 (15%) infants were classified into GMFCS levels I, II, III, and IV, respectively. The κ value for the interrater reliability of the GMFCS assessment was 0.83. There was no significant difference in gestational age and birth weight between infants with PWMI with bilateral SCP and those without CP. The gestational age and birth weight in infants with PWMI with bilateral SCP and without CP were less than those in controls. Moreover, there was no significant difference across groups in the sex constituent ratio or age at MR imaging (Table 1). More details about the perinatal characteristics of the infants with PWMI are listed in On-line Table 1.

TBSS Analysis of Intergroup Differences

FA in widespread areas of white matter was lower in the infants with PWMI with bilateral SCP than that in the controls. In infants with PWMI without CP, decreased FA was found mainly in regions adjacent to the anterior horns, trigone, and posterior horns of the lateral ventricles. Infants with PWMI with bilateral SCP

Table 1: Demographics of infants with PWMI with bilateral spastic cerebral palsy, infants with PWMI without cerebral palsy, and controls[a]

	Infants with PWMI with SCP (*n* = 20)	Infants with PWMI without CP (*n* = 19)	Controls (*n* = 33)	*P* Values		
				PWMI with SCP vs PWMI without CP	PWMI with SCP vs Control	PWMI without CP vs Control
Sex (female/male)	8:12	6:13	11:22	.58	.62	.90
Gestational age (wk)	37.38 ± 1.90 (34.57–41.29)	38.07 ± 2.45 (34.00–41.43)	39.16 ± 1.14 (37.00–41.00)	.23	<.01	.04
Age at MRI (mo)	11.74 ± 2.13 (8.00–17.66)	11.56 ± 2.99 (6.40–16.76)	11.06 ± 2.38 (6.10–17.33)	.83	.34	.49
Birth weight (kg)	2.95 ± 0.47 (2.20–3.70)	3.06 ± 0.60 (1.80–3.85)	3.38 ± 0.50 (2.50–4.38)	.50	.01	.04

[a] Data in columns 2–4 are mean and standard deviation (range).

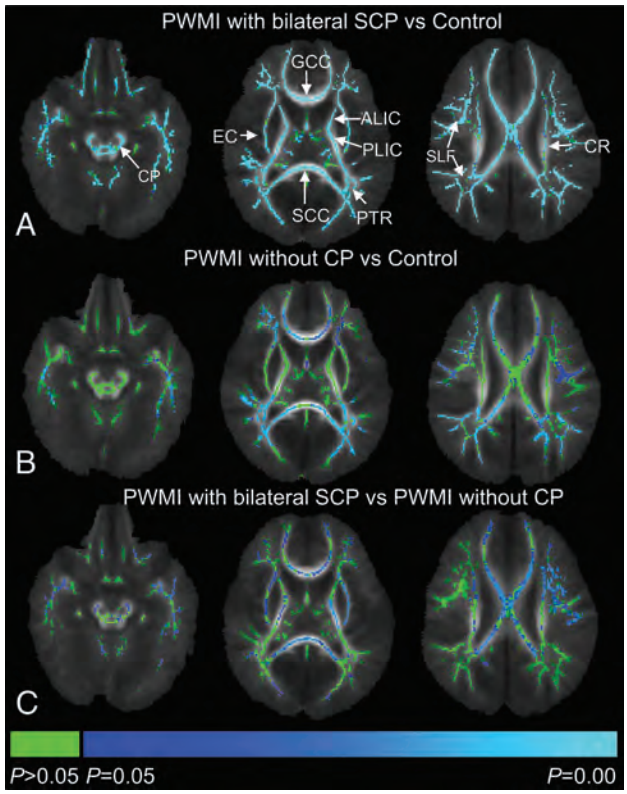


FIG 2. Intergroup comparisons of fractional anisotropy values using TBSS. Blue and light blue regions show major white matter tracts with significantly decreased FA ($P < .05$) in infants with PWMI with bilateral SCP (*A*), infants with PWMI without CP (*B*) relative to controls, and infants with PWMI with bilateral SCP relative to those without CP (*C*). Green represents regions without significant differences ($P > .05$). ALIC indicates anterior limb of the internal capsule; CR, corona radiata; EC, external capsule; PLIC, posterior limb of the internal capsule; SLF, superior longitudinal fasciculus.

showed significantly lower FA values than infants with PWMI without CP in the bilateral CST-CP, CST-IC, CST-CR, external capsule, GCC, and SCC (Fig 2).

AFQ Analysis of Intergroup Differences along Tracts

The tract profiles of infants with PWMI with bilateral SCP, infants with PWMI without CP, and controls are presented in Fig 3. Obvious deceases in FA values in infants with PWMI with bilateral SCP were found along the CST, mainly located at the internal capsule level, while these changes in FA were not found in infants with PWMI without CP. The PTR, GCC, and SCC showed significantly decreased FA values in both groups.

The receiver operating characteristic curves for the diagnostic performances of FA in different regions for differentiation of infants with PWMI with bilateral SCP from those without CP are depicted in On-line Fig 2. The area under the curve was larger in the CST than that in the PTR, GCC, or SCC (Table 2). Among the 3 parts of the CST, an FA threshold of 0.53 at the internal capsule level had the highest area under the curve value, which demonstrated high specificity and sensitivity.

Correlation between Motor Function Scores and FA

The correlations between motor function scores and FA values along white matter tracts are presented in On-line Table 2. There were more significant correlations along the CST than along the PTR, GCC, and SCC ($P < .05$); and GMFCS levels were negatively correlated with FA values within the CST at the level of the internal capsule (left $r = -0.80$; right $r = -0.79$). The Psychomotor Development Index scores showed a positive correlation with FA values within the CST at the level of internal capsule (left $r = 0.69$; right $r = 0.53$).

DISCUSSION

This study found that the spatial distribution of white matter alterations was different between the 2 PWMI groups. Structural integrity was more severely damaged in more widespread white matter areas in infants with PWMI with bilateral SCP than in those without CP. Furthermore, injured CST was found only in infants with PWMI with bilateral SCP. The FA threshold of 0.53 in the CST at the internal capsule level was useful for the differentiation of infants with PWMI with SCP from those without CP, and it had significant correlations with motor functions.

DTI with TBSS analysis is a powerful approach that effectively reveals alterations in the main white matter tracts of the whole brain. Several studies have reported the presence of FA reduction within widespread white matter areas, especially in the central and posterior white matter, in infants with PWMI with SCP.[10,12,22] However, knowledge of the spatial distribution characteristics of white matter lesions in infants with PWMI with and without CP is limited. In this study, the infants with PWMI with SCP had lower FA and more widespread distribution than those without CP. Additionally, the age distribution of patients here was different from that in previous studies.[10,12,15,22] We focused on 6- to 18-month-old infants with PWMI, which is the period before the definite diagnosis. Identification of the differences between infants with PWMI with SCP and those without CP during this time interval is beneficial for determining candidates at risk of SCP.

In this study, the white matter tracts potentially involved in SCP were further analyzed using the AFQ. Previously, Hoon et al[13] had suggested that PTR injury altered the sensorimotor connections to the motor cortex and attenuated descending CST. However, our study showed that the PTR, SCC, and GCC were the

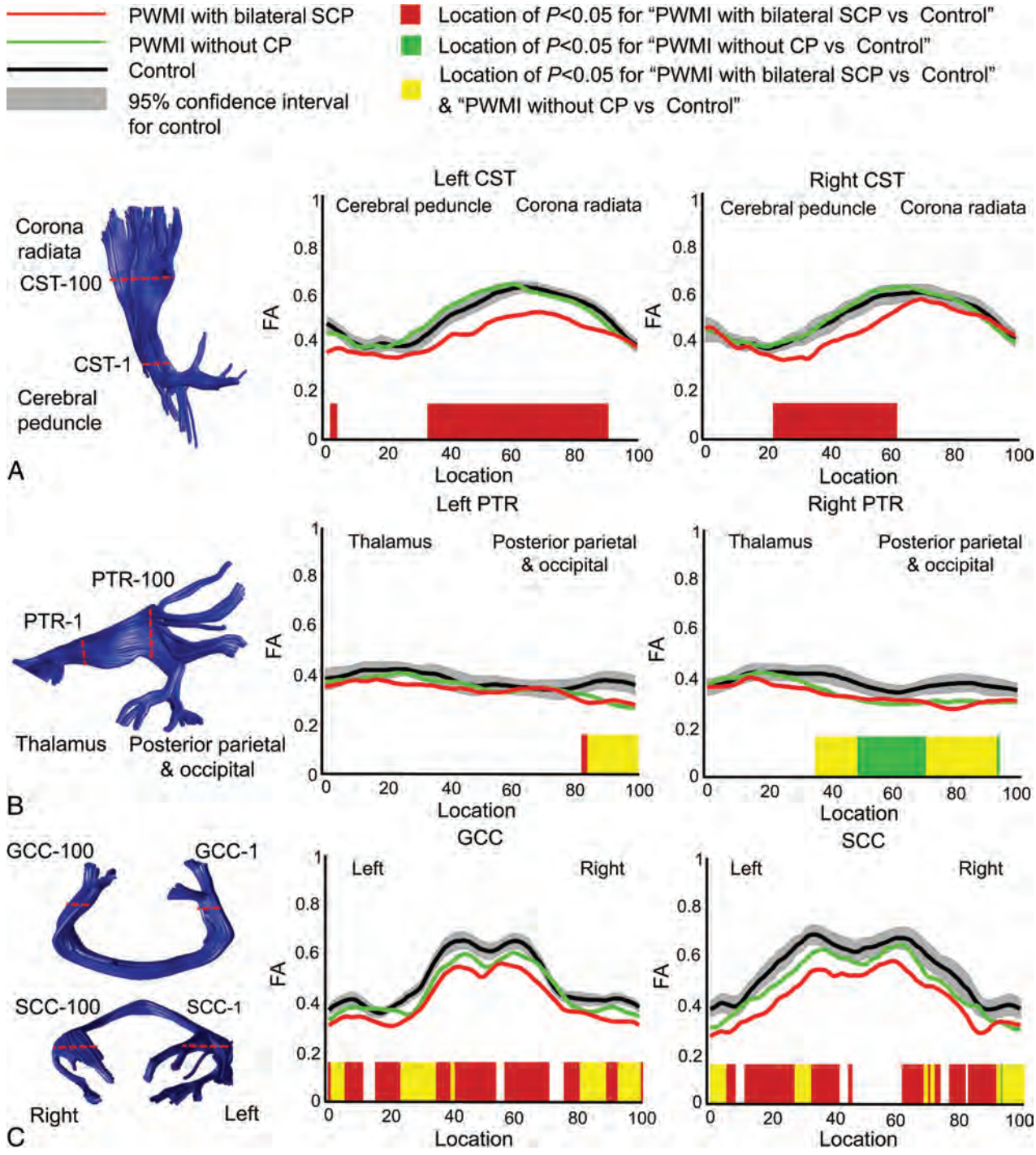


FIG 3. Intergroup comparisons of fractional anisotropy among infants with PWMI with bilateral SCP, infants with PWMI without CP, and controls along tracts of the bilateral corticospinal tract (*A*), bilateral posterior thalamic radiation (*B*), and genu of the corpus callosum and splenium of the corpus callosum (*C*).

injured regions in both PWMI subgroups, whereas decreased FA in CST was observed only in infants with PWMI with SCP. These results indicate that PTR injury may not be directly responsible for SCP. The CST, as the major projectional motor tract situated in close proximity to the periventricular white matter region, is vulnerable in patients with PWMI.[23] CST lesions can interrupt the corticomotor circuit in executing movement and have also been shown to be involved in spasticity.[24] These findings suggest that CST lesions are a prerequisite for bilateral SCP in infants with PWMI and provide clues for determining whether patients with PWMI will develop SCP.

Previous research has shown that an FA threshold of 0.5 within the CST was effective in differentiating CP and non-CP groups.[25] Similarly, this study demonstrated that the FA thresholds in CST for the diagnosis of SCP were 0.48 (left) and 0.49 (right). Additionally, the current study revealed that the FA thresholds in both the left and right CST at the internal capsule level were 0.53, with high sensitivity and specificity. These were higher than results in

Table 2: Diagnostic performance of fractional anisotropy in different regions for differentiation of infants with PWMI with bilateral spastic cerebral palsy and those without cerebral palsy

	PWMI with SCP (Mean and standard deviation)	PWMI without CP (Mean and standard deviation)	Area under the Curve (95% CI)	Threshold	Sensitivity	Specificity
CST-L	0.42 ± 0.06	0.50 ± 0.04	0.90 (0.75–0.98)	0.48	94%	72%
CST-R	0.43 ± 0.06	0.50 ± 0.04	0.85 (0.68–0.95)	0.49	93%	67%
PTR-L	0.34 ± 0.06	0.35 ± 0.05	0.55 (0.37–0.73)	0.27	27%	100%
PTR-R	0.33 ± 0.07	0.34 ± 0.05	0.54 (0.35–0.72)	0.36	67%	50%
GCC	0.41 ± 0.05	0.45 ± 0.05	0.72 (0.59–0.82)	0.46	87%	50%
SCC	0.43 ± 0.11	0.49 ± 0.09	0.62 (0.49–0.74)	0.40	40%	88%
CST-CP-L	0.36 ± 0.08	0.42 ± 0.04	0.74 (0.57–0.88)	0.40	69%	83%
CST-CP-R	0.36 ± 0.06	0.42 ± 0.06	0.77 (0.60–0.90)	0.40	73%	72%
CST-IC-L	0.48 ± 0.06	0.59 ± 0.05	0.95 (0.82–0.99)	0.53	94%	89%
CST-IC-R	0.48 ± 0.07	0.58 ± 0.05	0.88 (0.72–0.97)	0.53	93%	78%
CST-CR-L	0.43 ± 0.09	0.46 ± 0.09	0.60 (0.42–0.77)	0.49	81%	50%
CST-CR-R	0.47 ± 0.10	0.48 ± 0.11	0.52 (0.34–0.70)	0.59	100%	17%

Note:—L indicates left; R, right.

previous MR imaging studies for detecting the risk of CP.[8,26] The results obtained in this study suggest that CST at the internal capsule level is more suitable than the whole tract for the early diagnosis of SCP. Furthermore, the multiple regression results revealed that FA values of infants with PWMI were not correlated to age (On-line Table 3). This finding suggests that the FA threshold is applicable to the individual infants with PWMI between 6 and 18 months of age.

Apart from the determination of the risk of CP, the evaluation of the degree of motor impairment in CP can further facilitate the targeting of interventional strategies.[27,28] Consistent with previous findings,[23,24] FA in the CST was significantly correlated with motor function scores. A strong correlation was found along the CST at the internal capsule level, which may be associated with the vulnerability of white matter in the internal capsule, where the descending motor axons are densely concentrated.[29,30] Moreover, there were more significant correlations along the CST than along the PTR, GCC, and SCC. It has been suggested that PTR lesions are responsible for the weakness in motor function in patients with CP because of the role of PTR in visuospatial performance.[24] The corpus callosum may play an important role in dexterity and bimanual motor coordination.[31] Overall, the results suggest that FA of the CST at the internal capsule level is suitable for assessing the motor function in infants with PWMI with SCP.

Nevertheless, this study has several limitations. First, it excluded severe PWMI because these cases accounted for a minority (3 subjects) of all the infants with PWMI. Moreover, it was relatively easy to determine the risk of CP in these infants. Second, unilateral SCP and nonspastic CP were not included due to the small sample sizes (2 with unilateral SCP and 1 with athetoid CP) and divergent MR imaging signs, such as asymmetric brain lesions in unilateral SCP and concurrent basal ganglia lesions in athetoid CP. Third, the FA threshold in the internal capsule of the CST was objectively extracted from an Automated Fiber Quantification. This threshold may be different from those obtained by other analysis methods, such as the ROI analysis. Fourth, this work tried to investigate the characteristics of DTI metric changes associated with SCP before the definite diagnosis. However, DTI was performed just several months ahead of the diagnosis of SCP. Whether SCP can be predicted by DTI in earlier periods needs further research. Finally, this was a retrospective study. Prospective research is required to test the accuracy of the early diagnosis of SCP.

CONCLUSIONS

CST lesions play a crucial role in motor impairment associated with SCP in infants with PWMI. FA in the CST at the internal capsule level could aid in the early diagnosis of SCP with high diagnostic accuracy.

ACKNOWLEDGMENTS

The authors are grateful to Drs Liming Liu and Jie Yue in the Department of Child Health Care for the diagnosis of CP.

REFERENCES

1. Bax M, Tydeman C, Flodmark O. **Clinical and MRI correlates of cerebral palsy: the European Cerebral Palsy Study.** *JAMA* 2006;296:1602–08 CrossRef Medline
2. Back SA, Rivkees SA. **Emerging concepts in periventricular white matter injury.** *Semin Perinatol* 2004;28:405–14 CrossRef Medline
3. Jauhari P, Singhi P, Sankhyan N, et al. **A comparison of spastic diplegia in term and preterm-born children.** *J Child Neurol* 2018;33:333–39 CrossRef Medline
4. Imamura T, Ariga H, Kaneko M, et al. **Neurodevelopmental outcomes of children with periventricular leukomalacia.** *Pediatr Neonatol* 2013;54:367–72 CrossRef Medline
5. Woodward LJ, Anderson PJ, Austin NC, et al. **Neonatal MRI to predict neurodevelopmental outcomes in preterm infants.** *N Engl J Med* 2006;355:685–94 CrossRef Medline
6. Hadders-Algra M. **Early diagnosis and early intervention in cerebral palsy.** *Front Neurol* 2014;5:185 CrossRef Medline
7. Morgan C, Crowle C, Goyen TA, et al. **Sensitivity and specificity of General Movements Assessment for diagnostic accuracy of detecting cerebral palsy early in an Australian context.** *J Paediatr Child Health* 2016;52:54–59 CrossRef Medline
8. Novak I, Morgan C, Adde L, et al. **Early, accurate diagnosis and early intervention in cerebral palsy: advances in diagnosis and treatment.** *JAMA Pediatr* 2017;171:897–907 CrossRef Medline
9. Son SM, Park SH, Moon HK, et al. **Diffusion tensor tractography can predict hemiparesis in infants with high risk factors.** *Neurosci Lett* 2009;451:94–97 CrossRef Medline
10. Lee JD, Park HJ, Park ES, et al. **Motor pathway injury in patients with periventricular leucomalacia and spastic diplegia.** *Brain* 2011;134:1199–210 CrossRef Medline
11. Skranes J, Vangberg T, Kulseng S, et al. **Clinical findings and white matter abnormalities seen on diffusion tensor imaging in adoles-**

cents with very low birth weight. *Brain* 2007;130:654–66 CrossRef Medline
12. Ceschin R, Lee VK, Schmithorst V, et al. **Regional vulnerability of longitudinal cortical association connectivity: associated with structural network topology alterations in preterm children with cerebral palsy.** *Neuroimage Clin* 2015;9:322–37 CrossRef Medline
13. Hoon AH Jr, Stashinko EE, Nagae LM, et al. **Sensory and motor deficits in children with cerebral palsy born preterm correlate with diffusion tensor imaging abnormalities in thalamocortical pathways.** *Dev Med Child Neurol* 2009;51:697–704 CrossRef Medline
14. Trivedi R, Agarwal S, Shah V, et al. **Correlation of quantitative sensorimotor tractography with clinical grade of cerebral palsy.** *Neuroradiology* 2010;52:759–65 CrossRef Medline
15. Nagae LM, Hoon AH Jr, Stashinko EE, et al. **Diffusion tensor imaging in children with periventricular leukomalacia: variability of injuries to white matter tracts.** *AJNR Am J Neuroradiol* 2007;28:1213–22 CrossRef Medline
16. Colver A, Fairhurst C, Pharoah PO. **Cerebral palsy.** *Lancet* 2014;383:1240–49 CrossRef Medline
17. Rosenbaum P, Paneth N, Leviton A, et al. **A report: the definition and classification of cerebral palsy April 2006.** *Dev Med Child Neurol* 2007;109:8–14 Medline
18. Coté CJ, Wilson S; American Academy of Pediatrics; American Academy of Pediatric Dentistry. **Guidelines for monitoring and management of pediatric patients during and after sedation for diagnostic and therapeutic procedures: update 2016.** *Pediatrics* 2016;138. pii: e20161212 CrossRef Medline
19. Smith SM, Jenkinson M, Woolrich MW, et al. **Advances in functional and structural MR image analysis and implementation as FSL.** *Neuroimage* 2004;23:S208–19 CrossRef Medline
20. Yeatman JD, Dougherty RF, Myall NJ, et al. **Tract profiles of white matter properties: automating fiber-tract quantification.** *PLoS One* 2012;7:e49790 CrossRef Medline
21. Oishi K, Faria AV, Yoshida S, et al. **Quantitative evaluation of brain development using anatomical MRI and diffusion tensor imaging.** *Int J Dev Neurosci* 2013;31:512–24 CrossRef Medline
22. Arrigoni F, Peruzzo D, Gagliardi C, et al. **Whole-brain DTI assessment of white matter damage in children with bilateral cerebral palsy: evidence of involvement beyond the primary target of the anoxic insult.** *AJNR Am J Neuroradiol* 2016;37:1347–53 CrossRef Medline
23. Rešić B, Tomasović M, Kuzmanić-Šamija R, et al. **Neurodevelopmental outcome in children with periventricular leukomalacia.** *Coll Antropol* 008;32(Suppl 1):143–47 Medline
24. Scheck SM, Boyd RN, Rose SE. **New insights into the pathology of white matter tracts in cerebral palsy from diffusion magnetic resonance imaging: a systematic review.** *Dev Med Child Neurol* 2012;54:684–96 CrossRef Medline
25. Murakami A, Morimoto M, Yamada K, et al. **Fiber-tracking techniques can predict the degree of neurologic impairment for periventricular leukomalacia.** *Pediatrics* 2008;122:500–06 CrossRef Medline
26. Thayyil S, Chandrasekaran M, Taylor A, et al. **Cerebral magnetic resonance biomarkers in neonatal encephalopathy: a meta-analysis.** *Pediatrics* 2010;125:e382–95 CrossRef Medline
27. Kendall GS, Melbourne A, Johnson S, et al. **White matter NAA/Cho and Cho/Cr ratios at MR spectroscopy are predictive of motor outcome in preterm infants.** *Radiology* 2014;271:230–38 CrossRef Medline
28. Melhem ER, Hoon AH Jr, Ferrucci JT Jr, et al. **Periventricular leukomalacia: relationship between lateral ventricular volume on brain MR images and severity of cognitive and motor impairment.** *Radiology* 2000;214:199–204 CrossRef Medline
29. Jaspers E, Byblow WD, Feys H, et al. **The corticospinal tract: a biomarker to categorize upper limb functional potential in unilateral cerebral palsy.** *Front Pediatr* 2016;3:112 CrossRef Medline
30. Rose J, Mirmiran M, Butler EE, et al. **Neonatal microstructural development of the internal capsule on diffusion tensor imaging correlates with severity of gait and motor deficits.** *Dev Med Child Neurol* 2007;49:745–50 CrossRef Medline
31. Weinstein M, Green D, Geva R, et al. **Interhemispheric and intrahemispheric connectivity and manual skills in children with unilateral cerebral palsy.** *Brain Struct Funct* 2014;219:1025–40 CrossRef Medline

Cesarean Delivery Impacts Infant Brain Development

S.C. Deoni, S.H. Adams, X. Li, T.M. Badger, R.T. Pivik, C.M. Glasier, R.H. Ramakrishnaiah, A.C. Rowell, and X. Ou

ABSTRACT

BACKGROUND AND PURPOSE: The cesarean delivery rate has increased globally in the past few decades. Neurodevelopmental outcomes associated with cesarean delivery are still unclear. This study investigated whether cesarean delivery has any effect on the brain development of offspring.

MATERIALS AND METHODS: A total of 306 healthy children were studied retrospectively. We included 3 cohorts: 2-week-old neonates (cohort 1, $n = 32/11$ for vaginal delivery/cesarean delivery) and 8-year-old children (cohort 2, $n = 37/23$ for vaginal delivery/cesarean delivery) studied at Arkansas Children's Hospital, and a longitudinal cohort of 3-month to 5-year-old children (cohort 3, $n = 164/39$ for vaginal delivery/cesarean delivery) studied independently at Brown University. Diffusion tensor imaging, myelin water fraction imaging, voxel-based morphometry, and/or resting-state fMRI data were analyzed to evaluate white matter integrity, myelination, gray matter volume, and/or functional connectivity, respectively.

RESULTS: While not all MR imaging techniques were shared across the institutions/cohorts, post hoc analyses showed similar results of potential effects of cesarean delivery. The cesarean delivery group in cohort 1 showed significantly lower white matter development in widespread brain regions and significantly lower functional connectivity in the brain default mode network, controlled for a number of potential confounders. No group differences were found in cohort 2 in white matter integrity or gray matter volume. Cohort 3 had significantly different trajectories of white matter myelination between groups, with those born by cesarean delivery having reduced myelin in infancy but normalizing with age.

CONCLUSIONS: Cesarean delivery may influence infant brain development. The impact may be transient because similar effects were not observed in older children. Further prospective and longitudinal studies may be needed to confirm these novel findings.

ABBREVIATIONS: BMI = body mass index; DMN = default mode network; FA = fractional anisotropy; ICA = independent component analysis; IQ = intelligence quotient; MWF = myelin water fraction; RS-fMRI = resting-state fMRI; TFCE = threshold-free cluster enhancement; TBSS = Tract-Based Spatial Statistics; VBM = voxel-based morphometry

Cesarean delivery has become increasingly prevalent throughout the world.[1] In the United States, about one-third of all births are by cesarean delivery,[2] a 60% increase from the rate in 1996.[3] Elective cesarean delivery without a medical indication is not uncommon[4,5] and contributed to a considerable percentage (8%) of the increase of cesarean deliveries in recent years.[6] In addition, the large range of cesarean delivery rates across hospitals[7] suggests that practice variation, a modifiable factor, also played an important role in the increased prevalence of cesarean delivery. The World Health Organization recently reassessed available evidence and concluded that cesarean delivery should ideally be undertaken only when medically necessary, acknowledging that the relationship between the cesarean delivery rate and pediatric outcome is unclear.[1] An important aspect to con-

Received July 24, 2018; accepted after revision October 6.

From the Arkansas Children's Nutrition Center (S.H.A., T.M.B., R.T.P., X.O.), Little Rock, Arkansas; Departments of Radiology (X.L., C.M.G., R.H.R., A.C.R., X.O.) and Pediatrics (S.H.A., T.M.B., R.T.P., C.M.G., X.O.), University of Arkansas for Medical Sciences, Little Rock, Arkansas; Arkansas Children's Research Institute (C.M.G., R.H.R., A.C.R., X.O.), Little Rock, Arkansas; and School of Engineering (S.C.D.), Brown University, Providence, Rhode Island.

The studies were supported, in part, by the US Department of Agriculture–Agricultural Research Service Project (6026-51000-010-05S) at the Arkansas Children's Nutrition Center, the Marion B. Lyon Award at the Arkansas Children's Research Institute, and grants from the National Institute of Mental Health (R01MH087510) and the Bill and Melinda Gates Foundation. Dr Xiawei Ou is also supported by National Institutes of Health Centers of Biomedical Research Excellence grants P20GM121293 (Center for Translational Pediatric Research at Arkansas Children's Research Institute) and P30GM110702 (Center for Translational Neuroscience at University of Arkansas for Medical Sciences).

Please address correspondence to Xiawei Ou, PhD, One Children's Way, Slot 105, Little Rock, AR 72202; e-mail: ouxiawei@uams.edu

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

Indicates article with supplemental on-line appendix.

http://dx.doi.org/10.3174/ajnr.A5887

Table 1: Study population demographics for cohort 1—two-week-old neonates ($n = 43$)

	Vaginal Delivery (Mean ± SD) ($n = 32$)	Cesarean Delivery (Mean ± SD) ($n = 11$)	*P* Value
DTI TBSS study ($n = 43$)	32	11	
RS-fMRI study ($n = 37$)	27	10	
Mother's age at delivery (yr)	29.0 ± 3.9	30.6 ± 4.4	.26
	29.2 ± 4.1	30.6 ± 4.6	.35
Mother's body fat % at early/prepregnancy	33.4 ± 8.6	36.9 ± 8.5	.24
	34.5 ± 8.8	38.4 ± 7.3	.23
Mother's BMI at 12 wk of pregnancy	25.1 ± 5.2	28.8 ± 5.1	.04
	25.5 ± 5.5	29.2 ± 5.1	.08
Mother's gestational weight gain (kg)	11.79 ± 3.07	9.02 ± 3.37	.02[c]
	11.51 ± 2.75	8.87 ± 3.51	.03[c]
Mother's IQ[a]	107.7 ± 9.6	104.2 ± 6.9	.18
	108.6 ± 9.2	104.8 ± 6.9	.22
Gestational age at MRI[b] (days)	290 ± 6	287 ± 6	.20
	291 ± 6	288 ± 5	.22
Sex (M/F)	20/12	3/8	.04[c]
	19/8	3/7	.03[c]
Birth weight (kg)	3.51 ± 0.46	3.56 ± 0.54	.79
	3.49 ± 0.36	3.63 ± 0.52	.35
Birth length (cm)	51.1 ± 2.6	49.3 ± 2.6	.05[c]
	51.2 ± 2.1	49.3 ± 2.7	.02[c]
Head circumference (cm)	36.2 ± 1.1	35.9 ± 0.5	.22
	36.3 ± 0.9	35.9 ± 0.6	.31
Diet for the first 2 wk (breastmilk/formula)	29/3	10/1	.98
	24/3	9/1	.92

For each cell in the table, the first row is for the DTI TBSS study, the second row is for the RS-fMRI study.
[a] Not available for 1 subject with vaginal delivery.
[b] Defined as gestational age at birth plus postnatal age at MRI.
[c] Consistently different between groups for both TBSS and RF-fMRI and therefore included as a covariate in the analysis.

sider is long-term neurodevelopment of children born by cesarean delivery. While some studies have not shown significant differences in intelligence quotient (IQ) between children born by cesarean delivery compared with vaginal delivery,[8,9] others have suggested an increased risk of autism spectrum disorder and/or attention deficit/hyperactivity disorder associated with cesarean delivery,[10-13] though it is unlikely that cesarean delivery is the single or main cause of these outcomes.[14,15]

The mixed and inconclusive associations between delivery mode and neurodevelopment may be partly attributed to numerous confounding factors that come into play during childhood (eg, diet and life-style, family environment and enrichment, and so forth), which may be avoided by evaluating the effects of cesarean delivery during early infancy. While neurobehavioral assessments in young infants are often not sufficiently sensitive for detecting small differences or for predicting long-term neurodevelopmental outcomes in the healthy population, quantitative neuroimaging by MR imaging can provide a comprehensive evaluation of the developing brain and reveal subtle changes.[16-18] Examples of using several advanced MR imaging methods to reveal minor changes in infant brain development are given in On-line Appendix.

To test whether cesarean delivery has significant effects on infant brain development and whether these effects persist into and throughout childhood, we analyzed advanced brain MR imaging data in 3 cohorts of healthy children—cohort 1: two-week-old neonates with well-documented pregnancy and infant perinatal profiles; cohort 2: eight-year-old children who had undergone a suite of neuropsychological tests; and cohort 3: three-month to 5-year-old children with MR imaging data acquired and independently analyzed at a separate institution.

MATERIALS AND METHODS

All study procedures were approved by the local institutional review boards at the University of Arkansas for Medical Sciences (cohorts 1 and 2) and Brown University (cohort 3). Written informed consent was obtained from all parents or legal guardians of the study subjects.

Study Population

Cohort 1 consisted of 2-week-old, full-term, healthy neonates ($n = 43$, 32/11 for vaginal and cesarean delivery, respectively) from uncomplicated pregnancies. They were scanned at the Radiology Department of Arkansas Children's Hospital at ~2 weeks of age. The scan included regular imaging to screen for apparent brain abnormalities and diffusion tensor imaging and resting-state functional MR imaging (RS-fMRI) to quantitatively evaluate brain structural and functional connectivity. In total, 43 infants (32/11 for vaginal and cesarean delivery, respectively) had valid DTI data and 37 infants (27/10 for vaginal and cesarean delivery, respectively) had valid RS-fMRI data and were included in the analysis. Their demographic information is presented in Table 1.

Cohort 2 consisted of healthy 7.5- to 8.5-year-old children ($n = 60$, 37/23 for vaginal and cesarean delivery, respectively). MR imaging examinations at Arkansas Children's Hospital included DTI evaluation of brain white matter and T1-weighted high-resolution structural images for voxel-based morphometry (VBM) evaluation of brain gray matter volume. Demographic information for subjects with valid DTI ($n = 35/21$ for vaginal and cesarean delivery, respectively) and VBM data ($n = 36/21$ for vaginal and cesarean delivery, respectively) is presented in Table 2.

Cohort 3 consisted of healthy, typically developing, 3- to 60-month-old children ($n = 203$, 164/39 for vaginal and cesarean delivery, respectively). All subjects had at least 1 brain MR imaging at the Advanced Baby Imaging Lab at Brown University with valid myelin water fraction (MWF) imaging data. In total, 482 (377/105 for vaginal and cesarean delivery, respectively) datasets were acquired. Demographic information for this cohort is listed in Table 3.

Details such as inclusion/exclusion or selection criteria for these 3 cohorts and methods for demographic data measurements are presented in On-line Appendix.

MR Imaging Data Acquisition and Analysis

The data acquisition and analysis were based on established protocols at Arkansas Children's Hospital and Brown University and were similar to those in previous publications.[16-23] Details are provided in On-line Appendix.

Table 2: Study population demographics for cohort 2—eight-year-olds ($n = 60$)

	Vaginal Delivery ($n = 37$)	Cesarean Delivery ($n = 23$)	*P* Value
DTI TBSS study ($n = 56$)	35	21	
VBM study ($n = 57$)	36	21	
Birth weight (mean ± SD) (kg)	3.37 ± 0.40	3.44 ± 0.49	.79
	3.34 ± 0.43	3.45 ± 0.48	.55
Gestational age at birth (mean ± SD) (days)	276 ± 9	275 ± 9	.43
	275 ± 10	274 ± 10	.75
Age at MRI (mean ± SD) (yr)	7.93 ± 0.26	7.87 ± 0.24	.42
	7.94 ± 0.26	7.92 ± 0.27	.94
Sex (M/F)	13/22	11/10	.26
	14/22	9/12	.77
Infant diet (breastmilk/formula)	14/21	4/17	.10
	15/21	4/17	.08
BMI (mean ± SD)	16.8 ± 2.9	17.1 ± 2.3	.49
	16.7 ± 2.9	17.4 ± 2.8	.27
SES[a] mother's education: high school/college/graduate degree	4/15/8	2/5/3	.99
SES[a] father's education: high school/college/graduate degree	7/17/3	3/6/1	.99
SES[a] mother's annual income: NA/<$20k/$20k–$50k/>$50k	12/5/9/1	1/2/3/4	.03
SES[a] father's annual income: NA/<$20k/$20k–$50k/>$50k	1/0/11/15	0/0/3/7	.79
IQ (mean ± SD)	110 ± 10	112 ± 17	.53
	112 ± 11	112 ± 16	.72
CELF-4 language scores (mean ± SD)[b]	103 ± 14	104 ± 16	.31
	103 ± 12	105 ± 14	.24
CMS general memory index (mean ± SD)[b]	113 ± 12	109 ± 13	.14
	115 ± 11	111 ± 13	.24

Note:—SES indicates socioeconomic status; CELF-4, Clinical Evaluation of Language Fundamentals; CMS, Childhood Memory Scale; NA, not applicable. For each cell in the table, the first row is for the DTI TBSS study, the second row is for the VBM study.

[a] Only partial socioeconomic status data were available for all subjects ($n = 27$ for vaginal delivery, $n = 10$ for cesarean delivery).

[b] Not available for 1 subject with cesarean delivery.

Table 3: Study population demographics for cohort 3—three-month- to 5-year-olds ($n = 203$)

	Vaginal Delivery (Mean ± SD)	Cesarean Delivery (Mean ± SD)	*P* Value
MWF study ($n = 203$)	164	39	
Mother's post-secondary education (yr)	5.8 ± 1.0	6.0 ± 0.9	.22
Mother's age at pregnancy (yr)	29.0 ± 5.6	33.8 ± 5.1	<.001
Sex (M/F)	93/71	23/16	.86
Gestational age at birth (days)	277 ± 8	273 ± 4	.001
Birth weight (kg)	3.54 ± 0.45	3.37 ± 0.48	.06
Birth length (cm)	50.8 ± 4.1	50.8 ± 4.1	.63
No. of scans	2.2 ± 1.2	2.4 ± 1.3	.32
Mean interscan period (days)	296 ± 144	333 ± 147	.12

Statistics

For the comparison of demographic/anthropometric/neuropsychological data between delivery mode groups in cohorts 1 and 2, Wilcoxon rank sum tests were used for numeric parameters and Fisher exact tests (or χ^2 tests) were used for categoric variables. For the voxelwise comparison of DTI parameters in Tract-Based Spatial Statistics (TBSS; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/TBSS) analysis and RS-fMRI parameters in dual-regression analysis, randomization with 5000 permutations was used with the threshold-free cluster enhancement (TFCE) option (testing of larger amounts of permutations did not change the results). To correct for multiple comparisons for the voxelwise analysis, we compared the observed TFCE image with the empiric null distribution computed across permutations of the maximum voxel-specific TFCE scores.[24,25] $P < .05$ corrected for multiple comparisons (voxelwise) was regarded as significant.

Because multiple independent component analysis (ICA) components were defined as meaningful functional networks and were fed into the dual-regression program for group comparison in the RS-fMRI analyses, an additional threshold of cluster sizes of >5 imaging voxels was used. For post hoc ROI analyses of DTI and RS-fMRI parameters, independent *t* tests were used after confirming normal distribution of data and testing for equality of variance, and general linear model univariate analyses were used to compare differences, with covariates controlled. For the VBM analyses of regional gray matter volume, 2-sample *t* tests with unequal variance were performed with multiple-comparison correction to control for family-wise error. $P < .05$ family-wise error–corrected was regarded as significant. For cohort 3, the Gompertz function parameters for the MWF curve of the 6 regions were compared between delivery groups using nonparametric tests, with significance defined as $P < .001$.

Potential confounders were controlled as covariates in statistical analyses. Specifically, for cohort 1, the TBSS and RS-fMRI dual-regression analyses were performed, respectively, when there were no potential confounders added as covariates; when adding demographic parameters that were consistently different between groups ($P \leq .05$ for both the TBSS and the RS-fMRI subsets) as covariates (gestational weight gain, infant sex, and birth length); and when adding an additional 2 parameters that were known potential confounders identified by our previous studies (gestational age at MR imaging and maternal body mass index (BMI) at early pregnancy[17,18,26]) as covariates. For the post hoc ROI comparisons, significances with and without controlling for all of these 5 covariates were both evaluated. For cohort 2, the TBSS and VBM analyses were performed without adding covari-

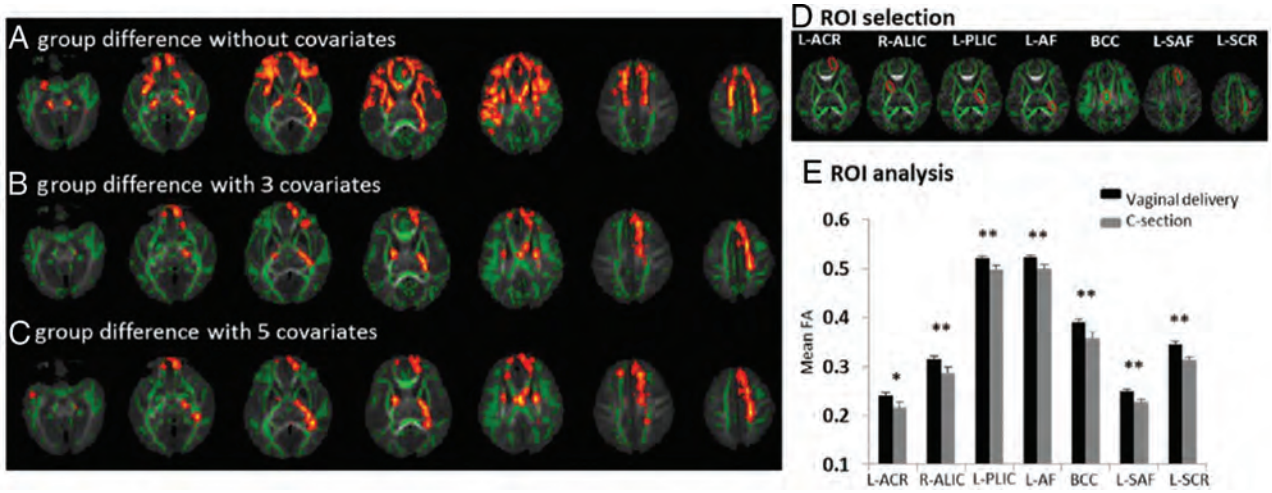


FIG 1. DTI TBSS results for the comparison of FA values between 2-week-old neonates born by cesarean delivery or vaginal delivery in cohort 1. Green represents major white matter tracts in the brain; orange shows voxels that have different FA values between groups. *A*, Group differences ($P < .05$, corrected for the voxelwise multiple comparisons) when no covariates were added. *B*, Group differences ($P < .05$, corrected for the voxelwise multiple comparisons) when 3 demographic parameters that were different between groups (gestational weight gain, infant sex, and birth length) were added as covariates. *C*, Trend of group differences ($P < .15$, corrected for the voxelwise multiple comparisons) when an additional 2 potential confounders (gestational age at MR imaging and maternal BMI at early pregnancy) were included as covariates. *D*, Illustration of ROIs (*red outlined*) selected for further post hoc analysis. *E*, Group comparison of mean FA values in these white matter ROIs. L indicates left; R, right; ACR, anterior corona radiata; ALIC, anterior limb of internal capsule; PLIC, posterior limb of internal capsule; AF, arcuate fasciculus; BCC, body of the corpus callosum; SAF, short association fibers; SCR, superior corona radiata; *asterisk*, $P < .05$ without controlling for covariates; *double asterisks*, $P < .05$ with and without controlling for covariates.

ates and then were repeated with age, sex,[27] and infant diet[22] added as covariates. Socioeconomic variables were not included as covariates because of incomplete data (group comparison on the available data did not show group differences other than mother's income).

RESULTS

The demographic/anthropometric/neuropsychological parameters are listed in Tables 1–3 for the 3 cohorts. For cohort 1, the 2 groups of neonates did not differ with respect to mother's age, maternal IQ, and body fat percentage at early pregnancy, infant gestational age at MR imaging, birth weight, head circumference, and diet at 2 weeks of age. There was a trend toward differences in maternal BMI measured at 12 weeks of pregnancy (significant for the subjects included in the TBSS analysis but not significant for the subjects included in the RS-fMRI analysis). There were group differences in the mother's gestational weight gain, infant sex, and birth length. For cohort 2, the 2 groups of 8-year-old children did not differ in any demographic or family socioeconomic status measures except for mother's income. The neuropsychological test scores including IQ, language skills, and memory index were also not significantly different between groups. For cohort 3, the 2 groups of children did not differ in birth weight, birth length, maternal education, number of scans per child, and mean interscan period. However, the cesarean delivery group had a higher maternal age and slightly lower gestational age at birth.

Imaging Findings for Cohort 1. DTI TBSS revealed widespread white matter regions in the frontal, parietal, and temporal lobes that had higher fractional anisotropy (FA) values ($P < .05$, corrected for voxelwise multiple comparisons, indicating better white matter microstructural integrity and connectivity) in the vaginally delivered infants compared with those born by cesarean delivery (Fig 1*A*) when no covariates were added to the voxelwise analysis. Many of these differences were still significant ($P < .05$, corrected for voxelwise multiple comparisons) when the 3 demographic parameters that differed between groups (gestational weight gain, infant sex, and birth length) were added as covariates to the analysis (Fig 1*B*). Gestational age at MR imaging and maternal BMI at early pregnancy are potentially additional confounders influencing infant brain development.[17,18,26] When they were also added as covariates into the analysis, the trend toward differences remained (at a level of $P < .15$ corrected for voxelwise multiple comparisons) for many clusters in this voxelwise analysis (Fig 1*C*) involving white matter tracts such as the anterior corona radiata, anterior/posterior internal capsule, arcuate fasciculus, short association fibers, superior corona radiata, and body of the corpus callosum. Further post hoc ROI analyses for these anatomic regions/white matter tracts (see Fig 1*D* for an illustration of ROI selections) did show statistically significant differences in mean FA values between groups (Fig 1*E*), with a mean FA in infants born by vaginal delivery 4%–10% higher than that in infants born by cesarean delivery for different ROIs. For 6 of 7 ROIs, the differences in mean FA values were significant both with and without controlling for all covariates.

RS-fMRI analyses revealed functional connectivity differences between the 2 groups in the default mode network (DMN). Fig 2*A* shows DMNs in neonates (obtained from the respective group ICA analyses), typical for the developing neonatal brain (that posterior regions such as posterior cingulate cortex and precuneus were well-recruited but anterior brain involvement was weak compared with that in older children).[28] Dual-regression analysis for this ICA network showed a cluster in the precuneus region that had higher functional connectivity ($P < .05$, corrected for voxelwise multiple comparisons) in the vaginal delivery compared with

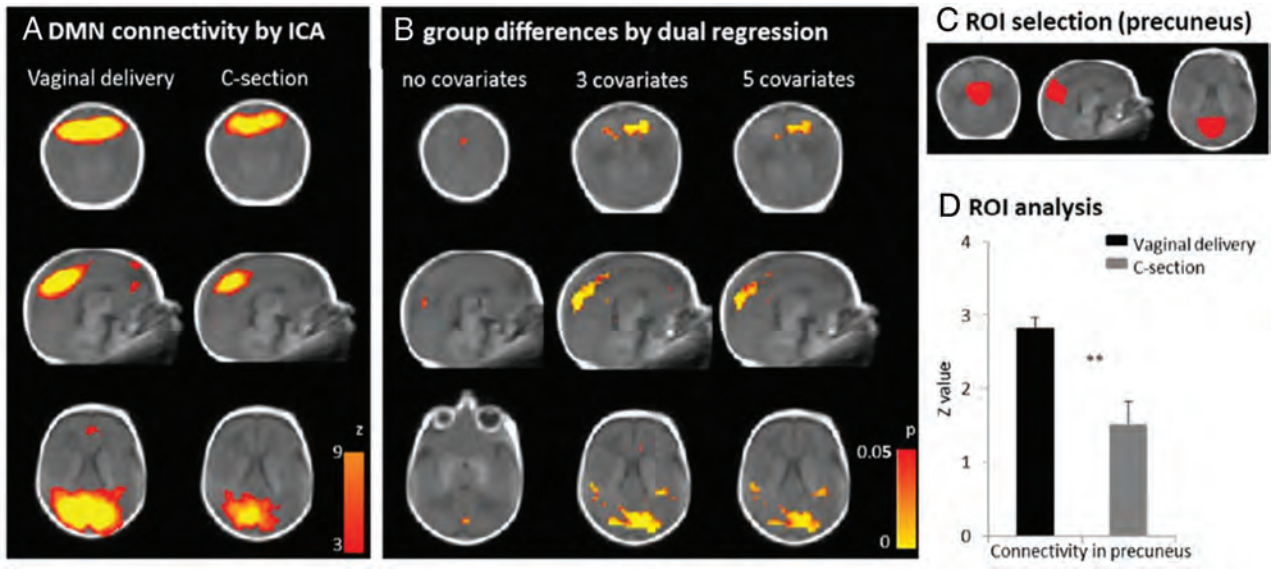


FIG 2. Resting-sate fMRI analyses of functional connectivity in the DMN in 2-week-old neonates in cohort 1. *A*, The DMN for vaginal and cesarean delivery groups obtained from respective independent component analysis (ICA). *B*, Regions in which the vaginal delivery group had higher functional connectivity (*z* score) in the DMN than the cesarean delivery group ($P < .05$, corrected for the voxelwise multiple comparisons) when excluding or including potential confounders as covariates. *C*, Illustration of ROI (precuneus selected) for further post hoc analysis. *D*, Comparison of the mean *z* score values in the precuneus between the 2 groups. *Double asterisks* indicate $P < .05$ with and without controlling for covariates.

the cesarean delivery group (Fig 2*B*). After controlling for potential confounders (3 demographic parameters that differed between groups and 2 more parameters known to influence neonatal brain development), the differences were more prominent with a larger cluster in the precuneus and additional bilateral regions in the parietal lobes involved (Fig 2*B*). ROI analysis of mean *z* scores in the precuneus (see Fig 2*C* for an illustration of the ROI selection) confirmed lower functional connectivity ($P < .05$, with and without controlling for covariates) (Fig 2*D*). No other meaningful functional components obtained from the ICA showed significant group differences in the dual-regression analyses.

Imaging Findings for Cohort 2. DTI TBSS analysis did not show any clusters or imaging voxels with significant white matter FA value differences between children born by cesarean or vaginal delivery. Likewise, VBM did not show any clusters or imaging voxels with significant regional gray matter volume differences between groups. Excluding/including potential confounders (age, sex, infant diet) as covariates did not change the results.

Imaging Findings for Cohort 3. Significantly different brain developmental trajectories were observed for the 2 groups (Fig 3). Specifically, at ~3 months of age, the MWF for the infants born by vaginal delivery was higher (indicating better myelination) in the frontal, temporal, parietal, and occipital white matter and the body of corpus callosum compared with those born by cesarean delivery. The differences of MWF in white matter regions between groups were consistent during infancy but gradually decreased and were not observable at age ~3 years or beyond when the MWF for both groups reached the same plateau.

DISCUSSION

The imaging results from cohort 1 showed striking differences in brain structural connectivity (measured by DTI-TBSS) and functional connectivity (measured by RS-fMRI) in 2-week-old healthy neonates born by cesarean delivery compared with vaginal delivery, whereas measured imaging parameters (structural connectivity and brain volume) in 8-year-old healthy children in cohort 2 did not differ. These novel findings provide the first-ever evidence that cesarean delivery may be associated with a shift in brain development, at least during early infancy. Considering the potential importance of these findings and the post hoc nature of our analyses using data derived from a larger study, we sought confirmation in a separate cohort (cohort 3) in which imaging results from a different institution were evaluated with the investigator blinded to findings from cohorts 1 and 2. The study in cohort 3 confirmed the delivery mode–associated differences in brain white matter development during infancy and was consistent in showing that differences dissipated with age. The differences in white matter development (as reflected by FA and MWF values) during infancy associated with delivery mode were widespread in both cohorts 1 and 3, involving the frontal, temporal, and parietal lobes and the corpus callosum, but not the cerebellum; the latter is one of the first regions to be myelinated, usually before birth. The differences were not prominent beyond age ~3 years, when myelination in most brain white matter regions approaches completion. Differences in resting-state functional connectivity were observed for the DMN network for neonates in cohort 1, while at this age most brain functional networks are still being developed.

Birth mode–associated brain developmental differences were not observed in later childhood (ie, 8-year-old healthy children in cohort 2 and 5-year-old healthy children in cohort 3). One possibility is that the differences were ameliorated during the years of postnatal development. Myelination for the cesarean delivery children may eventually catch up with that in children delivered vaginally when the developmental curve reaches a plateau after

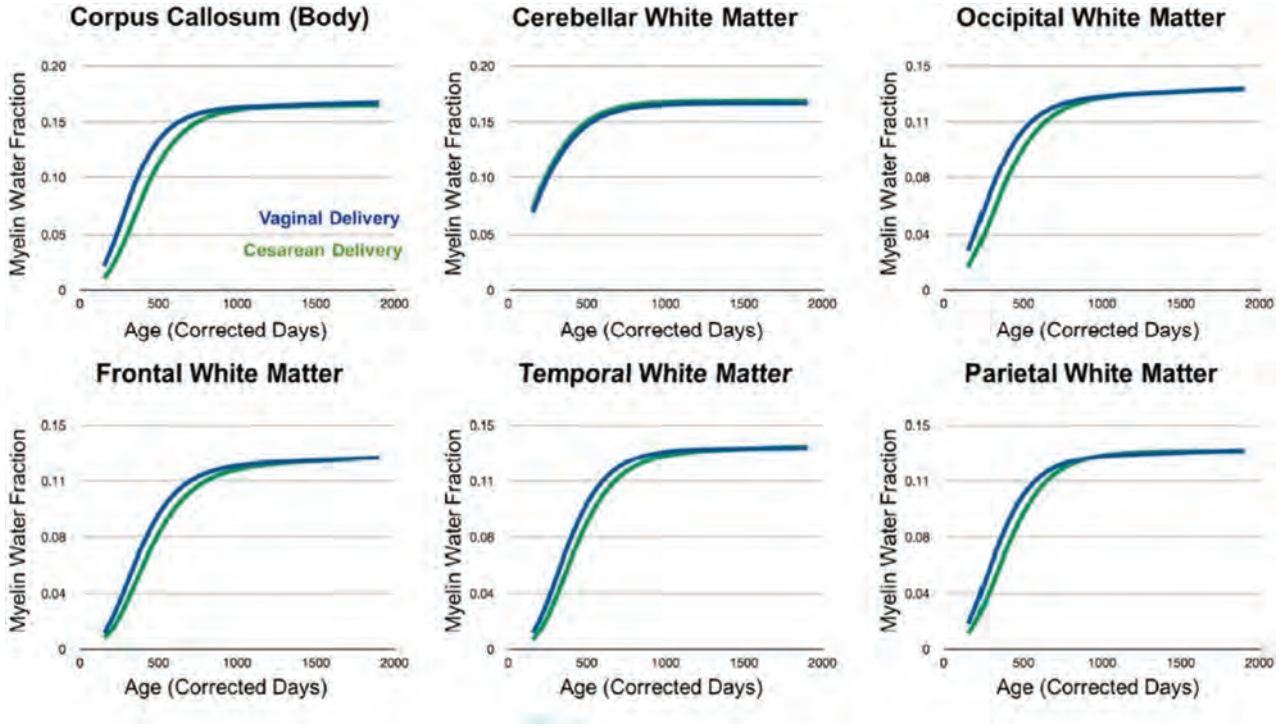


FIG 3. Longitudinal MWF in different brain regions for the vaginal delivery (blue) and cesarean delivery (green) groups in cohort 3. Children born by vaginal delivery had higher MWF (indicative of better white matter development) in most of the brain regions at young infancy, while the differences between groups gradually disappeared with age.

the first few years of life. In addition, breastfeeding (which reportedly promotes brain development in children[20,22,29]) may have driven enough changes to offset differences caused by cesarean delivery. Another possibility is the influence of the environment. While age, sex, and infant diet were included as covariates in cohort 2 and socioeconomic parameters were, in general, not different for the 2 groups in cohort 3, extensive exposures to environmental factors (childhood diet, social enrichment, or other aspects) may have contributed to changes in neurodevelopment with long-term effects stronger than birth mode. For example, cognitive development in healthy children has been associated with breakfast consumption[30] and composition,[31] physical activity,[32] television-viewing time,[33] and sleep duration.[34] Many of these variables during childhood are difficult to quantify and were not considered in our study of older children. Recent population studies showed an increased risk of autism spectrum disorder and attention deficit/hyperactivity disorder associated with cesarean delivery that did not persist if using sibling controls[11,35] and partially confirmed the confounding factors of family environment during childhood. In addition, while advanced and quantitative imaging methods were used in the evaluation of children's brain development in our study, it is possible that early brain structural differences associated with delivery mode introduced variations in brain function that could be apparent, for example, in the association of brain responses with stimulus-related information-processing.

The mechanisms underlying the effects of cesarean delivery on early life brain development are not known and could not be directly addressed in the current studies. Although cesarean delivery is generally considered a safe procedure, both mother and neonate are directly or indirectly subjected to factors that differ significantly from vaginal delivery, such as anesthesia/analgesics, surgical incisions, labor, trauma, stress, and so forth. It is not unreasonable to assume that these factors could have significant effects on the neonate's brain development. While speculative, we consider here a novel concept that associates cesarean delivery effects on neurodevelopment with an altered microbial environment. Several lines of evidence form the premise of this hypothesis. First, mode of delivery is a major determinant of gut microbiome composition in infants.[36,37] Vaginally delivered neonates acquire bacterial communities resembling maternal vaginal microbiota, while cesarean delivery neonates have bacterial communities similar to those found on the skin surface,[38] suggesting that the exposure to vaginal microbial environment during natural delivery is important for the initial establishment of neonate microbiota. Second, there is increasing evidence suggesting that gut microbiota modify central nervous system function and behavior, and microbiome-associated factors impact host immune activation, neural pathways, tryptophan metabolism and serotonin, gut hormone responses, and systemic exposures to bacterial metabolites.[39] One or more of these events could, in theory, impact neurophysiology and development. A recent study of microbiota-deficient adult rodents showed differences in the regulation of genes linked to myelination and myelin plasticity in the prefrontal cortex, which was reversed by colonization with conventional microbiota.[40] Third, a role for microglia and brain remodeling may be involved in response to the stressors and/or microbial-derived signals noted above.

One study in neonatal mice showed lower mitochondrial uncoupling protein 2 messenger RNA expression in the brain hippocampus associated with cesarean delivery.[41] Uncoupling protein 2 expression during the early postnatal period is important for neuronal differen-

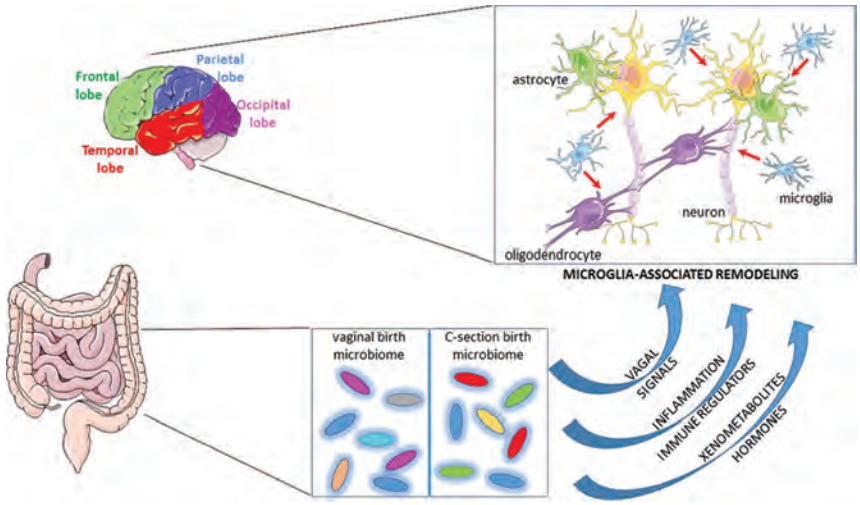


FIG 4. Illustration of a hypothetic model for mechanisms that drive differences in brains between infants with vaginal birth and cesarean delivery. Under this model, birth mode–associated changes to the gut microbiome lead to alterations in signals to the brain that regulate normal remodeling processes. Because microglia are implicated in this process, the model emphasizes this cell type as a major player in cell-cell cross-talk (*red arrows*) that modifies overall brain structure and function during the neonatal period. The model provides a theoretic framework for how birth mode contributes to normal neonatal brain development. The figure was generated, in part, by using components from the Servier Medical Art Powerpoint image bank (https://smart.servier.com/), which is under an open-use Creative Commons license.

tiation, axonal outgrowth, and synapse formation,[42] and uncoupling protein 2 is highly expressed in microglia and important in microglia function. Most important, microglia are implicated in modifying neural circuits during early brain development.[43] Thus, a working model is that cesarean delivery–associated changes in stress signals and/or the gut microbiome elicit a change in local signals of gut-brain cross-talk, in turn modifying brain microglia activities that impact brain development (Fig 4). Further study on animal models is necessary to evaluate this working model and explore the molecular mechanisms underlying cesarean delivery–associated changes in brain development and function.

There are several limitations in this study. First, it was based on secondary, post hoc analysis of MR imaging data for 3 cohorts, and the original study designs were not focused on detecting cesarean delivery effects on brain development. Second, because of age differences and institutional preferences, MR imaging methods were not the same for the 3 cohorts, and a complete characterization of the brain (ie, white/gray matter structure and function at rest and during tasks) for all subjects was not available. Nevertheless, the results from both institutions were consistent in illustrating brain developmental differences in infants and young children delivered by cesarean delivery that waned with age. Third, our results showed weakening delivery mode–associated differences on DTI measures after controlling for potential confounders. Gestational age, even for term pregnancy, and maternal BMI during pregnancy would also impact an infant's brain development, as shown by our previous publications. Nevertheless, the effects associated with cesarean delivery remained significant in cohort 1 after controlling for confounders for all measurements except the voxelwise TBSS. It is possible that there are additional confounders that may also interfere with the observation. In addition, reasons for cesarean delivery (eg, elective or emergent) and type of vaginal delivery (spontaneous or induced/instrument assisted) were not included as covariates because these parameters were not available for some subjects, and stratified analysis was not performed due to the limited sample size. Other factors during uncomplicated delivery, such as anesthesia use and length of labor, may also be potential confounders but were not considered. Despite limitations, this study provides the first evidence that cesarean delivery (which is usually complicated by multiple obstetric factors) may impact offspring's brain development. A prospective study specifically focused on this question, with a larger sample size, will be necessary to fully validate or refute this concept. Finally, our results were derived from healthy cohorts, so the association of delivery mode and brain development in compromised or at-risk infants could not be addressed. The strengths of our approach include statistical analyses considering relevant covariates and reproducibility in 2 independent cohorts at different sites.

CONCLUSIONS

Studies separately conducted at 2 independent institutions revealed significant effects of cesarean delivery on infant brain development, including reduced white matter microstructural integrity, weaker resting-state functional connectivity, and less myelination. While obstetric situations necessitating cesarean delivery not addressed in this study could also be potential confounding factors, these novel findings may have important clinical practice implications, given the increasing global prevalence of cesarean delivery. Nevertheless, there is no evidence from the current study that the effects are long-lasting at the brain anatomic level, considering that different structural measurements across age were used and functional connectivity were not assessed in the older cohorts. Additional studies are needed that look specifically at the impact of delivery mode on short- and long-term brain biology, neurocognition, learning, and behavior phenotypes.

ACKNOWLEDGMENTS

We thank the staff of the Arkansas Children's Nutrition Center Clinical Research Core, the Arkansas Children's Hospital Radiology MR imaging team, and the staff of the Advanced Baby Imaging Lab at Brown University for their assistance with these studies. We also thank Drs Aline Andres, Kartik Shankar, Mario Cleves, and Elisabet Borsheim for helpful discussion and/or reviewing this manuscript.

Disclosures: Sean C. Deoni—*RELATED*: National Institutes of Health, *Comments*: National Institute of Mental Health award as listed in the funding section*; *UNRELATED*: *Consultancy*: Nestlé Nutrition, *Comments*: consultant fees associated with nutritional impacts on neurodevelopment; *Grants/Grants Pending*: Nestlé Nutrition, *Comments*: study of nutritional impact on brain development*; *Payment for Lectures Including Service on Speakers Bureaus*: Wyeth Nutrition, Nestlé Nutrition, *Comments*: speaking fees for lectures on early brain development. Sean H. Adams—

RELATED: *Grant*: US Department of Agriculture–Agricultural Research Service, *Comments*: The Arkansas Children's Nutrition Center is funded as a national research center by the US Department of Agriculture–Agricultural Research Service*; *Support for Travel to Meetings for the Study or Other Purposes*: US Department of Agriculture–Agricultural Research Service, *Comments*: The Arkansas Children's Nutrition Center is funded as a national research center by the US Department of Agriculture–Agricultural Research Service; the Center has travel support for its investigators to attend scientific conferences*; *Provision of Writing Assistance, Medicines, Equipment, or Administrative Support*: US Department of Agriculture–Agricultural Research Service, *Comments*: The Arkansas Children's Nutrition Center is funded as a national research center by the US Department of Agriculture–Agricultural Research Service. The Center has administrative support for its investigators*; *UNRELATED*: *Consultancy*: roundtable participant (eg, International Life Sciences Institute) and meeting organizer (eg, for the National Dairy Council); *Employment*: University of Arkansas for Medical Sciences; *Grants/Grants Pending*: US Department of Agriculture–Agricultural Research Service, *Comments*: The Arkansas Children's Nutrition Center is funded as a national research center by the US Department of Agriculture–Agricultural Research Service. This is not in conflict with the current article; *Payment for Lectures Including Service on Speakers Bureaus*, *Comments*: Various honoraria for invited seminars at universities or giving talks at national scientific conferences are not in conflict; topic areas have nothing to do with brain function or the subject matter of the current article; *Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed*: The Arkansas Children's Nutrition Center is funded as a national research center by the US Department of Agriculture–Agricultural Research Service; the Center has travel support for its investigators to attend scientific conferences.* Thomas M. Badger—*RELATED*: *Grant*: National Institutes of Health and US Department of Agriculture*. Rudolph T. Pivik—*RELATED*: *Grant*: US Department of Agriculture–Agricultural Research Service Project 6026-51000-010-05S*; *UNRELATED*: *Employment*: University of Arkansas for Medical Sciences, *Comments*: only my salary as Research Professor, Department of Pediatrics. Charles M. Glasier—*RELATED*: *Grant*: Arkansas Children's Hospital; *UNRELATED*: *Expert Testimony*: medicolegal cases; *Payment for Lectures Including Service on Speakers Bureaus*: Cincinnati Children's Hospital. Amy C. Rowell—*RELATED*: *Grant*: US Department of Agriculture.* Xiawei Ou—*RELATED*: *Grant*: US Department of Agriculture–Agricultural Research Service Project 6026-51000-010-05S.* *Money paid to the institution.

REFERENCES

1. World Health Organization Human Reproduction Programme, 10 April 2015. **WHO Statement on caesarean section rates.** *Reprod Health Matters* 2015;23:149–50 CrossRef Medline
2. Martin JA, Hamilton BE, Osterman MJ, et al. **Births: final data for 2013**. *Natl Vital Stat Rep* 2015;64:1–65
3. Osterman MJ, Martin JA. **Trends in low-risk cesarean delivery in the United States, 1990–2013**. *Natl Vital Stat Rep* 2014;63:1–16
4. Cesarean delivery on maternal request. Committee Opinion No. 559. American College of Obstetricians and Gynecologists. Obstet Gynecol 2013:121;904–7.
5. **National Institutes of Health state-of-the-science conference statement: cesarean delivery on maternal request March 27–29, 2006**. *Obstet Gynecol* 2006;107:1386–97 CrossRef Medline
6. Barber EL, Lundsberg LS, Belanger K, et al. **Indications contributing to the increasing cesarean delivery rate.** *Obstet Gynecol* 2011;118:29–38 CrossRef Medline
7. Kozhimannil KB, Law MR, Virnig BA. **Cesarean delivery rates vary tenfold among US hospitals; reducing variation may address quality and cost issues.** *Health Aff (Millwood)* 2013;32:527–35 CrossRef Medline
8. Hasab Allah MF, El Adawy AR, Moustafa MF, et al. **Effect of mode of delivery on children intelligence quotient at pre-school age in El-Minia City.** *J Am Sci* 2012;8:1188–98
9. Khadem N, Khadivzadeh T. **The intelligence quotient of school aged children delivered by cesarean section and vaginal delivery.** *Iran J Nurs Midwifery Res* 2010;15:135–40 Medline
10. Amiri S, Malek A, Sadegfard M, et al. **Pregnancy-related maternal risk factors of attention-deficit hyperactivity disorder: a case-control study.** *ISRN Pediatr* 2012;2012:458064 CrossRef Medline
11. Curran EA, Dalman C, Kearney PM, et al. **Association between obstetric mode of delivery and autism spectrum disorder a population-based sibling design study.** *JAMA Psychiatry* 2015;72:935–42 CrossRef Medline
12. Curran EA, O'Neill SM, Cryan JF, et al. **Research review: birth by caesarean section and development of autism spectrum disorder and attention-deficit/hyperactivity disorder—a systematic review and meta-analysis.** *J Child Psychol Psychiatry* 2015;56:500–08 CrossRef Medline
13. Talge NM, Allswede DM, Holzman C. **Gestational age at term, delivery circumstance, and their association with childhood attention deficit hyperactivity disorder symptoms.** *Paediatr Perinat Epidemiol* 2016;30:171–80 CrossRef Medline
14. Glasson EJ, Bower C, Petterson B, et al. **Perinatal factors and the development of autism: a population study.** *Arch Gen Psychiatry* 2004;61:618–27 CrossRef Medline
15. Chien LN, Lin HC, Shao YH, et al. **Risk of autism associated with general anesthesia during cesarean delivery: a population-based birth-cohort analysis.** *J Autism Dev Disord* 2015;45:932–42 CrossRef Medline
16. Dean DC 3rd, Jerskey BA, Chen KW, et al. **Brain differences in infants at differential genetic risk for late-onset Alzheimer disease: a cross-sectional imaging study.** *JAMA Neurol* 2014;71:11–22 CrossRef Medline
17. Li X, Andres A, Shankar K, et al. **Differences in brain functional connectivity at resting-state in neonates born to healthy obese or normal-weight mothers.** *Int J Obes (Lond)* 2016;40:1931–34 CrossRef Medline
18. Ou X, Thakali KM, Shankar K, et al. **Maternal adiposity negatively influences infant brain white matter development.** *Obesity* 2015;23:1047–54 CrossRef Medline
19. Deoni SC, Dean DC, O'Muircheartaigh J, et al. **Investigating white matter development in infancy and early childhood using myelin water faction and relaxation time mapping.** *Neuroimage* 2012;63:1038–53 CrossRef Medline
20. Ou X, Andres A, Cleves MA, et al. **Sex specific association between infant diet and white matter integrity in 8-y-old children.** *Pediatr Res* 2014;76:535–43 CrossRef Medline
21. Ou X, Andres A, Pivik RT, et al. **Brain grey and white matter differences in healthy normal weight and obese children.** *J Magn Reson Imaging* 2015;42:1205–13 CrossRef Medline
22. Ou X, Andres A, Pivik RT, et al. **Voxel-based morphometry and fMRI revealed differences in brain gray matter in breastfed and milk formula-fed children.** *AJNR Am J Neuroradiol* 2016;37:713–19 CrossRef Medline
23. Ou X, Glasier CM, Ramakrishnaiah RH, et al. **Impaired white matter development in extremely low-birth-weight infants with previous brain hemorrhage.** *AJNR Am J Neuroradiol* 2014;35:1983–89 CrossRef Medline
24. Smith SM, Jenkinson M, Johansen-Berg H, et al. **Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data.** *Neuroimage* 2006;31:1487–505 CrossRef Medline
25. Smith SM, Nichols TE. **Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference.** *Neuroimage* 2009;44:83–98 CrossRef Medline
26. Ou X, Glasier CM, Ramakrishnaiah RH, et al. **Gestational age at birth and brain white matter development in term-born infants and children.** *AJNR Am J Neuroradiol* 2017;38:2373–79 CrossRef Medline
27. Reiss AL, Abrams MT, Singer HS, et al. **Brain development, gender and IQ in children: a volumetric imaging study.** *Brain* 1996;119:1763–74 CrossRef Medline
28. Gao W, Alcauter S, Smith JK, et al. **Development of human brain cortical network architecture during infancy.** *Brain Struct Funct* 2015;220:1173–86 CrossRef Medline
29. Deoni SC, Mercure E, Blasi A, et al. **Mapping infant brain myelination with magnetic resonance imaging.** *J Neurosci* 2011;31:784–91 CrossRef Medline
30. Liu JH, Hwang WT, Dickerman B, et al. **Regular breakfast consumption is associated with increased IQ in kindergarten children.** *Early Hum Dev* 2013;89:257–62 CrossRef Medline
31. Taki Y, Hashizume H, Sassa Y, et al. **Breakfast staple types affect**

brain gray matter volume and cognitive function in healthy children. *PLoS One* 2010;5:e15213 CrossRef Medline
32. Sibley BA, Etnier JL. **The relationship between physical activity and cognition in children: a meta-analysis.** *Pediatric Exercise Science* 2003;15:243–56 CrossRef
33. Zimmerman FJ, Christakis DA. **Children's television viewing and cognitive outcomes: a longitudinal analysis of national data.** *Arch Pediatr Adolesc Med* 2005;159:619–25 CrossRef Medline
34. Dahl RE. **The impact of inadequate sleep on children's daytime cognitive function.** *Semin Pediatr Neurol* 1996;3:44–50 CrossRef Medline
35. Curran EA, Khashan AS, Dalman C, et al. **Obstetric mode of delivery and attention-deficit/hyperactivity disorder: a sibling-matched study.** *Int J Epidemiol* 2016;45:532–42 CrossRef Medline
36. Bäckhed F, Roswall J, Peng YQ, et al. **Dynamics and stabilization of the human gut microbiome during the first year of life.** *Cell Host Microbe* 2015;17:690–703 CrossRef Medline
37. Penders J, Thijs C, Vink C, et al. **Factors influencing the composition of the intestinal microbiota in early infancy.** *Pediatrics* 2006;118:511–21 CrossRef Medline
38. Dominguez-Bello MG, Costello EK, Contreras M, et al. **Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns.** *Proc Natl Acad Sci U S A* 2010;107:11971–75 CrossRef Medline
39. Borre YE, O'Keeffe GW, Clarke G, et al. **Microbiota and neurodevelopmental windows: implications for brain disorders.** *Trends Mol Med* 2014;20:509–18 CrossRef Medline
40. Hoban AE, Stilling RM, Ryan FJ, et al. **Regulation of prefrontal cortex myelination by the microbiota.** *Transl Psychiatry* 2016;6:e774 CrossRef Medline
41. Simon-Areces J, Dietrich MO, Hermes G, et al. **UCP2 induced by natural birth regulates neuronal differentiation of the hippocampus and related adult behavior.** *PLoS One* 2012;7:e42911 CrossRef Medline
42. Seli E, Horvath TL. **Natural birth-induced UCP2 in brain development.** *Rev Endocr Metab Disord* 2013;14:347–50 CrossRef Medline
43. Tremblay MÈ, Stevens B, Sierra A, et al. **The role of microglia in the healthy brain.** *J Neurosci* 2011;31:16064–69 CrossRef Medline

MR Imaging of the Brain in Neurologic Wilson Disease

X.-E. Yu, S. Gao, R.-M. Yang, and Y.-Z. Han

ABSTRACTS

BACKGROUND AND PURPOSE: Neurologic Wilson disease is an inherited disease characterized by a copper metabolic disorder that causes damage to many organs, especially the brain. Few studies report the relationships between these neurologic symptoms and MR imaging of the brain. Therefore, we investigated the correlation of brain abnormalities in patients with neurologic Wilson disease with their clinical symptoms, age of onset, and lag time to diagnosis.

MATERIALS AND METHODS: A cohort of 364 patients was recruited in China between January 2003 and December 2017. Age of onset, lag time until diagnosis, and neurologic symptoms were recorded, and cranial MR imaging was performed. Patients were divided into groups within each of these factors for correlation analysis with the MR imaging brain scans.

RESULTS: Abnormal signals in the MR imaging brain scans were seen in all 364 cases. Affected regions included the putamen, pons, midbrain, and thalamus, while the medulla and occipital lobe were unaffected. The putamen was the most frequently damaged brain region in this study. With the age of onset younger than 10 years, cranial MR imaging scans showed only impairment in the putamen. Patients with a longer lag time before diagnosis were more likely to have impairment in the pons, midbrain, and cortex. Among neurologic symptoms of Wilson disease, torsion spasm is associated with the midbrain and cortex, and choreoathetosis is related to the caudate nucleus.

CONCLUSIONS: Abnormalities in the putamen, pons, midbrain, and thalamus are part of the neuroimaging spectrum of Wilson disease. There is a significant correlation between the site of brain injury and diagnosis lag time.

ABBREVIATION: WD = Wilson disease

Wilson disease (WD), also known as hepatolenticular degeneration, is an autosomal recessive disorder of human copper metabolism,[1,2] caused by pathogenic variants in the copper-transporting gene *ATP7B*.[3-5] WD leads to intracellular copper accumulation, causing damage to many organs, especially the brain.[6-8] Neurologic WD is one of the main forms of the disease, with some patients showing severe neurologic symptoms that persist despite treatment. Further neurologic deterioration may be observed even after treatment initiation.[9-12]

MR imaging is a sensitive method to evaluate the brains of patients with neurologic WD.[8,13] Whereas abnormalities in the putamen are the most common feature of neurologic WD,[14] brain shrinkage is also frequently observed.[15] It has been suggested that the clinical manifestations of WD are region-specific,[16] and there may be specific differences among patients with WD in China. Therefore, this study aimed to explore the characteristics of cranial MR imaging among patients with WD in China. Many reports have recently published MR imaging data of patients with WD, but the sample sizes of these studies have been limited.[17-19] Moreover, investigations of the relationship between abnormal findings on MR imaging and neurologic symptoms in patients with WD are sparse. In this study, we conducted MR imaging in 364 patients with WD before they received any treatment for the disease. We provide a detailed report of their MR imaging in conjunction with factors such as age of onset, lag time until diagnosis, and clinical symptoms.

MATERIALS AND METHODS

Study Participants

The study included 364 patients recruited from the Affiliated Hospital of the Institute of Neurology at Anhui University of Chi-

Received June 26, 2018; accepted after revision October 30.

From the Department of Pharmacology (Y.X.-E., S.G.), Basic Medical College, Anhui Medical University, Hefei, China; and Hospital of the Institute of Neurology (X.-E.Y., R.-M.Y., Y.-Z.H.), Anhui University of Chinese Medicine, Hefei, China.

This work was funded by the Natural Science Foundation of Anhui Province (1508085MH153).

Please address correspondence to Shan Gao, PhD, Department of Pharmacology, Basic Medical College, Anhui Medical University, Hefei 230032, China; e-mail: gaoshan1746@sina.com

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

http://dx.doi.org/10.3174/ajnr.A5936

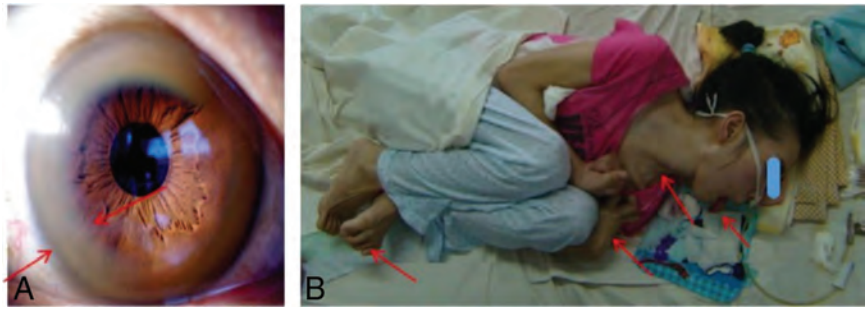


FIG 1. *A*, Kayser-Fleischer rings observable with the unaided eye (*arrows*). *B*, Wilson disease in a 23-year-old women with torsion spasm (*arrows*).

nese Medicine between January 2003 and December 2017. Within this cohort, 219 patients were male and 145 were female. All data were collected at the time of the patient's hospitalization, and the data were summarized in a timely manner.

Diagnostic Criteria

The criteria used for recruitment to this study, as is standard for WD diagnoses, were as follows[7,20-22]:

1) Family heredity: parental consanguinity, having a sibling with WD, or having a sibling who died from unexplained liver disease.
2) Presence of an extracorticospinal tract symptom such as dystonia, parkinsonism, chorea, torsion spasm, or liver symptoms.
3) Kayser-Fleischer rings visible to the naked eye or with slit-lamp examination.
4) Ceruloplasmin level <2.16 μmol/L or serum copper oxidase level <0.20 optical density units.
5) Urinary copper excretion level of >1.6 μmol every 24 hours.
6) Hepatic copper concentration via needle biopsy of >250 μg/g.

Participants fulfilling both criteria 1 and 4 or both criteria 2 and 4 were diagnosed with symptomatic WD. Participants fulfilling both criteria 3 and 5, only criterion 4, or only criterion 6 were diagnosed with asymptomatic WD. Only patients with symptomatic WD were included in the cohort of this study.

Study Measures and Group Analysis

Age of disease onset, diagnosis lag time (time from symptom onset to definite diagnosis), and type of clinical neurologic symptoms were recorded, and groups were established for subsequent correlation analysis. Patients were divided into 5 groups based on the age of onset, as follows: younger than 10 years, 10–14 years, 15–30 years, 31–40 years, and older than 40 years. In terms of diagnosis lag time, patients were divided into 5 groups: <3 months, 4–6 months, 7–12 months, 1–3 years, and > years. Finally, the 5 neurologic symptom groups used for correlation analysis were dystonia, parkinsonism, torsion spasm, and choreoathetosis.

MR Imaging

A 1.5T MR imaging scanner (Avanto; Siemens, Erlangen, Germany) (T1–TR = 210.0 ms, TE = 3.8 ms, time of acquisition = 46.83 seconds, bandwidth = 160.0 Hz/Px, T2–TR = 3300.0 ms, TE = 99.0 ms, time of acquisition = 56.27 seconds, bandwidth = 195.0 Hz/Px) was used to image the T1 and T2 signals in the basal ganglia, brain stem, and cerebral cortex of participants. The basal ganglia images included the putamen, caudate, and globus pallidus as well as the thalamus. Images from the brain stem included the midbrain, pons, and medulla. Scans of the cerebral cortex showed the frontal, temporal, parietal, and occipital lobes. All MR images were analyzed separately by 2 researchers (X.Y. and R.Y.). Disagreements about the findings were resolved by discussion until a consensus was reached. This MR imaging protocol was followed for the entire 15-year study period.

Statistical Analyses

The Kruskal-Wallis test was used to assess the statistical difference in the mean age of onset among the groups. The MR imaging manifestations were used to determine the statistical differences in neurologic symptoms among the groups. Continuous data with a normal distribution are presented as mean ± SD, and the frequencies of categoric variables are presented as sample size with percentage. All data were analyzed using SPSS, Version 10.01 (IBM, Armonk, New York), and $P < .05$ was considered statistically significant.

RESULTS

After diagnostic testing, we found that all patients showed ceruloplasmin levels of <2.16 μmol/L, serum copper oxidase levels of <0.2 optical density units, and urinary copper excretion levels of >1.6 μmol per 24 hours. All participants were positive on slit-lamp examination for Kayser-Fleischer rings (Fig 1*A*). In addition to the presence of neurologic symptoms, the above-mentioned values resulted in the clear diagnosis of symptomatic WD in all 364 patients included in our study. According to participants' neurologic symptoms, we found 197 cases of dystonia, 127 cases of parkinsonism, 32 cases of torsion spasm, and 8 cases of choreoathetosis (Fig 1*B*).

All patients underwent cranial MR imaging before their initial copper chelation therapy. The scans showed long T1 and long T2 abnormal signal in the brain (Fig 2*A*). All 364 patients with WD showed abnormal signal in at least 1 examined brain region (Table 1). Most patients showed abnormal signal in the following brain regions: 81.0% in the putamen (Fig 2*A*), 46.4% in the pons (Fig 2*B*), 36.0% in the thalamus (Fig 2*C*), 34.3% in the caudate nucleus (Fig 2*D*), 33.5% in the midbrain (Fig 2*E*), and 15.4% in the globus pallidus (Fig 2*F*). Brain regions that were less often affected by WD were the temporal lobe (in 3.3% of patients, Fig 2*G*), parietal lobe (3.3%, Fig 2*H*), frontal lobe (3.2%, Fig 2*I*), cerebellum (2.2%; Fig 2*J*), and corpus callosum (in 2.2%, Fig 2*K*). No abnormalities were observed in the occipital lobe (Fig 2*L*) or medulla oblongata (Fig 2*M*).

The most common areas with brain damage in our patients with WD with dystonia were the putamen, pons, and thalamus. However, cranial MR imaging of patients with WD with torsion spasm showed extensive impairment in the putamen, pons, and midbrain. Compared with patients with WD who had other neurologic symptoms, the cerebral cortex was more often impaired in

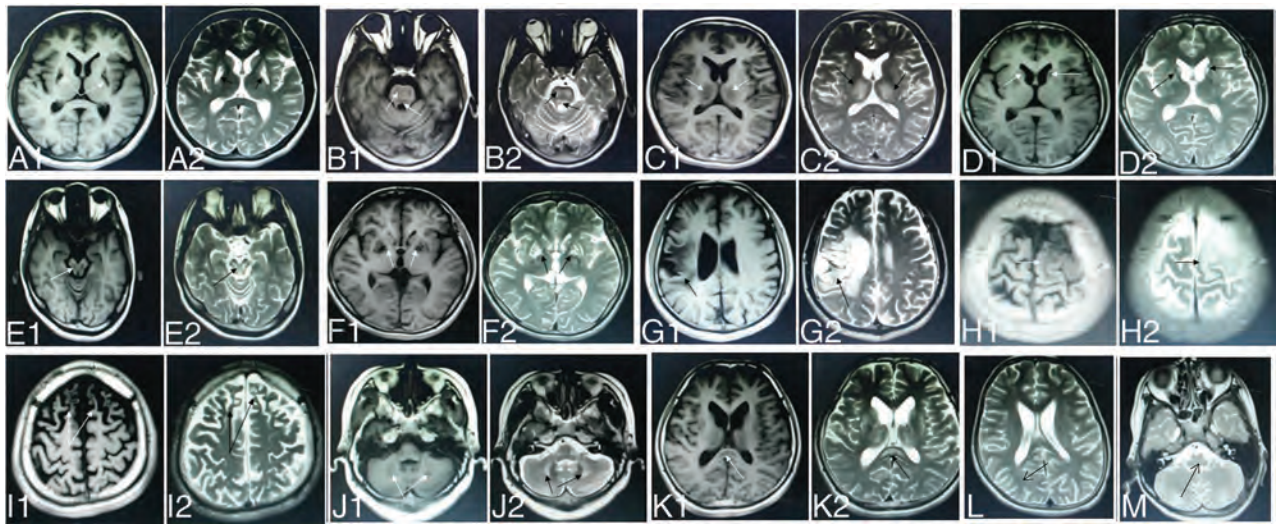


FIG 2. WD in a 14-year-old boy with dysarthria and dystonia with abnormal signal in the putamen (*A1*, *A2*). WD in a 19-year-old woman with dysphagia and dystonia with abnormal signal in the pons (*B1*, *B2*). WD in a 24-year-old woman with dysarthria and parkinsonism with abnormal signal in the thalamus (*C1*, *C2*). WD in an 18-year-old man with choreoathetosis with abnormal signal in the caudate nucleus (*D1*, *D2*). WD in a 22-year-old woman with dysarthria and parkinsonism with abnormal signal in the midbrain (*E1*, *E2*). WD in a 19-year-old woman with dystonia with abnormal signal in the globus pallidus (*F1*, *F2*). WD in a 27-year-old man with torsion spasm with abnormal signal in the temporal lobe (*G1*, *G2*). WD in a 19-year-old man with epilepsy and dysarthria with abnormal signal in the parietal lobe (*H1*, *H2*). WD in a 17-year-old boy with psychiatric symptoms with abnormal signal in the frontal lobe (*I1*, *I2*). WD in an 18-year-old man with dysarthria and parkinsonism with abnormal signal in the cerebellum (*J1*, *J2*). WD in a 13-year-old girl with parkinsonism with abnormal signal in the corpus callosum (*K1*, *K2*). WD in a 23-year-old woman with dystonia with normal signal in the medulla oblongata (*L*). WD in a 27-year-old man with dysarthria and dystonia with normal signal in the occipital lobe (*arrows*) (*M*).

Table 1: Correlations between region-specific brain damage in neurologic WD and accompanying neurologic symptoms[a]

Neurologic Symptoms	No. of Cases	Putamen	Globus Pallidus	Caudate Nucleus	Thalamus	Midbrain	Pons	Cerebellum	Frontal Lobe	Temporal Lobe	Parietal Lobe	Corpus Callosum
Dystonia	197	159 (80.7)	34 (17.3)	66 (33.5)	73 (37.1)	67 (34.0)	95 (48.2)	4 (2.0)	4 (2.0)	3 (1.5)	3 (1.5)	
Parkinsonism	127	104 (81.9)	19 (15.0)	41 (32.3)	45 (35.4)	31 (15.7)	52 (26.4)	4 (3.1)		3 (2.4)	3 (2.4)	6 (4.7)
Torsion spasm	32	25 (78.1)	2 (6.3)	10 (31.2)	8 (25)	18 (56.3)	20 (62.5)		7 (21.9)	6 (18.8)	6 (18.8)	2 (6.3)
Choreoathetosis	8	7 (87.5)	1 (12.5)	8 (100)	5 (62.5)	6 (75)	2 (25)					
All patients	364	295 (81.0)	56 (15.4)	125 (34.3)	131 (36.0)	122 (33.5)	169 (46.4)	8 (2.2)	11 (3.2)	12 (3.3)	12 (3.3)	8 (2.2)

[a] Data are number and percentage.

Table 2: Correlations between age of onset and region-specific brain damage in neurologic WD[a]

Age of Onset (yr)	No. of Cases	Putamen	Globus Pallidus	Caudate Nucleus	Thalamus	Midbrain	Pons	Cerebellum	Frontal Lobe	Temporal Lobe	Parietal Lobe	Corpus Callosum
Younger than 10	25	25 (100)										
10–14	112	107 (95.5)		31 (27.7)	47 (42.0)	52 (46.4)	74 (66.1)	2 (1.8)	3 (2.7)	2 (1.8)	3 (2.7)	8 (7.1)
15–30	212	151 (71.2)	48 (22.6)	88 (41.5)	78 (36.8)	67 (31.6)	81 (38.2)	5 (2.4)	8 (3.8)	9 (4.2)	8 (3.8)	
31–40	14	11 (78.6)	8 (57.1)	5 (35.7)	7 (50)	3 (21.4)	3 (21.4)	1 (7.1)		1 (7.1)	1 (7.1)	
Older than 40	1	1 (100)		1 (100)	1 (100)		1					

[a] Data are number and percentage.

the group with torsion spasm; the probability of damage in ≥1 of the cortical lobes in this group was approximately 18%–21%. Taken together, our MR imaging data show that the putamen is the most frequently damaged brain region in patients with WD, regardless of the accompanying neurologic symptoms, followed by the pons and the thalamus. Whereas the symptom of torsion spasm in WD seems to indicate abnormalities in the midbrain and cortex, the symptom of choreoathetosis predominantly suggests damage in the caudate nucleus.

In our patient cohort, age of disease onset varied from 5 to 42 years, with an average onset at 17.04 ± 6.13 years of age. Table 2 shows that the age of onset in most patients with WD was younger than 30 years (95.9% of all patients), among whom 6.9% were younger than 10 years at disease onset, 30.8% were 10–14 years of age, and 58.2% were 15–30 years of age. In the group with age of onset younger than 10 years, cranial MR imaging scans showed impairment only in the putamen. However, in the group with age of onset of 10–14 years, cranial MR imaging scans very frequently revealed impairment in the brain stem.

The diagnosis lag time, defined as the time between initial symptom onset and confirmed diagnosis, averaged 3.78 ± 4.46 years and varied from 1 month to 23 years. Table 3 shows that only 74 patients were diagnosed within 6 months of symptom onset. Our data showed that an increased frequency of the presence of brain damage correlated with an increased delay in treatment initiation. In the group with a lag time of <3 months, 90.6% of participants showed putamen impairment (Fig 2*A*) and a few had impaired thalami (Fig 2*C*) and pontes (Fig 2*B*). The latter patients had no defects in the midbrain or cerebral cortex. Our research revealed that diagnosis lag time does not correlate with

Table 3: Correlations between diagnosis lag time and region-specific brain damage in neurologic WD[a]

Time Lag to Diagnosis	No. of Cases	Putamen	Pallidum	Caudate Nucleus	Thalamus	Midbrain	Pons	Cerebellum	Frontal Lobe	Temporal Lobe	Parietal Lobe	Corpus Callosum
≤3 mo	32	29 (90.6)		3 (9.4)	11 (34.4)		8 (25)					3 (9.4)
4–6 mo	42	37 (88.1)	8 (19.0)	5 (11.9)	12 (28.6)	9 (21.4)	10 (23.8)		1 (2.4)	2 (4.8)	1 (2.4)	5 (11.9)
7–12 mo	76	69 (90.8)	14 (18.4)	31 (40.8)	17 (22.4)	20 (26.3)	28 (36.8)	2 (2.6)	2 (2.6)	2 (2.6)	1 (1.3)	
1–3 yr	87	69 (79.3)	5 (5.7)	13 (14.9)	24 (27.6)	19 (21.8)	37 (42.5)	5 (5.7)	5 (5.7)	5 (5.7)	4 (4.6)	
>3 yr	127	91 (71.7)	29 (22.8)	73 (57.5)	67 (52.8)	74 (58.3)	86 (67.7)	1 (0.8)	3 (2.4)	3 (2.4)	6 (4.7)	

[a] Data are number and percentage.

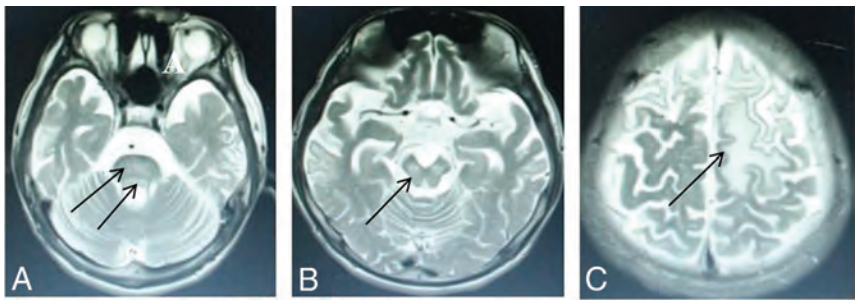


FIG 3. High signal intensity lesions on T2WI. Wilson disease in a 14-year-old girl with dystonia, with a diagnosis lag time 5 years (participant group >3 years) and abnormal signal in the pons (*D*), midbrain (*E*), and frontal and parietal lobes (*F*) (*arrows*).

impairment of the putamen in WD but does correlate with the severity of accompanying neurologic symptoms and the likelihood of impairment in the pons (Fig 3*A*), midbrain (Fig 3*B*), and cerebral cortex (Fig 3*C*). The *P* value obtained by the Fisher exact probability method was .001, and the difference was statistically significant. The MR imaging influence in the brain region was related to the course of disease.

DISCUSSION

Neurologic WD is a genetic disorder that leads to intracellular accumulation of copper in the body, primarily causing damage to the liver and various regions of the brain. Previous studies of MR imaging brain scans in patients with WD have shown abnormal bilateral long T1 and T2 signals in the basal ganglia, especially in the putamen nucleus and head of the caudate nucleus and often in some portion of the thalamus.[23] Abnormal bilateral long T1 and T2 signals in the basal ganglia (with or without brain stem impairment) have become one of the important signs of WD.[7,23,24] Consistent with the literature, our results also showed long T1 and long T2 abnormal signals in the brain, with the most frequently affected regions including the putamen and caudate nucleus of the basal ganglia (Table 1). The aberrant signal in the basal ganglia is said to be elicited by glial cell hyperplasia as well as edema, necrosis, and lacunae caused by copper deposition.[25] However, a study by Prayer et al[26] showed that some abnormal signals caused by edema and demyelination can disappear after proper de-coppering therapy. The medulla and occipital lobe were entirely spared, with none of our patients with WD showing abnormalities in these regions (Table 1).

Dystonia is among the most debilitating clinical neurologic symptoms commonly observed in neurologic WD.[11,27,28] In our study, the most common areas of brain damage in patients with WD with dystonia were the putamen, pons, and thalamus. Regardless of etiology, symptomatic treatment of dystonia generally depends on its severity and localization.[29] According to our findings, dystonia may be related to the putamen, pons, and thalamus.

Similar to dystonia, parkinsonism occurs quite frequently in neurologic WD. Patients with WD show combined presynaptic and postsynaptic nigrostriatal deficits.[30] In general neurology, tremor treatment varies depending on the tremor type.[31,32] In our patients, the unanticipated finding of corpus callosum abnormalities in the group with parkinsonism resulted in corpus callosum repair in some individuals, with marked improvement of their symptoms. In our study, the most common areas of brain damage in patients with WD with parkinsonism were the putamen, thalamus, and caudate nucleus. Our study findings suggest that parkinsonism may be related to the putamen, thalamus, and caudate nucleus.

Not all of our patients with WD with abnormal caudate nuclei had choreoathetosis. We found that 7 of the 8 patients with WD with choreoathetosis symptoms had an abnormal putamen, and all 8 had an abnormal caudate nucleus. This is basically consistent with the authors' previous research[8] indicating that the symptom of choreoathetosis strongly correlates with impairment of the caudate nucleus. Thus, the presence of choreoathetosis predominantly suggests damage in the caudate nucleus.

Torsion spasms are a rare neurologic symptom of WD, which contribute to patient disability when present.[7] MR imaging data from our patients showed frequent damage to the midbrain and pons but no damage to the remaining regions of the brain stem and medulla oblongata. Taken together, our MR imaging data show that the putamen is the most frequently damaged brain region in patients with WD, regardless of the accompanying neurologic symptoms, followed by the pons and the thalamus.

The data in Table 2 seem to imply that the disease gradually worsens with age before of 30 years of age, but older age of onset is a mitigating factor in WD; however, we cannot ignore the possibility that the lower frequency of brain stem abnormalities and improved prognosis is owing to some other factor that we did not examine in this study, such as disease duration.

The prognosis of WD is closely related to the diagnosis lag time for each patient.[33] As our correlation analysis showed, the frequency of damage observed in the pons, midbrain, and cortex is directly correlated with the diagnosis lag time. Specifically, the longer the period between symptom onset and confirmed diagnosis, the greater the likelihood that the patient will develop impairment in these brain regions. In the present study, the features

of WD characterized by torsion spasm included age of onset usually younger than 15 years and diagnosis lag time of at least 2 years (the longest was 23 years). From the trend in the length of diagnosis lag time seen in Table 3, we can speculate that the initial sites of brain injury in neurologic WD are the putamen and thalamus, then developing in the pons and midbrain, and culminating with cortical damage. Therefore, it is crucial for individuals with WD to seek medical care, receive a confirmed diagnosis, and initiate treatment as early as possible.

Differences between our results and those of other MR imaging studies of WD could be due to many reasons: Limited sample sizes, diversity of study participants, variations in the magnetic field strength of the MR imaging scanner used, and duration of therapy all contribute to variations in the results reported in the clinic. The clinical manifestations of WD may be region-specific,[16] making the demographic composition of the study cohort an important factor for interpretation of the collected data. With 364 patients with WD recruited only from China and MR images collected before the start of any systemic de-coppering therapy, our study has the crucial advantages of a large sample size focused within a single geographic region and being free from the confounding effects of prior treatment.

This study spans a long time, and MR imaging data from gradient-echo images and diffusion images have only become available in recent years. Therefore, most early studies lack these data, and this feature is a drawback of this study.

CONCLUSIONS

Our study findings showed that the putamen, pons, midbrain, and thalamus are highly prone to damage in neurologic WD, whereas the medulla oblongata and occipital lobe seem completely unaffected. The putamen is the most frequently damaged brain region in this study. In patients with age of onset younger than 10 years, cranial MR imaging scans showed impairment only in the putamen. Patients with a longer lag time before diagnosis were more likely to have impairment in the pons, midbrain, and cortex. Among neurologic symptoms of WD, torsion spasm is associated with the midbrain and cortex, and choreoathetosis is primarily related to the caudate nucleus. The brain abnormalities in patients with neurologic Wilson disease are correlated with their clinical symptoms, age of onset, and lag time to diagnosis.

Disclosures: Xu-En Yu—*RELATED*: *Grant*: Natural Science Foundation of Anhui Province, *Comments*: mitochondrial damage mechanism of mouse brain neurons in Wilson disease model TX (grant No. 1508085MH153).* Shan Gao—*RELATED*: *Grant*: Natural Science Foundation of Anhui Province, *Comments*: mitochondrial damage mechanism of mouse brain neurons in Wilson disease model TX (grant No. 1508085MH153).* Ren-Min Yang—*RELATED*: *Grant*: Natural Science Foundation of Anhui Province, *Comments*: mitochondrial damage mechanism of mouse brain neurons in Wilson disease model TX (grant No. 1508085MH153).* Yong-Zhu Han—*RELATED*: *Grant:* Natural Science Foundation of Anhui Province, *Comments*: mitochondrial damage mechanism of mouse brain neurons in Wilson disease model TX (grant No. 1508085MH153).* *Money paid to the institution.

REFERENCES

1. Bandmann O, Weiss KH, Kaler SG. **Wilson's disease and other neurological copper disorders.** *Lancet Neurol* 2015;14:103–13 CrossRef Medline
2. Ferenci P. **Pathophysiology and clinical features of Wilson disease.** *Metab Brain Dis* 2004;19:229–39 CrossRef Medline
3. Gromadzka G, Schmidt HH, Genschel J, et al. **p.H1069Q mutation in ATP7B and biochemical parameters of copper metabolism and clinical manifestation of Wilson's disease.** *Mov Disord* 2006;21:245–48 CrossRef Medline
4. Thomas GR, Forbes JR, Roberts EA, et al. **The Wilson disease gene: spectrum of mutations and their consequences.** *Nat Genet* 1995;9:210–17 CrossRef Medline
5. Zong Y, Kong X. **Mutation analysis of 35 Wilson's disease pedigrees** [in Chinese]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 2016;33:30–33 CrossRef Medline
6. Schilsky ML. **Wilson disease: diagnosis, treatment, and follow-up.** *Clin Liver Dis* 2017;21:755–67 CrossRef Medline
7. Yang RM. *Hepatolenticular Degeneration.* Beijing: People's Medical Publishing House; 2015
8. Yu XE, Yang RM. **132 cases brain imaging of hepatolenticular degeneration.** *J Apopl Nerv Dis* 2007;1:30–34
9. Litwin T, Dušek P, Czlonkowska A. **Symptomatic treatment of neurologic symptoms in Wilson disease.** *Handb Clin Neurol* 2017;142:211–23 CrossRef Medline
10. Litwin T, Dzieżyc K, Karliński M, et al. **Early neurological worsening in patients with Wilson's disease.** *J Neurol Sci* 2015;355:162–67 CrossRef Medline
11. Lorincz MT. **Neurologic Wilson's disease.** *Ann N Y Acad Sci* 2010;1184:173–87 CrossRef Medline
12. Lucato LT, Otaduy MC, Barbosa ER, et al. **Proton MR spectroscopy in Wilson disease: analysis of 36 cases.** *AJNR Am J Neuroradiol* 2005;26:1066–71 Medline
13. van Wassenaer-van Hall HN, van den Heuvel AG, Algra A, et al. **Wilson disease: findings at MR imaging and CT of the brain with clinical correlation.** *Radiology* 1996;198:531–36 CrossRef Medline
14. Ranjan A, Kalita J, Kumar S, et al. **A study of MRI changes in Wilson disease and its correlation with clinical features and outcome.** *Clin Neurol Neurosurg* 2015;138:31–36 CrossRef Medline
15. Kalita J, Naik S, Bhoi SK, et al. **Pontomesencephalic atrophy and postural instability in Wilson disease.** *AJNR Am J Neuroradiol* 2017;38:1343–47 CrossRef Medline
16. Yu XE, Hu WB, Han YZ, et al. **The regionalism of Wilson disease in Anhui Province.** *Chin J Nerv Ment Dis* 2012;38:435–38
17. Das M, Misra UK, Kalita J. **A study of clinical, MRI and multimodality evoked potentials in neurologic Wilson disease.** *Eur J Neurol* 2007;14:498–504 CrossRef Medline
18. Kim S, Song IU, Chung YA, et al. **Brain MRI, Tc-99m HMPAO SPECT and F-18 FP-CIT PET/CT findings in a patient with Wilson disease: a case report.** *Nucl Med Mol Imaging* 2014;48:303–05 CrossRef Medline
19. Park HK, Lee JH, Lee MC, et al. **Teaching NeuroImages: MRI reversal in Wilson disease with trientine treatment.** *Neurology* 2010;74:e72 CrossRef Medline
20. Chinese Medical Association Neurology Branch. **A guide to diagnosis and treatment of hepatolenticular degeneration.** *Chin J Neurol* 2008;41:566–69
21. Roberts EA, Schilsky ML. **Diagnosis and treatment of Wilson disease: an update.** *Hepatology* 2008;47:2089–111 CrossRef Medline
22. European Association for Study of Liver. **EASL Clinical Practice Guidelines: Wilson's disease.** *J Hepatol* 2012;56:671–85 CrossRef Medline
23. King AD, Walshe JM, Kendall BE, et al. **Cranial MR imaging in Wilson's disease.** *AJR Am J Roentgenol* 1996;167:1579–84 CrossRef Medline
24. Thuomas KA, Aquilonius SM, Bergström K, et al. **Magnetic resonance imaging of the brain in Wilson's disease.** *Neuroradiology* 1993;35:134–41 CrossRef Medline
25. Scheiber IF, Brůha R, Dušek P. **Pathogenesis of Wilson disease.** *Handb Clin Neurol* 2017;142:43–55 CrossRef Medline
26. Prayer L, Wimberger D, Kramer J, et al. **Cranial MRI in Wilson's disease.** *Neuroradiology* 1990;32:211–14 CrossRef Medline

27. Burke JF, Dayalu P, Nan B, et al. **Prognostic significance of neurologic examination findings in Wilson disease.** *Parkinsonism Relat Disord* 2011;17:551–56 CrossRef Medline
28. Dusek P, Litwin T, Czlonkowska A. **Wilson disease and other neurodegenerations with metal accumulations.** *Neurol Clin* 2015;33:175–204 CrossRef Medline
29. Thenganatt MA, Jankovic J. **Treatment of dystonia.** *Neurotherapeutics* 2014;11:139–52 CrossRef Medline
30. Oder W, Brücke T, Kollegger H, et al. **Dopamine D2 receptor binding is reduced in Wilson's disease: correlation of neurological deficits with striatal 123I-iodobenzamide binding.** *J Neural Transm (Vienna)* 1996;103:1093–103 CrossRef Medline
31. Puschmann A, Wszolek ZK. **Diagnosis and treatment of common forms of tremor.** *Semin Neurol* 2011;31:65–77 CrossRef Medline
32. Schneider SA, Deuschl G. **The treatment of tremor.** *Neurotherapeutics* 2014;11:128–38 CrossRef Medline
33. Yu XE, Hu WB, Han YZ, et al. **The course and prognosis of Wilson's disease.** *Chin J Neurol* 2012;45

MRI Abnormalities Predominate in the Bottom Part of the Sulcus with Type II Focal Cortical Dysplasia: A Quantitative Study

Z. Liu, W. Hu, Z. Sun, X. Wang, L. Liu, X. Shao, K. Zhang, Y. Ma, and J. Zhang

ABSTRACT

BACKGROUND AND PURPOSE: Type II focal cortical dysplasia is a common histopathological substrate in focal epilepsy. This study explored the spatial distribution of abnormal findings on MR imaging across the sulcus with type II focal cortical dysplasia using quantitative MR imaging postprocessing techniques.

MATERIALS AND METHODS: The morphometric analysis program and normalized FLAIR signal intensity analysis were applied to retrospectively analyze the MR imaging data of 58 patients with histopathologically confirmed type II focal cortical dysplasia. We divided the dysplastic sulcus into the bottom and nonbottom parts. Then spatial distribution types 1, 2, and 3 were arbitrarily defined as the abnormal findings on MR imaging (*z*-value > threshold) located in the bottom part, both the bottom and nonbottom parts, and the nonbottom part, respectively. For type 2, the mean *z*-values and standardized volumes of abnormal findings on MR imaging were compared between the bottom and nonbottom parts.

RESULTS: Abnormal findings on MR imaging were detected by quantitative techniques in 42 of 58 enrolled patients. Among these 42 patients, 38 and 26 patients showed gray-white matter junction blurring and cortical FLAIR hyperintensity, respectively, which were the 2 most common abnormal MR imaging features. Gray-white matter junction blurring manifested as types 1, 2, and 3 in 24, 13, and 1 patient, respectively, and the corresponding counts for cortical FLAIR hyperintensity were 12, 13, and 1 patient. For the 2 most common abnormal findings on MR imaging spatially manifested as type 2, higher mean *z*-values and larger corresponding standardized volumes of abnormalities were found in the bottom part.

CONCLUSIONS: Abnormal findings on MR imaging predominate in the bottom part of the sulcus with type II focal cortical dysplasia, which indicates that this malformation is bottom-of-sulcus-rooted.

ABBREVIATIONS: FCD = focal cortical dysplasia; MAP = morphometric analysis program; nFSI = normalized FLAIR signal intensity

Type II focal cortical dysplasia (FCD) is a highly epileptogenic lesion associated with pharmacoresistant epilepsy, and it is characterized by dysmorphic neurons and balloon cells.[1] FCD type II was first described by Taylor et al, in 1971,[2] and increasing reports have indicated that it is a common histopathologic substrate in epilepsy surgery.[3-6] Clinical characteristics, including severe partial epilepsy beginning in childhood, stereotyped seizures, high seizure frequency, and extratemporal location, have been associated with this specific malformation.[6] Electrophysiologic studies have also revealed that repetitive subcontinuous spikes, spikes and waves, polyspikes, or bursts of fast rhythms are reliable biomarkers of FCD type II.[7]

In addition to the above-mentioned electroclinical features, neuroimaging studies have demonstrated MR imaging characteristics of FCD type II, including gray-white matter junction blurring, increased cortical FLAIR signal, abnormal gyration/sulcation, increased cortical thickness, and the transmantle sign.[8-10] However, most previous studies were based on conventional vi-

Received August 14, 2018; accepted after revision October 30.

From the Departments of Neurosurgery (Z.L., X.W., K.Z., J.Z.) and Neurology (X.S.), Beijing Tiantan Hospital, Capital Medical University, Beijing, China; Beijing Neurosurgical Institute (W.H., J.Z.), Capital Medical University, Beijing, China; Department of Neurosurgery (Z.S.), Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing, China; and Department of Pathology (L.L.) and Neurosurgery (Y.M.), Beijing Fengtai Hospital, Beijing, China.

This work is partly supported by grants from the Capital (China) Health Research and Development of Special Fund (2016–1-1071), the Beijing Municipal Science & Technology Commission (Z161100000216130), the Beijing Municipal Administration of Hospitals' Ascent Plan (DFL20150503), and the Capital Medical University Basic and Clinical Cooperative Research Project (17JL05).

Zhifeng Liu and Wenhan Hu contributed equally to this work.

Please address correspondence to Jianguo Zhang, MD, PhD, Beijing Neurosurgical Institute, Capital Medical University, Tiantan xili 6, Beijing, 100050, China; e-mail: zjguo73@126.com

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

http://dx.doi.org/10.3174/ajnr.A5919

sual analysis, which strongly depends on the training and experience of the interpreting rater and has limited ability to recognize the existence and delineate the extent of FCD lesions. With the development of neuroimaging and computer technology, a large number of image-postprocessing methods have been developed to detect the abnormal findings on MR imaging of FCD quantitatively.[11-15] In a morphometric analysis program (MAP) first introduced by Huppertz et al,[14] 3 feature images (ie, junction, extension, and thickness images) are derived from a statistical parametric mapping–based algorithm, which was designed to measure the severity of gray-white matter junction blurring, abnormal gyration/sulcation, and cortical thickening, respectively.[14,16] Normalized FLAIR signal intensity (nFSI) analysis compares individual MR imaging data with a healthy control data base to detect changes in FLAIR signal intensity.[12] These 2 quantitative methods have high degrees of sensitivity in FCD detection compared with conventional visual inspection.[12,13,17]

Besson et al[18] used techniques of automated sulcal extraction and morphometry to explore the spatial relationship between FCD lesions and brain sulci, and they concluded that small FCDs were located at the bottom of a deep sulcus. However, only small lesions were investigated; large FCD lesions with abnormal findings on MR imaging continuously extending to the wall or crown of the sulci remained unexplored in their series. The present study aimed to describe the patterns of the distribution of abnormal findings on MR imaging across the type II dysplastic sulci using quantitative techniques. We believe our work can provide further helpful information to better understand the mechanism of FCD pathogenesis as well as principles of intracranial electrode implantation and resective surgery planning.

MATERIALS AND METHODS

Patients

We retrospectively included patients who underwent a resective epilepsy operation in Beijing Tiantan-Fengtai Epilepsy Center from January 2015 to December 2017 using the following criteria: 1) The histopathologic finding was type II FCD, according to the 2011 International League Against Epilepsy FCD classification system. FCD type II was defined as abnormal cortical lamination with dysmorphic neurons alone (IIa) or together with balloon cells (IIb).[1] 2) Presurgical 3D T1 and FLAIR images were available. Patients with low-quality MR images due to noise or movement artifacts were excluded. 3) The patient was older than 5 years of age,[15,19] and 4) no cranial surgery had been performed previously.

This research was approved by the institutional review board of the Beijing Tiantan Hospital (approval number KY2017–043-02), and informed consent was obtained from all included participants.

MR Imaging Acquisition

The MR images were obtained on a 3T Verio scanner (Siemens, Erlangen, Germany), including 3D T1 sagittal MPRAGE (TR/TE = 1900/2.53 ms, TI = 900 ms, matrix = 256 × 256, thickness = 1.0 mm), T2 axial (TR/TE = 7030/110 ms, matrix = 256 × 320, thickness = 3 mm), FLAIR axial (TR/TE = 8000/94 ms, TI = 2371.5 ms, matrix = 424 × 512, thickness = 3 mm), FLAIR sagittal (TR/TE = 8000/96 ms, TI = 2371.2 ms, matrix = 236 × 256, thickness = 3 mm), and FLAIR coronal (TR/TE = 8000/96 ms, TI = 2371.2 ms, matrix = 408 × 512, thickness = 3 mm) sequences.

Visual Inspection

Two epileptologists (X.S., W.H.) performed conventional visual analysis of the MR images independently, and discrepancies were resolved through discussion. The 2 reviewers defined positive findings on MR imaging when FCD-associated abnormal findings were identified during this procedure.

MR Imaging Postprocessing and Dysplastic Sulci Labeling

Two SPM8 (http://www.fil.ion.ucl.ac.uk/spm/software/spm8) postprocessing algorithms, MAP and nFSI, were used to quantify the abnormal findings on MR imaging. Extension, junction, and thickness images composed of *z*-values were derived from MAP by comparing individual MR imaging data with a healthy control data base. These images are used to measure the severity of abnormal gyration/sulcation, gray-white matter junction blurring, and cortical thickening, respectively.[16] Gray-white matter junction blurring may lead to inaccurate segmentation of gray or white matter probability images, which can result in incorrect calculation of cortical thickness. To avoid the cortical pseudothickening induced by gray-white matter junction blurring, we adjusted the gray matter probability image segmented for cortical thickness calculation by subtracting the binary image (derived from the 3D T1 image in the workflow of junction image calculation). FLAIR intensity changes in gray and white (transmantle sign) matter were quantitatively measured by nFSI, resulting in a *z*-value-based image. An epileptologist (K.Z.) blinded to the conventional visual results used *z* score thresholds of 6 (extension image), 3.5 (junction image), 4 (thickness image), or 3 (nFSI image) to identify candidate postprocessing positive (postprocessing+) regions.[12,15,19,20] The coregistered structural MR imaging and postsurgical MR imaging or CT were visually inspected at the candidate postprocessing+ region. The patient was classified as postprocessing+ if the candidate region was considered abnormal on the structural scan and was removed according to the postsurgical MR imaging or CT. The details of image-processing steps and threshold selection were described in previous articles.[12,15-17,19,20] The dysplastic sulcus with abnormal findings on MR imaging was manually labeled in the 3D T1 MR image according to the postprocessing results, and the corresponding ROI was drawn using VOI tools of MRIcron (http://people.cas.sc.edu/rorden/mricron/index.html) (Fig 1).

Division of the Dysplastic Sulcus

There is no previous literature describing the extent of the bottom part of a sulcus. We first delineated this extent according to our knowledge. On the basis of the structural MR image and corresponding ROI image, a reviewer (Z.L.) blinded to the postprocessing results drew a line intersecting the bottom point of the sulcus and parallel to the tangent to the interior border of the gray matter. This line divided the ROI of the dysplastic sulcus into the bottom and nonbottom parts (Fig 2*A*).

Patterns of Distribution of Abnormal Findings on MR Imaging

We defined spatial distribution types 1, 2, and 3 as abnormal findings on MR imaging (*z*-value > corresponding threshold)

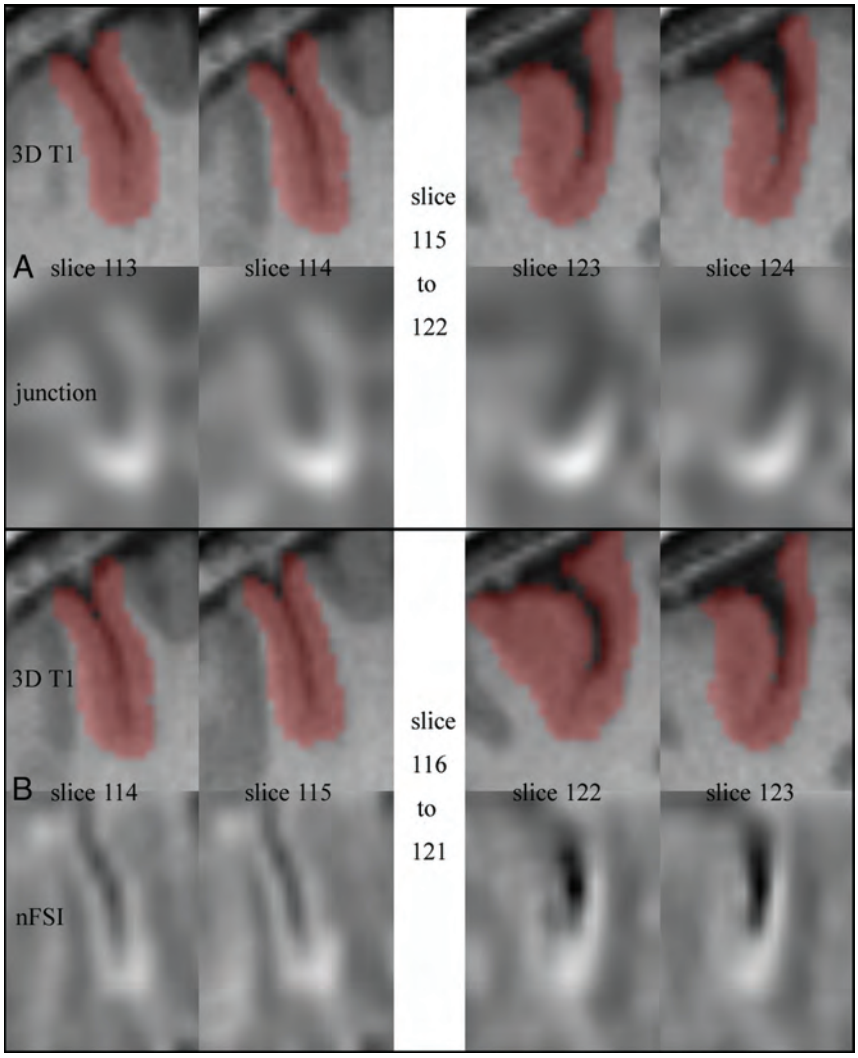


FIG 1. Dysplastic sulcus labeling. *A*, Dysplastic sulcus with an abnormality of gray-white matter junction blurring (coronal). *B*, Dysplastic sulcus with an abnormality of cortical FLAIR hyperintensity (coronal).

located in the bottom part, both the bottom and nonbottom parts, and the nonbottom part, respectively (Fig 2*B–D*).

Comparisons between the Bottom and Nonbottom Parts

For patients with type 2 distribution, we further compared the severities of abnormal findings on MR imaging between the bottom and nonbottom parts. The corresponding postprocessing images were multiplied by the mask images of the bottom or nonbottom part to obtain the statistical images of the 2 parts. The mean *z*-value in the bottom or nonbottom part of each statistical image was calculated. Moreover, the number of voxels with a *z*-value more than the previously defined threshold in the bottom or nonbottom part was divided by the total number of voxels of the corresponding part to obtain the standardized volume of the abnormality. All calculations during this step were performed in the platform of Matlab R2013a (MathWorks, Natick, Massachusetts). The mean *z*-values and standardized volumes of abnormalities between the bottom and nonbottom parts were compared in patients with type 2 distribution.

Postoperative Outcome

The Engel Scale was used to rate postoperative outcome. Seizure freedom was strictly defined as Engel Ia (seizure-free, no auras).[21]

Statistical Analyses

The Fisher exact test (2-tailed) was performed for statistical comparisons of independent categoric data. The independent *t* test (2-tailed) was performed for comparisons of independent numeric data if the variables were normally distributed; otherwise, the Mann-Whitney *U* test was used. Significance was defined as $P \leq .05$. The statistical tests were performed with SPSS 22.0 (IBM, Armonk, New York).

RESULTS

General Information

Between January 2015 and December 2017, four hundred seventy-one patients underwent resective surgery for refractory focal epilepsy at Beijing Tiantan-Fengtai Epilepsy Center. Among them, 72 patients had the histopathologic diagnosis of FCD II. One, 9, and 4 patients were excluded due to obvious artifacts on MR images, age younger than 5 years, and previous cranial operations, respectively. Fifty-eight patients (23 female and 35 male) were ultimately included in this study. The mean age at the operation was 18.78 ± 9.59 years, and the mean epilepsy duration was 7.89 ± 5.37 years. Thirty-seven patients underwent stereoelectroencephalography monitoring before resective surgery. The FCD lesions were located in the frontal lobe in 38 (65.52%) patients, parietal lobe in 8 (13.79%) patients, temporal lobe in 4 (6.90%) patients, occipital lobe in 4 (6.90%) patients, insular lobe in 3 (5.17%) patients, and multiple lobes in 1 (1.72%) patient. At a mean postoperative follow-up of 23.81 ± 9.60 months, 50 patients (86.2%) became seizure-free (Engel Ia).

Detection Rates of Conventional Visual Analysis and Image Postprocessing Assistant Analysis

In general, 27 (46.55%) FCD lesions were detected by visual inspection. In addition to the 27 lesions, another 15 (25.86%) subtle FCDs were detected under the assistance of image postprocessing techniques, leading to the detection rate of 72.41%. The findings of remaining 16 (27.59%) FCDs were strictly negative on MR imaging and were detected by interictal PET with [^{18}F] FDG or stereoelectroencephalography (Fig 3).

Frequencies of Abnormal Findings on MR Imaging

As shown in the Table, among the 42 FCD lesions with detectable abnormal findings on MR imaging, gray-white matter junction

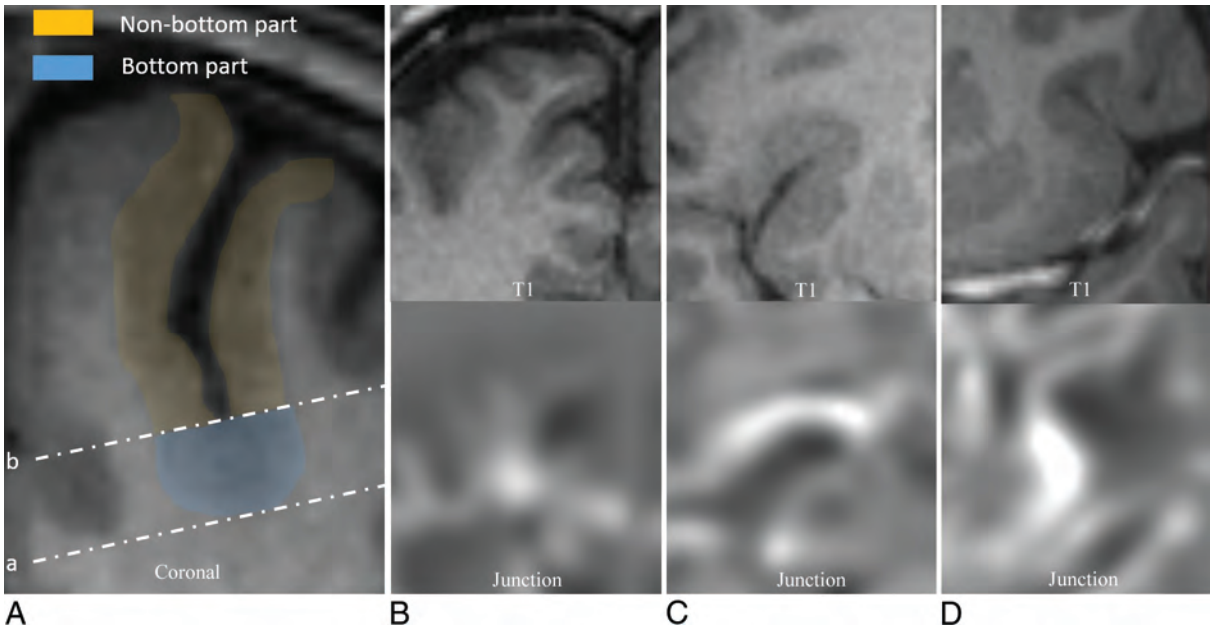


FIG 2. Schematic diagram of the divisions of the sulcus and the types of spatial distribution of abnormal findings on MR imaging. *A*, *Line a* is tangential to the interior border of the gray matter in the sulcus bottom, and *line b* is parallel to *line a*, intersecting the bottom point of the sulcus. Gray matter interior and exterior to *line b* is defined as the bottom and the nonbottom parts, respectively. If one considered gray-white matter junction blurring, for example, spatial distribution types 1 (*B*), 2 (*C*), and 3 (*D*) are defined as the abnormalities located in the bottom part, both the bottom and nonbottom parts, and the nonbottom part, respectively.

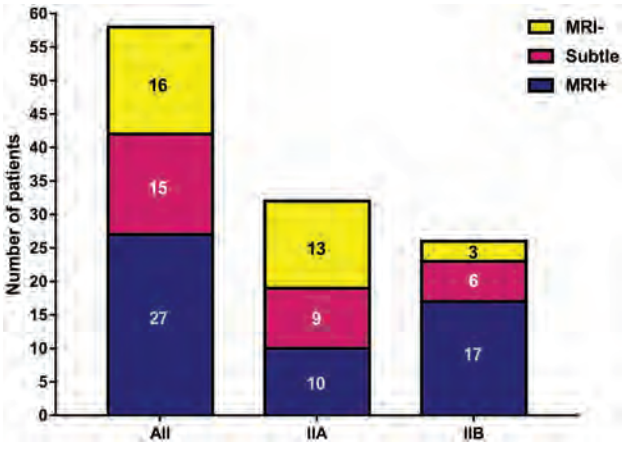


FIG 3. Abnormal findings on MR imaging detected by conventional visual analysis and postprocessing techniques. MRI+ indicates positive by conventional visual analysis; Subtle, negative by conventional visual analysis but positive by postprocessing; MRI−, negative by postprocessing.

blurring was the most common ($n = 38$), followed by cortical FLAIR hyperintensity ($n = 26$), the transmantle sign ($n = 13$), abnormal gyration or sulcation ($n = 7$), and cortical thickening ($n = 4$). Type IIb lesions significantly more often showed detectable abnormalities ($P = .015$), including gray-white matter junction blurring ($P = .029$) and cortical FLAIR hyperintensity ($P = .001$), compared with type IIa lesions; the transmantle sign was an exclusive feature of FCD IIb ($P < .001$).

Patterns of Distribution of Abnormal Findings on MR Imaging

The abnormalities of gray-white matter junction blurring located in the bottom part (type 1) of dysplastic sulci in 24 patients (IIa, $n = 10$; IIb, $n = 14$), consecutively extended to the wall or crown (type 2) in 13 patients (IIa, $n = 6$; IIb, $n = 7$) and were located in the nonbottom part (type 3) in 1 patient with IIa. For the abnormality of cortical FLAIR hyperintensity, 12 (IIa, $n = 2$; IIb, $n = 10$), 13 (IIa, $n = 6$; IIb, $n = 7$), and 1 (IIb, $n = 1$) lesion showed type 1, 2, and 3 spatial distribution, respectively (Fig 4). All abnormalities of the transmantle sign and abnormal gyration/sulcation showed a type 1 distribution. Regarding cortical thickening, 2 lesions showed a type 2 distribution and the other 2 showed a type 3 distribution.

Comparisons between the Bottom and Nonbottom Parts

Because all abnormalities of the transmantle sign and abnormal gyration/sulcation were located in the bottom part and cases with cortical thickening were rare in our cohort, only *z*-values of the junction and nFSI images were taken into account for comparisons. In patients with gray-white matter junction blurring consecutively extending from the bottom to the wall or crown (type 2 distribution), the mean *z*-value in the bottom part was higher than that in the nonbottom part (all: 3.91 ± 1.21 versus 0.43 ± 0.60, $P < .001$; subgroup of FCD IIa: 4.37 ± 1.41 versus 0.41 ± 0.41, $P < .001$; subgroup of FCD IIb: 3.51 ± 0.94 versus 0.44 ± 0.77, $P < .001$), and the standardized volume of abnormalities was also larger in the bottom part (all: 0.56 ± 0.19 versus 0.09 ± 0.06, $P < .001$; subgroup of FCD IIa: 0.60 ± 0.18 versus 0.09 ± 0.04, $P < .001$; subgroup of FCD IIb: 0.52 ± 0.20 versus 0.08 ± 0.08, $P < .001$). Regarding FLAIR signal changes, the mean *z*-value was higher (all: 2.52 ± 0.63 versus 0.92 ± 0.50, $P < .001$; subgroup of FCD IIa: 2.39 ± 0.87 versus 0.80 ± 0.24, $P < .001$; subgroup of FCD IIb: 2.63 ± 0.37 versus 1.03 ± 0.65, $P < .001$) and the standardized volume of abnormalities was larger (all: 0.34 ± 0.18 versus 0.06 ± 0.05, $P < .001$; subgroup of FCD IIa: 0.28 ± 0.22 versus 0.04 ± 0.04, $P < .001$; subgroup of FCD IIb:

Frequencies of abnormal findings on MRI measured by quantitative methods

	All (n = 58)	IIa (n = 32)	IIb (n = 26)	P (IIa vs IIb)
Gray-white matter junction blurring	38 (65.52%)	17 (53.13%)	21 (80.77%)	.029
Cortical FLAIR hyperintensity	26 (44.83%)	8 (25%)	18 (69.23%)	.001
Cortical thickening	4 (6.90%)	2 (6.25%)	2 (7.69%)	>.999
Abnormal gyrus or sulcus	7 (12.07%)	5 (15.63%)	2 (7.69%)	.614
Transmantle sign	13 (22.41%)	0 (0%)	13 (100%)	<.001
Normal findings	16 (27.59%)	13 (40.63%)	3 (11.54%)	.015

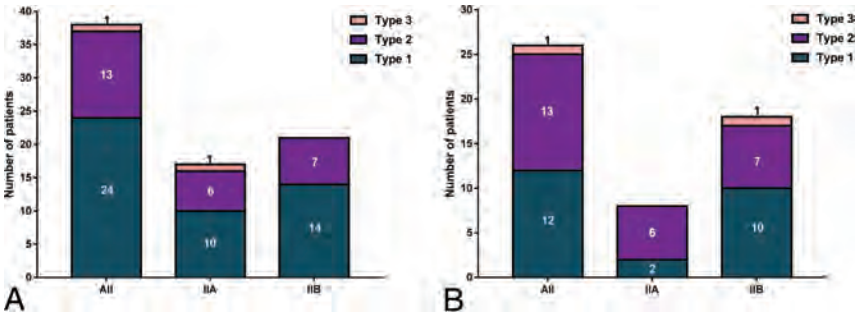


FIG 4. Statistical data of the 3 spatial distribution types of abnormalities on MR imaging. *A*, Gray-white matter junction blurring. *B*, Cortical FLAIR hyperintensity.

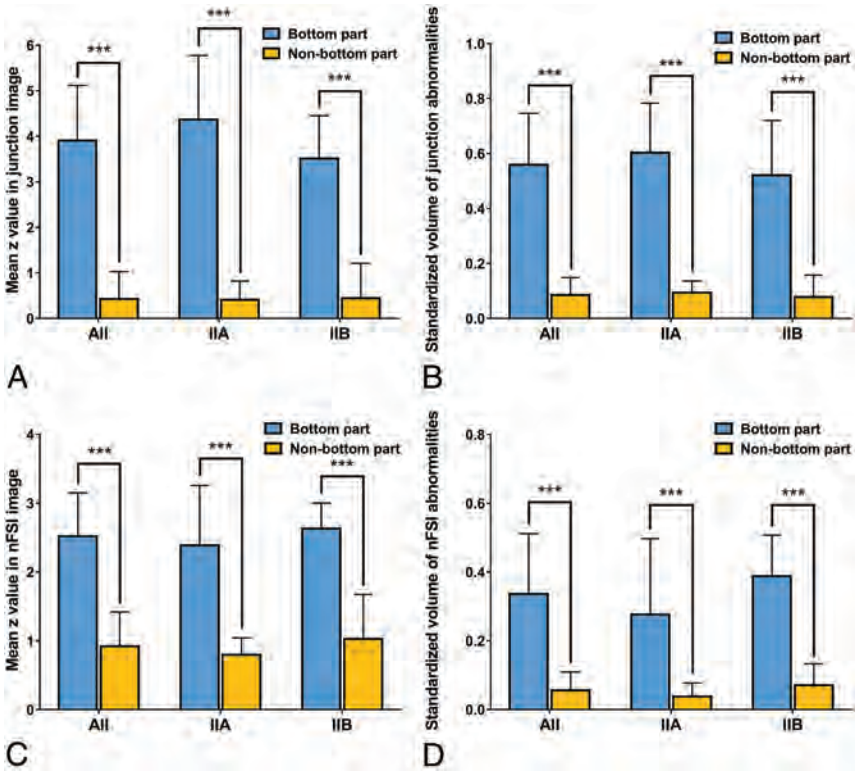


FIG 5. Comparisons between the bottom and nonbottom parts in patients with a type 2 abnormality distribution. *A*, Comparison of *z*-values derived from the junction image (designed to measure gray-white matter junction blurring) of MAP. *B*, Comparison of standardized volumes of abnormalities ($z > 3.5$) in the junction image. *C*, Comparison of *z*-values derived from the nFSI image (designed to measure cortical FLAIR hyperintensity). *D*, Comparison of standardized volumes of abnormalities ($z > 3$) in the nFSI image.

0.39 ± 0.12 versus 0.07 ± 0.06, $P < .001$) in the bottom part compared with that in the nonbottom part (Fig 5).

DISCUSSION

Knowledge of the MR imaging features is crucial for recognition and detection of FCD II lesions, and understanding the spatial distribution of abnormal findings on MR imaging is important for deep electrode implantation and lesion resection. The present retrospective study quantitatively revealed that gray-white matter junction blurring and cortical FLAIR hyperintensity were the 2 most common abnormal findings on MR imaging of type II FCD, and the transmantle sign was exclusive to FCD IIb. Moreover, abnormal findings on MR imaging predominated in the bottom part of the sulcus with type II FCD.

Bottom-of-sulcus dysplasia was first proposed by the 2005 revision of the Barkovich classification. It was defined as a subtype of FCD with balloon cells (FCD IIb).[22] FCD IIa was excluded because it had different imaging features, different outcomes after the seizure operation, and clearly different etiologies.[4,5,22-24] In 2008, a study by Besson et al[18] indicated that small FCD lesions were located at the bottom of a deep sulcus. The present study demonstrates that in addition to FCD IIb, the abnormal findings on MR imaging ($n = 10$ for gray-white matter junction blurring and $n = 2$ for cortical FLAIR hyperintensity) of FCD IIa lesions could also be located in the bottom part of the sulcus. From the perspective of spatial location, these FCD IIa lesions should not be excluded from bottom-of-sulcus dysplasia. The exclusion of FCD IIa in the classification by Barkovich is most likely explained by the use of low-field-strength or low-resolution MR images. Small FCD IIa lesions located in the bottom part of the sulcus, frequently manifesting as subtle gray-white matter junction blurring and lacking notable characteristics such as the transmantle sign, tend to be overlooked during conventional visual analysis. As a consequence of negative findings on MR imaging, it is difficult to localize those lesions and the seizure outcomes are poor.[25,26]

In addition to subtle lesions located at the bottom, other larger lesions with abnormalities continuously extending from the bottom to the wall were also analyzed in our study. Among these lesions, quantitative analyses indicated that abnormal findings on MR imaging predominated in the bottom part of the sulcus. Based on the above-mentioned findings, we postulated that the bottom of the sulcus was the "root" of type II FCD lesions and proposed the concept of

bottom-of-sulcus rooted dysplasia. The lesions with type 1 and 2 spatial distributions were all bottom-of-sulcus-rooted dysplasia. The only difference between them we could see was the volume of the lesions.

The mechanism of cortical folding is unclear. Two hypotheses have dominated views on this topic. The first is the intracortical differential growth hypothesis, whereby the differential expansion of upper-versus-lower neuronal layers causes the surface to buckle or fold.[27] The cortex is thicker at sites of growing gyri and thinner beneath developing sulci.[28] The cortex forms its 6-layered structure until gestational week 18, before cortical folding. We postulate that the sites where FCD lesions are located, being the weak points of tangential tension due to the disruption of cortical layers, tend to be folded into the sulci. The second is the axonal tension hypothesis, which posits that tension generated by axons drives cortical folding by pulling strongly interconnected regions together.[29] According to this theory, gyri are bulked between areas with strong axonal connections as axons draw them together, and sulci are folded between areas weakly connected or even without links. Decreased subcortical fiber connectivity in FCD has been demonstrated,[30-32] which may be the main reason that this kind of lesion is bottom-of-sulcus-rooted.

Our study also provided a rule for interpreting the results of postprocessing images in the detection of FCD lesions. The sample cases of FCD II presented in the literature of image postprocessing showed that most dysplastic areas highlighted by the extension, junction, and nFSI images were located in the inferior parts of the dysplastic sulci.[12-14] The results of the present study are consistent with the findings from previous publications. Huppertz,[16] the introducer of MAP, claimed that this method does not detect the lesion automatically, so a visual confirmation at the MAP-highlighted area on coregistered structural MR imaging was still necessary. The interpretation of the postprocessing images requires some experience. Moreover, false-positives exist because the postprocessing images may highlight areas that have no pathologic correlates in the structural MR images.[16] If the electroclinical data indicate an FCD II lesion, areas at the crown of the gyri highlighted by any postprocessing images need extremely careful inspection because this kind of lesion is bottom-of-sulcus-rooted.

Subdural and depth electrodes are widely used in localizing epileptogenic zones. Although subdural electrodes have advantages, including large cortical coverage and convenience in cortical mapping,[33] contacts of subdural electrodes, which are placed on the surface of the brain, cannot directly record the signals of FCD deeply buried in the sulcus. Thus, depth electrodes are preferred in localizing FCD II during the phase 2 evaluation. To go close to the FCD lesion rooted at the bottom of the sulcus, the trajectory of the depth electrode should run through the inferior portion of the dysplastic sulcus. For lesion resection, the depth of resection of the sulcus directly affects the postsurgical seizure outcome. In other words, the gray matter at the bottom should be completely removed to reveal the underlying white matter, especially for the sulcus, where vessels pass at the bottom.

The main limitation of the present study is the exclusion of pediatric patients younger than 5 years of age due to the absence of healthy controls with a similar age distribution. Younger brains show significant structural differences compared with adult brains, and bias might have occurred during the unequal comparisons. However, the prevalence rate of FCD II in these pediatric patients undergoing an epilepsy operation was high. Another limitation is that some cases with electroclinical or neuroimaging data strongly indicating FCD II were not included due to negative histopathologic findings. The main cause, we postulate, is the fragmentation of specimens during the resective surgery. Selection bias might have been induced by the above-mentioned limitations in our study.

CONCLUSIONS

We explored the spatial distribution of abnormal findings on MR imaging across the sulci containing FCD II lesions using quantitative techniques. Our results suggest that gray-white matter junction blurring and cortical FLAIR hyperintensity are the 2 most common characteristics, which predominate in the bottom part of the sulcus with type II FCD. The findings of our study indicate that FCD II is bottom-of-sulcus-rooted and provide helpful information for postprocessing image interpretation, intracranial electrode implantation, and resection of this kind of lesion.

Disclosures: Kai Zhang—*RELATED*: *Grant*: Capital Medical University Beijing, *Comments*: Capital Medical University Basic and Clinical Cooperative Research Project (17JL05).* Jianguo Zhang—*RELATED*: *Grant*: Beijing Municipal Commission of Health and Family Planning, Beijing Municipal Science & Technology Commission, Beijing Municipal Administration of Hospitals, *Comments*: Capital (China) Health Research and Development of Special Fund (2016-1-1071), the Beijing Municipal Science & Technology Commission (Z161100000216130), Beijing Municipal Administration of Hospitals' Ascent Plan (DFL20150503)*; *Consulting Fee or Honorarium*: Beijing Municipal Administration of Hospitals, *Comments*: Beijing Municipal Administration of Hospitals' Ascent Plan (DFL20150503)*; *Support for Travel to Meetings for the Study or Other Purposes*: Beijing Municipal Commission of Health and Family Planning, *Comments*: Capital (China) Health Research and Development of Special Fund (2016-1-1071)*; *Fees for Participation in Review Activities such as Data Monitoring Boards, Statistical Analysis, Endpoint Committees, and the Like*: Beijing Municipal Administration of Hospitals, *Comments*: Beijing Municipal Administration of Hospitals' Ascent Plan (DFL20150503).* *Money paid to the institution.

REFERENCES

1. Blümcke I, Thom M, Aronica E, et al. **The clinicopathologic spectrum of focal cortical dysplasias: a consensus classification proposed by an ad hoc Task Force of the ILAE Diagnostic Methods Commission.** *Epilepsia* 2011;52:158–74 CrossRef Medline
2. Taylor DC, Falconer MA, Bruton CJ, et al. **Focal dysplasia of the cerebral cortex in epilepsy.** *J Neurol Neurosurg Psychiatry* 1971;34:369–87 CrossRef Medline
3. Lerner JT, Salamon N, Hauptman JS, et al. **Assessment and surgical outcomes for mild type I and severe type II cortical dysplasia: a critical review and the UCLA experience.** *Epilepsia* 2009;50:1310–35 CrossRef Medline
4. Tassi L, Colombo N, Garbelli R, et al. **Focal cortical dysplasia: neuropathological subtypes, EEG, neuroimaging and surgical outcome.** *Brain* 2002;125:1719–32 CrossRef Medline
5. Lawson JA, Birchansky S, Pacheco E, et al. **Distinct clinicopathologic subtypes of cortical dysplasia of Taylor.** *Neurology* 2005;64:55–61 CrossRef Medline
6. Chassoux F, Landré E, Mellerio C, et al. **Type II focal cortical dysplasia: electroclinical phenotype and surgical outcome related to imaging.** *Epilepsia* 2012;53:349–58 CrossRef Medline
7. Guerrini R, Duchowny M, Jayakar P, et al. **Diagnostic methods and treatment options for focal cortical dysplasia.** *Epilepsia* 2015;56:1669–86 CrossRef Medline
8. Colombo N, Tassi L, Deleo F, et al. **Focal cortical dysplasia type IIa**

and IIb: MRI aspects in 118 cases proven by histopathology. *Neuroradiology* 2012;54:1065–77 CrossRef Medline
9. Krsek P, Maton B, Korman B, et al. **Different features of histopathological subtypes of pediatric focal cortical dysplasia.** *Ann Neurol* 2008;63:758–69 CrossRef Medline
10. Krsek P, Pieper T, Karlmeier A, et al. **Different presurgical characteristics and seizure outcomes in children with focal cortical dysplasia type I or II.** *Epilepsia* 2009;50:125–37 CrossRef Medline
11. Bernasconi A, Antel SB, Collins DL, et al. **Texture analysis and morphological processing of magnetic resonance imaging assist detection of focal cortical dysplasia in extra-temporal partial epilepsy.** *Ann Neurol* 2001;49:770–75 CrossRef Medline
12. Focke NK, Symms MR, Burdett JL, et al. **Voxel-based analysis of whole brain FLAIR at 3T detects focal cortical dysplasia.** *Epilepsia* 2008;49:786–93 CrossRef Medline
13. Wagner J, Weber B, Urbach H, et al. **Morphometric MRI analysis improves detection of focal cortical dysplasia type II.** *Brain* 2011; 134:2844–54 CrossRef Medline
14. Huppertz HJ, Grimm C, Fauser S, et al. **Enhanced visualization of blurred gray-white matter junctions in focal cortical dysplasia by voxel-based 3D MRI analysis.** *Epilepsy Res* 2005;67:35–50 CrossRef Medline
15. Wang ZI, Jones SE, Jaisani Z, et al. **Voxel-based morphometric magnetic resonance imaging (MRI) postprocessing in MRI-negative epilepsies.** *Ann Neurol* 2015;77:1060–75 CrossRef Medline
16. Huppertz HJ. **Morphometric MRI analysis.** In: Urbach H, ed. *MRI in Epilepsy*. Berlin: Springer-Verlag; 2013:73–84
17. Hu WH, Wang X, Liu LN, et al. **Multimodality image post-processing in detection of extratemporal MRI-negative cortical dysplasia.** *Front Neurol* 2018;9:450 CrossRef Medline
18. Besson P, Andermann F, Dubeau F, et al. **Small focal cortical dysplasia lesions are located at the bottom of a deep sulcus.** *Brain* 2008;131: 3246–55 CrossRef Medline
19. Wang ZI, Alexopoulos AV, Jones SE, et al. **Linking MRI postprocessing with magnetic source imaging in MRI-negative epilepsy.** *Ann Neurol* 2014;75:759–70 CrossRef Medline
20. Wellmer J, Parpaley Y, von Lehe M, et al. **Integrating magnetic resonance imaging postprocessing results into neuronavigation for electrode implantation and resection of subtle focal cortical dysplasia in previously cryptogenic epilepsy.** *Neurosurgery* 2010;66:187–94; discussion 194–95 CrossRef Medline
21. Engel JJ, Van Ness PC, Rasmussen TB, et al. **Outcome with respect to epileptic seizures.** In: Engel JJ, ed. *Surgical Treatment of the Epilepsies.* 2nd ed. New York: Raven Press; 1993:609–21
22. Barkovich AJ, Kuzniecky RI, Jackson GD, et al. **A developmental and genetic classification for malformations of cortical development.** *Neurology* 2005;65:1873–87 CrossRef Medline
23. Marusic P, Najm IM, Ying Z, et al. **Focal cortical dysplasias in eloquent cortex: functional characteristics and correlation with MRI and histopathologic changes.** *Epilepsia* 2002;43:27–32 Medline
24. Urbach H, Scheffler B, Heinrichsmeier T, et al. **Focal cortical dysplasia of Taylor's balloon cell type: a clinicopathological entity with characteristic neuroimaging and histopathological features, and favorable postsurgical outcome.** *Epilepsia* 2002;43:33–40 Medline
25. Noe K, Sulc V, Wong-Kisiel L, et al. **Long-term outcomes after nonlesional extratemporal lobe epilepsy surgery.** *JAMA Neurol* 2013;70: 1003–08 CrossRef Medline
26. Téllez-Zenteno JF, Hernández 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
27. Richman DP, Stewart RM, Hutchinson JW, et al. **Mechanical model of brain convolutional development.** *Science* 1975;189:18–21 CrossRef Medline
28. Budday S, Steinmann P, Kuhl E. **Physical biology of human brain development.** *Front Cell Neurosci* 2015;9:257 CrossRef Medline
29. Van Essen DC. **A tension-based theory of morphogenesis and compact wiring in the central nervous system.** *Nature* 1997;385:313–18 CrossRef Medline
30. Diehl B, Tkach J, Piao Z, et al. **Diffusion tensor imaging in patients with focal epilepsy due to cortical dysplasia in the temporo-occipital region: electro-clinico-pathological correlations.** *Epilepsy Res* 2010;90:178–87 CrossRef Medline
31. Lee SK, Kim DI, Mori S, et al. **Diffusion tensor MRI visualizes decreased subcortical fiber connectivity in focal cortical dysplasia.** *Neuroimage* 2004;22:1826–29 CrossRef Medline
32. Widjaja E, Zarei Mahmoodabadi S, Otsubo H, et al. **Subcortical alterations in tissue microstructure adjacent to focal cortical dysplasia: detection at diffusion-tensor MR imaging by using magnetoencephalographic dipole cluster localization.** *Radiology* 2009; 251:206–15 CrossRef Medline
33. Najm IM, Bingaman WE, Luders HO. **The use of subdural grids in the management of focal malformations due to abnormal cortical development.** *Neurosurg Clin N Am* 2002;13:87–92, viii-ix Medline

Reliability of MR Imaging–Based Posterior Fossa and Brain Stem Measurements in Open Spinal Dysraphism in the Era of Fetal Surgery

M. Aertsen, J. Verduyckt, F. De Keyzer, T. Vercauteren, F. Van Calenbergh, L. De Catte, S. Dymarkowski, P. Demaerel, and J. Deprest

ABSTRACT

BACKGROUND AND PURPOSE: Fetal MR imaging is part of the comprehensive prenatal assessment of fetuses with open spinal dysraphism. We aimed to assess the reliability of brain stem and posterior fossa measurements; use the reliable measurements to characterize fetuses with open spinal dysraphism versus what can be observed in healthy age-matched controls; and document changes in those within 1 week after prenatal repair.

MATERIALS AND METHODS: Retrospective evaluation of 349 MR imaging examinations took place, including 274 in controls and 52 in fetuses with open spinal dysraphism, of whom 23 underwent prenatal repair and had additional early postoperative MR images. We evaluated measurements of the brain stem and the posterior fossa and the ventricular width in all populations for their reliability and differences between the groups.

RESULTS: The transverse cerebellar diameter, cerebellar herniation level, clivus-supraocciput angle, transverse diameter of the posterior fossa, posterior fossa area, and ventricular width showed an acceptable intra- and interobserver reliability (intraclass correlation coefficient > 0.5). In fetuses with open spinal dysraphism, these measurements were significantly different from those of healthy fetuses (all with $P < .0001$). Furthermore, they also changed significantly (P value range $= .01$ to $< .0001$) within 1 week after the fetal operation with an evolution toward normal, most evident for the clivus-supraocciput angle ($65.9 \pm 12.5°$; 76.6 ± 10.9; $P < .0001$) and cerebellar herniation level (-9.9 ± 4.2 mm; -0.7 ± 5.2; $P < .0001$).

CONCLUSIONS: In fetuses with open spinal dysraphism, brain stem measurements varied substantially between observers. However, measurements characterizing the posterior fossa could be reliably assessed and were significantly different from normal. Following a fetal operation, these deviations from normal values changed significantly within 1 week.

ABBREVIATIONS: ACi = atriocerebral index; CHL = cerebellar herniation level; CSA = clivus-supraocciput angle; GA = gestational age; ICC = intraclass correlation coefficient; OSD = open spinal dysraphism; PF = posterior fossa; TCD = transverse cerebellar diameter; TDPF = transverse diameter of the PF; VW = ventricular width

Open spinal dysraphism (OSD), subdivided into myelomeningocele and myeloschisis, is a nonlethal congenital malformation with complex physical and neurodevelopmental sequelae. Its prevalence is approximately 4.9 per 10,000 live births in Europe and 3.17 in the United States.[1-3] OSD results in motor and sensory deficits, their extension being defined by the upper level of the anatomic defect. These range, as the level increases, from bladder, bowel, and sexual dysfunction to involvement of the lower and even upper extremities and secondary orthopedic disabilities.[4,5] Children with OSD almost invariably have an associated Chiari II hindbrain malformation and ventriculomegaly.[6] The Chiari II malformation is characterized by posterior fossa (PF) and brain stem abnormalities with downward displacement and compression of the cerebellum and brain stem.[7]

Geerdink et al[8] demonstrated that morphometric measures reliably quantify the morphologic distortions of Chiari II malformation on postnatal MR images. The mamillopontine distance and the cerebellar width were the most sensitive and specific de-

Received May 29, 2018; accepted after revision October 6.

From the Department of Imaging and Pathology (M.A., J.V., F.D.K., S.D., P.D.), Clinical Department of Radiology, University Hospitals KU Leuven, Leuven, Belgium; School of Biomedical Engineering and Imaging Sciences (T.V.), King's College, London; Department of Neurosurgery (F.V.C.), University Hospitals Leuven, Leuven, Belgium; Academic Department of Development and Regeneration, Cluster Woman and Child (L.D.C., J.D.), Group Biomedical Sciences, KU Leuven, Leuven, Belgium; and Institute for Women's Health, University College London, (J.D.), London, UK.

Philippe Demaerel and Jan Deprest are shared last authors.

J. Deprest was partly funded by the Great Ormond Street Hospital Charity Fund. The work was also supported by the Engineering and Physical Sciences Research Council and the Innovative Engineering for Health award by the Wellcome Trust.

Please address correspondence to Michael Aertsen, MD, Department of Radiology, Herestraat 49, 3000 Leuven, Belgium; e-mail: michael.aertsen@uzleuven.be

Indicates article with supplemental on-line tables.

http://dx.doi.org/10.3174/ajnr.A5930

terminants of Chiari II.[9] Some fetuses with OSD have ventriculomegaly, and its degree is believed to be predictive of the need for postnatal shunting.[10,11]

In 2011, the Management of Myelomeningocele Study (MOMS) demonstrated the benefit of in utero repair of myelomeningocele because the need for ventricular shunting at 12 months was reduced and motor outcome at 30 months improved.[12] Fetuses with the suspicion of OSD should be assessed comprehensively to counsel parents about the expected outcome and possibility of fetal surgery. In this assessment, fetal MR imaging has a crucial role to characterize the brain and spinal abnormalities and rule out additional anomalies in fetuses with OSD.[13,14] For fetal surgery eligibility, the presence of Chiari II hindbrain malformation on MR imaging is a necessary finding.[12,15] Many measurements have been proposed to describe the typical PF changes in fetuses with OSD, yet the reproducibility of these has rarely been studied.[14,16-23] These parameters have also been shown to change after in utero repair of OSD in small series and at different time points after fetal surgery, yet no study has consistently reported early postoperative assessment in utero.[18]

The aims of this study were 3-fold: 1) to assess the reproducibility of measurements of the brain stem and PF that have been suggested to be representative on postnatal[8,9] and prenatal MR imaging[14,16-23]; 2) to apply those parameters that were shown to be reproducible, to discriminate fetuses with OSD from gestational age–matched fetuses with a normal PF; and 3) to document early changes in these measurements 1 week after a fetal operation.

MATERIALS AND METHODS

This was a single-center retrospective study at University Hospitals Leuven that was approved by its ethics committee (S60814). Patients eligible for the OSD group were those having fetal MR imaging examinations for additional assessment because of a prenatal diagnosis of OSD on ultrasound. Before MR imaging, patients had an ultrasound assessment, in which the lesion, secondary changes, and, when applicable, associated anomalies were characterized. Before the MR imaging, the radiologist was informed of the ultrasound findings. For the gestational age–matched controls, we included fetuses assessed for other congenital anomalies that do not affect the central nervous system or who were scanned for suspected CNS abnormalities with normal findings on prenatal ultrasound, fetal MR imaging, and postnatal evaluation (On-line Table 1).

Fetal Imaging and Quality Criteria

The routine protocol for this condition includes acquisition on a 1.5T MR imaging system (Aera; Siemens, Erlangen, Germany) with 2 small body coils placed adjacent to each other over the maternal abdomen. The mother was positioned in the supine or left lateral decubitus position. The images used were T2-weighted HASTE or balanced steady-state gradient-echo sequences in the sagittal, axial, and coronal planes relative to the fetal head. Before September 2015, maternal sedation (flunitrazepam, 0.5 mg orally 20–30 minutes before the examination) was used when the gestational age (GA) was <30 weeks.[24] Later, this was abandoned because we, like others, thought that this induced maternal adverse effects while not clinically required.[25] For this study, we searched our data base for all examinations performed in the setting of spinal dysraphism assessment, as well as for appropriate gestational age–matched controls. The image quality had to be good, consisting of at least 3 orthogonal T2-weighted HASTE series of the fetal brain with limited fetal motion, allowing adequate performance of the outcome measurements. The primary selection and review of images was performed by a single pediatric radiologist (M.A.) with >3 years of experience in fetal MR imaging. The main exclusion criteria were twin pregnancy, syndromal pathology, fetal hydrops, or anhydramnios. The number of patients and individual reasons for exclusion are shown in On-line Table 2. This exclusion left data from 349 MR imaging examinations of a total of 1006, including 274 examinations in 246 control fetuses. These data illustrate that some fetuses were scanned more than once. Additionally, we included 52 MR imaging examinations in fetuses with OSD, of whom 23 had a repeat MR imaging examination after the operation. The eligibility criteria for fetal surgery were those used in the MOMS trial.[12]

Outcome Measurements

Biometric variables included the transverse cerebellar diameter (TCD), pontine thickness, and pontine height, measured according to the standards defined by Garel[26] and Tilea et al.[27] The transverse diameter of the PF (TDPF) was measured according to Woitek et al,[17] who suggested that this would be a proxy for the TCD. The midsagittal PF area was measured according to Tsai et al.[20] The ventricular width (VW) was measured in the coronal plane according to Garel,[26] and in case of asymmetry, the largest value was taken into account. Mamillopontine distance, the level of kinking of the brain stem, medullary length, tentorial length, and width of the cisterna magna were measured as described by Geerdink et al.[8] The width of the foramen magnum was defined as the distance between the opisthion and the basion. The cerebellar herniation level (CHL) was measured by drawing a perpendicular line from the foramen magnum to the lowest cerebellar portion. In the presence of cerebellar herniation, the deepest portion was measured.[20] The clivus-supraocciput angle (CSA) was measured according to D'Addario et al.[28] The TCD, TDPF, mamillopontine distance, TL, PF area, and CSA are demonstrated in Fig 1.

Reproducibility Study

The reproducibility of measurements was determined on a randomly chosen subgroup of spinal dysraphism cases ($n = 15/52$; referred to as the pilot group). Images were anonymized and uploaded to a research server[29] for assessment by M.A. and J.V., a radiologist with 1-year specific training for fetal MR imaging. This radiologist was first trained with a training dataset from 5 other fetuses with spinal dysraphism, with the help of a purposely designed training document. For intraobserver assessment, M.A. read the images twice in a random order with a 2-week interval. J.V. measured parameters with at least moderate (intraclass correlation coefficient [ICC] > 0.5) intraobserver reliability once to obtain interobserver reliability.[30]

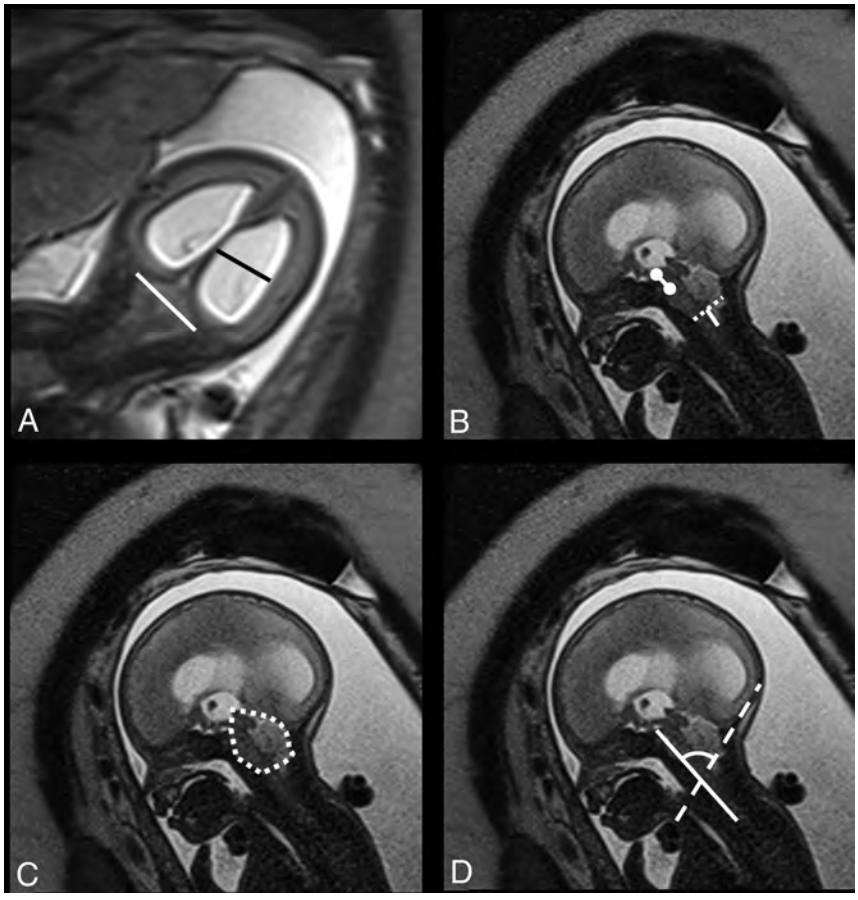


FIG 1. T2-weighted imaging (HASTE) in the coronal (*A*) plane shows the biparietal diameter (*white line* in *A*) and the ventricular width (*black line* in *A*). In the sagittal plane (*B–D*), the mamillopontine distance (*white line with circles at both ends in B*) is demonstrated along with the foramen magnum diameter (*dotted white line* in *B*) and the cerebellar herniation level (*white line perpendicular to the dotted line* in *B*). The midsagittal posterior fossa area is shown in *C* (*dotted free form*), and the clivus-supraocciput angle is demonstrated in *D* with a line according to the clivus (*white line*) and the occiput (*dotted line*).

Posterior Fossa Characteristics

The PF characteristics were determined on the presumed healthy population to obtain normative values. To compare cases of fetal myelomeningocele with the healthy population, we used these normative curves to calculate expected values for the given gestational ages, and the fetal myelomeningocele values were then expressed as observed over the expected ratio.

Short-Term (<7 Days) Postoperative Changes

In this part of the study, we looked at the difference between PF measurements shown to be reproducible in the above part of the study in 23 fetuses who had a fetal operation at our center. These fetuses had preoperative MR imaging and were imaged again within 1 week after the operation. All measurements were performed on T2-weighted images in the coronal or sagittal plane of the fetal head. Again, the values were expressed as observed over the expected ratio to determine changes after prenatal treatment that were not attributable to normal growth.

In addition to the posterior fossa, we also evaluated the ventricular width in fetuses with OSD and the difference from the control population. To describe the differential effect of a fetal operation on the parenchyma and ventricles, one can measure changes in the so-called atriocerebral index (ACi). The ACi is the ratio of the atrial diameter and the cerebral (parenchymal) biparietal diameter.[24] Others called this index the ventricular width index as used in the postnatal literature.[18]

Statistics

Intraobserver and interobserver variability were analyzed with a 2-way random ICC using SPSS for Windows, Version 22.0 (Released 2004; IBM, Armonk, New York). An ICC cutoff value of 0.5 as the lowest acceptable was chosen, taking into account the guidelines for interpretation by Cicchetti[31] and allowing some variation in view of the limited spatial resolution for assessing such small structures so that borderline parameters can be fully investigated. For the interpretation of ICC values, we followed the guidelines of Koo and Li,[30] with ICC values <0.5 indicative of poor reliability; values between 0.5 and 0.75, moderate reliability; values between 0.75 and 0.9, good reliability; and values >0.90, excellent reliability. The normality of GA in the 274 examinations was evaluated using the Shapiro-Wilk test, which indicated that GA was not normally distributed. We attempted several transformation models (logarithmic, polynomial, and square root) from which the square root transformation provided the most normally distributed data. Afterward, regression analysis was performed on all examined PF characteristics to find normative ranges in correlation with GA. Differences in the reliable parameters between the healthy cohort and the fetuses with OSD were calculated using the Wilcoxon–Mann-Whitney test. The Wilcoxon test was used to analyze differences in the paired measurements of individual fetuses before and after the operation. All statistics in the PF characteristics section were performed using Analyze-it (Analyze-it for Microsoft Excel 4.81.4; Analyze-it Software, Leeds, UK). A P value $< .05$ indicated statistical significance.

RESULTS

Demographics

MR imaging examinations in controls were performed at a mean GA of 27.9 ± 5.3 weeks (range,18.6–38.3 weeks). In fetuses with OSD, the mean GA at MR imaging was 23.6 ± 0.3 weeks (range, 19.3–27.3 weeks). Descriptive statistics for the study parameters in controls and cases with OSD are shown in Table 1.

Reproducibility Study

The intraobserver ICCs for the PF area, VW, TCD, and CHL were excellent. Conversely, TDPF (0.729), the pontine thickness (0.59),

Table 1: Mean and SDs of different parameters for the controls and fetuses with nonoperated and operated spinal dysraphism

	Controls (*n* = 274)	Nonoperated OSD (*n* = 52)	Postoperative OSD (*n* = 23)
GA (wk)	27.9 (5.3)	23.6 (0.3)	26.0 (1.1)
VW (mm)	6 (1.6)	11.9 (4.6)	12.3 (3.6)
PF area (cm^2)	7.4 (2.2)	3.1 (0.9)	4.44 (0.83)
TDPF (mm)	39.8 (7.1)	22.3 (3.4)	27.2 (4.1)
TCD (mm)	31.6 (6.4)	21.8 (3)	24.4 (3.3)
Ratio TCD/TDPF	0.79 (0.05)	1 (0.05)	0.9 (0.07)
CHL (mm)	8.8 (1.9)	−9.3 (5.4)	−0.7 (5.2)
CSA	87.1° (8.3°)	62.5° (9.7°)	76.6° (10.9°)

Table 2: Equations of the regression curves with their respective levels of significance and the R^2 for the TCD, CHL, CSA, TDPF, PF area, VW, and ratio TCD/TDPF

Parameter	Regression Curve	*P* Value[a]	R^2
TCD	$y = 64.38–30.58*x + 4.606*x^2$	<.0001	0.907
CHL	$y = −78.26 + 31.6*x–2.849*x^2$	<.0001	0.117
CSA	$y = 21.61 + 12.44*x$	<.0001	0.265
TDPF	$y = −65.28 + 19.95*x$	<.0001	0.894
PF area	$y = 14.25–8.537*x +1.369*x^2$	<.0001	0.875
VW	$y = 51.5–17.59*x + 1.692*x^2$	.0257	0.027
Ratio TCD/TDPF	$y = 3.87–1.215*x + 0.1192*x^2$	<.0001	0.243

Note:—x indicates square root (gestational age in weeks); R^2, coefficient of determination.

[a] *P* value according to Fisher *F* test.

the foramen magnum diameter (0.44), mamillopontine distance (0.66), CSA (0.60), and the width of the cisterna magna (0.48) had a fair-to-good reproducibility. Measurements of the other parameters showed a low reproducibility (tentorial length) or were unreliable (level of brain stem kinking, medullar length, and pontine length). Interobserver reproducibility was moderate for TCD, CHL, CSA, and TDPF. PF area and VW had good interrater reliability.

Posterior Fossa Characteristics

All parameters with an intra- and interobserver ICC ≥ 0.5 were taken into account for further analysis. Normative curves for VW, PF area, TDPF, TCD, CSA, TCD/TDPF, and CHL were calculated and are shown in Table 2. Figure 2 shows the individual observations for cases with OSD, which were all significantly different from what was measured in healthy fetuses ($P < .0001$).

Short-Term (<7 Days) Postoperative Changes

When we considered the observed over expected ratio values, fetal surgery was associated with a significant difference in cerebellar herniation ($P < .0001$), TCD/TDPF ($P = .0002$), TCD ($P = .0127$), PF area ($P = .0003$), TDPF ($P = .0127$), CSA ($P < .0001$), and VW ($P = .0002$). Figure 3 shows the individual observations and boxplots in patients with OSD for all tested parameters, both pre- and postoperatively. In 18/23 (78%) fetuses, the postoperative observed over expected ratios of the PF area were improving toward normal compared with the preoperative measurement. The same was true for the observed over expected ratios of the TDPF in 15/23 (65%), for the observed over expected ratios of the TCD in 7/23 (30%), and for the observed over expected TCD/TDPF ratios in 4/23 (17%) postoperative fetuses. The observed over expected ratio level of cerebellar herniation increased, meaning that the Chiari II–associated changes were reduced in 19/23 fetuses at 1 week after the operation. The observed over expected ratios of the CSA increased toward, the normal range in 16/23 (70%) postoperative fetuses. The VW increased in 17/23 (74%) fetuses, but there was no difference in the ACi between the pre- and postoperative examinations ($P = .46$), with a mean of the preoperative measurements of 0.22 and, postoperatively, of 0.23.

DISCUSSION

A large number of structural measurements on MR images of the PF have been suggested to characterize changes attributed to OSD. Some are believed to be clinically relevant due to their relation to the symptoms associated with a small PF, Chiari II malformation, and brain stem compression.[32] These changes are represented by the degree of cerebellar herniation, mamillopontine distance, brain stem kinking, tentorial hypoplasia, and the configuration of the fourth ventricle. These features can be measured reproducibly in the postnatal period.[8] In the era of prenatal diagnosis and fetal surgery, logically, the same measurements are also used. Yet, in case of prenatal surgery, these are measured on midgestational prenatal images at a time when the condition is still progressive and image quality is different. We therefore investigated whether those measurements were reproducible in the window of interest. For instance, we did not evaluate the fourth ventricle because it is hardly visible around 20–24 weeks in fetuses with OSD. Conversely, we looked at parameters characterizing the brain stem (mamillopontine distance, pontine thickness, pontine length, foramen magnum diameter, level of brain stem kinking, medulla length, tentorial length, and cisterna magna width) and the PF as a whole (TCD, TDPF, PF area, CHL, TCD/TDPF, and CSA). Due to its clinical relevance in fetuses with OSD and the impact on outcome after fetal surgery,[33] we also included the VW.

Herein, we conclude that brain stem parameters in fetuses with OSD cannot be measured reliably at midgestation; thus, we were not able to objectively measure brain stem elongation and displacement in utero. For fetuses with OSD, we believe this issue is due to the limited spatial resolution of fetal MR imaging at that point in gestation when the structures involved are in the millimetric range so that the slightest measurement error has a tremendous impact on statistics. Another reason is that in OSD, there is a decrease or even absence of extra-axial CSF, further limiting the contrast resolution of MR imaging.[34] Contrast resolution is essential for accurate evaluation of small PF structures as well as some additional cerebral lesions due to the presence of different structures (medulla oblongata, pons, vermis, cerebellum) in a small area.[35] There is no difference between balanced steady-state free-precession sequences and half-Fourier rapid relaxation with relaxation enhancement sequences for measuring the foramen magnum.[19] Conversely, when we looked at larger structures (PF area, TCD, and TPFD) and/or with more abundant contrast between fluid and soft tissue (CHL and VW), the ICC values were much better. We confirmed this finding for those parameters typical for the smaller posterior fossa and cerebellar descent in the pathologic subgroup, as others did.[17,18,20,22] Moreover, the difference with our gestational age–matched healthy patients was highly significant, again as previously described.[17]

Fetal surgery has been shown to reverse those posterior fossa

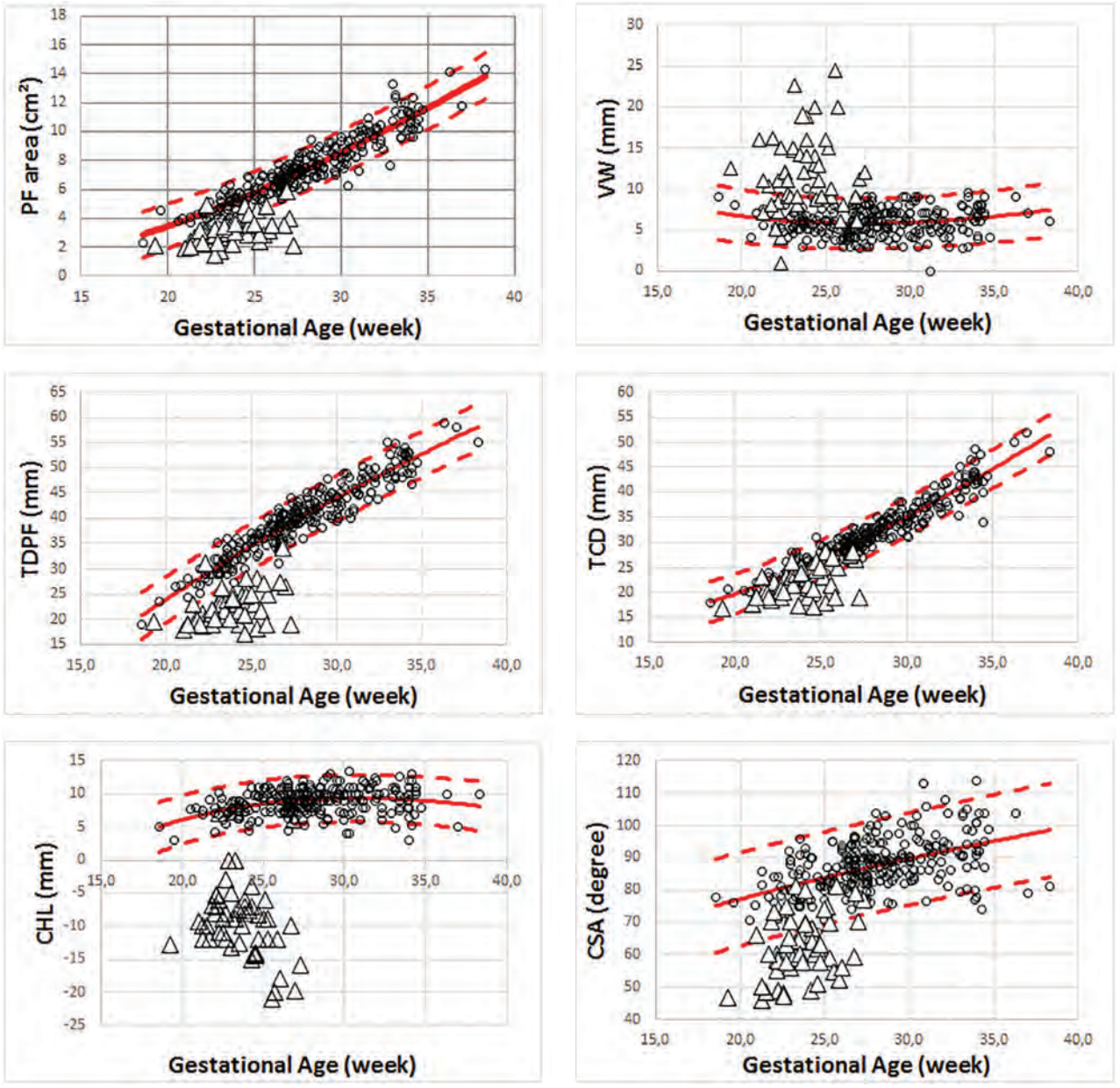


FIG 2. The individual observations in the control population (*black circles*) with the mean (*full line*) and 95% confidence interval (*dashed lines*) compared with the preoperative fetuses with open spinal dysraphism (*white triangles*) for the posterior fossa area, ventricular width, transverse diameter of the posterior fossa, transverse cerebellar diameter, cerebellar herniation level, and clivus-supraocciput angle.

changes.[11,14,18] Other investigators used the above parameters to measure those typically ≥4 weeks after the operation.[18] In the present study, we acquired images within 2 weeks. In line with ultrasound observations, we quantified significant changes during that short observation period. Within 1 week, 26% of operated fetuses had a PF area within the normal range, and in 52%, the TCD was normal. Furthermore, the cerebellar herniation level was at or above the foramen magnum in 52% of fetuses, and 70% had a normal CSA. These acute changes in the PF following closure of the defect are in line with the theory of McLone and Knepper.[36] In other words, it seems that the effects of fetal surgery on the PF are already evident and can be quantified very early postoperatively. They are very likely to persist because others observed the same effects later on and even confirmed them after birth.[14,18,37] This outcome might be an interesting proxy for measurement of the efficacy of fetal surgery in clinical studies.

In this short-term follow-up study, we observed a postoperative increase in ventricular width within 1 week in most patients. Such increase is in line with observations made by others, though several weeks after fetal surgery.[18,33] They suggest that there is still a certain degree of obstructive ventricular widening. The dynamics of CSF fluid production and resorption in OSD are still poorly understood. It may take some time after fetal surgery for CSF fluid circulation to normalize after stopping its egress.[11]

In healthy fetuses, the ACi drops dramatically between 24 and 27 weeks, which means that there is, during that time period, a proportional increase in the parenchymal brain component. In our patients undergoing fetal surgery, the ACi remained stable.

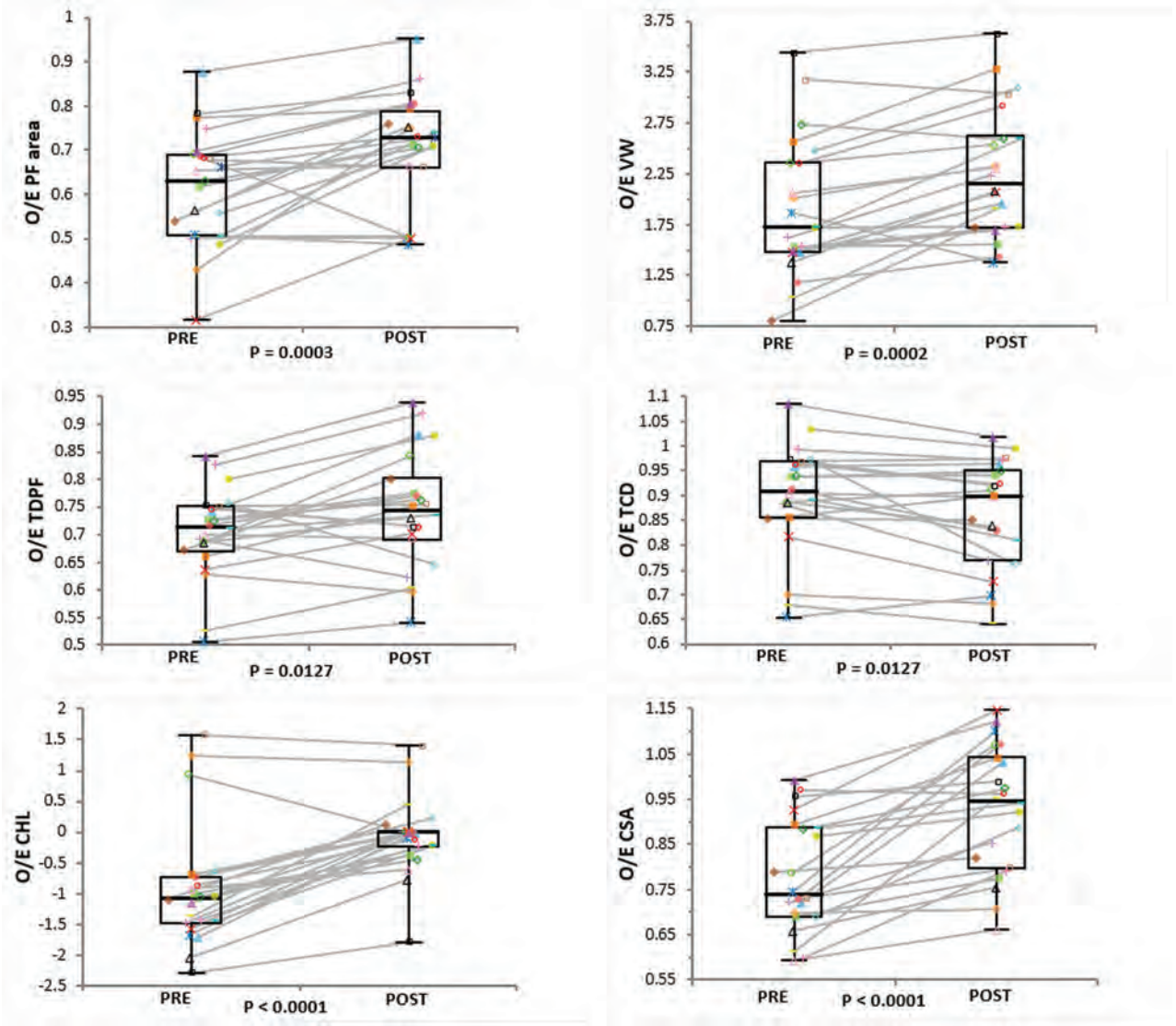


FIG 3. Boxplot demonstrating the minimum, first quartile, median, third quartile, and maximum of the observed over expected ratio in preoperative fetuses with open spinal dysraphism and postoperative fetuses with OSD at 1 week for the posterior fossa area, ventricular width, transverse diameter of the posterior fossa, transverse cerebellar diameter, cerebellar herniation level, and clivus-supraocciput angle.

This might be counterintuitive and contradicts the findings of Rethmann et al.[18] They measured the ACi and observed a drop in the ACi; yet, that was 4 weeks after the operation and continued after birth. These contrasting findings can be explained in different ways. A drop in ACi would suggest a proportional increase in biparietal cerebral diameter, hence a larger parenchymal component. Conversely, the increase in ACi for a comparable VW in our cohort would suggest that the parenchymal component decreases. Although tempting, both groups cannot be compared because measurements were performed at different gestational ages. In healthy fetuses, the ACi spontaneously declines between 24 and 27 weeks. If fetuses with OSD follow this normal evolution, our findings may eventually align with these of Rethmann et al and can be explained by spontaneous and normal evolution. Unfortunately, we have no longitudinal follow-up MR images to further study these observations.

We have looked into prenatal measurements characterizing the brain stem, which, to our knowledge, was not performed in detail before. We based our evaluation on our own normative values. Furthermore, we documented all parameters in a relatively large pathologic population in the narrow gestational age range that is relevant to prenatal spina bifida repair. Although the number of operated fetuses was not very large, we were able to describe early postoperative changes therein. There are, however, some shortcomings. First, our control fetuses definitely had normal CNS findings on both MR imaging and ultrasound but were not truly fully healthy fetuses. Controls underwent MR imaging because of other congenital abnormalities, presumed not to be associated with CNS abnormalities.

Second, we did not report on advanced MR images, such as DWI and DTI, which may also provide relevant information. We definitely acknowledge the potential of DWI and DTI because they may detect more subtle abnormalities below the anatomic level. Woitek et al[38] already showed that fetuses with spina bifida have increased fractional anisotropy compared with normally developing fetuses, however without reporting the functional im-

pact. Although we acquired such sequences in fetuses with spina bifida, we were lacking those in healthy fetuses or the controls used in this study; hence, we could not interpret the findings. Third, we describe in utero findings without correlation to postnatal short- or long-term follow-up or early postnatal MR imaging confirmation. Although relevant, such a follow-up was beyond the scope of this study as was a comparison of these posterior fossa measurements with those in a cohort that underwent postnatal repair. Postnatal evaluation would most likely have identified additional findings, such as subependymal heterotopias.[18,39] These are often missed in utero before as well as after fetal surgery.

CONCLUSIONS

This study showed that the brain stem cannot be reliably characterized using the current panel of measurements in fetuses with OSD. Conversely, posterior fossa measurements are demonstrated to be reliable in the evaluation of fetuses with OSD. In addition, these were significantly different from those in the healthy population and changed within 7 days after prenatal surgery. This finding advocates for their use in the evaluation of fetuses with OSD on fetal MR imaging on a routine basis before and shortly after a prenatal operation.

Disclosures: Tom Vercauteren—*UNRELATED*: *Employment*: University College London, *Comments*: main employer for the duration of this work; *Grant*: Wellcome Trust (WT101957)/Engineering and Physical Sciences Research Council (NS/A000027/1*), *Comments*: *Support for Travel to Meetings for the Study or Other Purposes*: Wellcome Trust/Engineering and Physical Sciences Research Council, *Comments*: travel support from the listed grant funding.* Luc De Catte—*UNRELATED*: *Employment*: consultant to Feto-Maternal Medicine AZ St-Jan Brugge, *Comments*: 1 day/2 weeks clinical activity.* Philippe Demaerel—*UNRELATED*: *Board Membership*: Editorial Board of *Neuroradiology*. Jan Deprest—*RELATED*: *Other:* Great Ormond Street Hospital Children's Charity, *Comments*: This grant pays part of my academic (research) time to my institution*; *UNRELATED*: *Grants/Grants Pending*: MEDRI, Wellcome Trust, *Comments*: We have a grant on the use of advanced imaging techniques in fetal surgery (Wellcome) and a research grant on novel implant materials for Holder of a research chair "POPART" (Pelvic Organ Prolapse - Advanced Research and Technology) - sponsored by MEDRI*; *Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed*: Chiesi Farmaceutici, *Comments*: paid travel and accommodations to the Sharing Progress in Neonatology 2018 meeting.* *Money paid to the institution.

REFERENCES

1. Garne E, Loane M, Addor M-C, et al. **Congenital hydrocephalus: prevalence, prenatal diagnosis and outcome of pregnancy in four European regions.** *Eur J Paediatr Neurol* 2010;14:150–55 CrossRef Medline
2. Khoshnood B, Loane M, de Walle H, et al. **Long term trends in prevalence of neural tube defects in Europe: population-based study.** *BMJ* 2015;351:h5949 CrossRef Medline
3. Canfield MA, Mai CT, Wang Y, et al; National Birth Defects Prevention Network. **The association between race/ethnicity and major birth defects in the United States, 1999–2007.** *Am J Public Health* 2014;104:e14e–23 CrossRef Medline
4. Oakeshott P, Hunt GM. **Long-term outcome in open spina bifida.** *Br J Gen Pract* 2003;53:632–36 Medline
5. Rintoul NE, Sutton LN, Hubbard AM, et al. **A new look at myelomeningoceles: functional level, vertebral level, shunting, and the implications for fetal intervention.** *Pediatrics* 2002;109:1–7 CrossRef Medline
6. Mitchell LE, Adzick NS, Melchionne J, et al. **Spina bifida.** *Lancet* 2004;364:1885–95 CrossRef Medline
7. Barkovich AJ, **Congenital malformations of the brain and skull**. In: Barkovich AJ, ed, Ovid Technologies, Inc. *Pediatric Neuroimaging*. 4th ed. Philadelphia: Lippincott Williams & Wilkins; 2005:374–84
8. Geerdink N, van der Vliet T, Rotteveel JJ, et al. **Essential features of Chiari II malformation in MR imaging: an interobserver reliability study, Part 1.** *Childs Nerv Syst* 2012;28:977–85 CrossRef Medline
9. Geerdink N, van der Vliet T, Rotteveel JJ, et al. **Interobserver reliability and diagnostic performance of Chiari II malformation measures in MR imaging, Part 2.** *Childs Nerv Syst* 2012;28:987–95 CrossRef Medline
10. Biggio JR Jr, Wenstrom KD, Owen J. **Fetal open spina bifida: a natural history of disease progression in utero.** *Prenat Diagn* 2004;24:287–89 CrossRef Medline
11. Bruner JP, Tulipan N, Reed G, et al. **Intrauterine repair of spina bifida: preoperative predictors of shunt-dependent hydrocephalus.** *Am J Obstet Gynecol* 2004;190:1305–12 CrossRef Medline
12. Adzick NS, Thom EA, Spong CY, et al; MOMS Investigators. **A randomized trial of prenatal versus postnatal repair of myelomeningocele.** *N Eng J Med* 2011;364:993–1004 CrossRef Medline
13. Saleem SN, Said AH, Abdel-Raouf M, et al. **Fetal MRI in the evaluation of fetuses referred for sonographically suspected neural tube defects (NTDs): impact on diagnosis and management decision.** *Neuroradiology* 2009;51:761–72 CrossRef Medline
14. Sutton LN, Adzick NS, Bilaniuk LT, et al. **Improvement in hindbrain herniation demonstrated by serial fetal magnetic resonance imaging following fetal surgery for myelomeningocele.** *JAMA* 1999;282:1826–31 CrossRef Medline
15. Heuer GG, Moldenhauer JS, Scott Adzick NS. **Prenatal surgery for myelomeningocele: review of the literature and future directions.** *Childs Nerv Syst* 2017;33:1149–55 CrossRef Medline
16. Danzer E, Adzick NS. **Fetal surgery for myelomeningocele: patient selection, perioperative management and outcomes.** *Fetal Diagn Ther* 2011;30:163–73 CrossRef Medline
17. Woitek R, Dvorak A, Weber M, et al. **MR-based morphometry of the posterior fossa in fetuses with neural tube defects of the spine.** *PLoS One* 2014;9:e112585 CrossRef Medline
18. Rethmann C, Scheer I, Meuli M, et al. **Evolution of posterior fossa and brain morphology after in utero repair of open neural tube defects assessed by MRI.** *Eur Radiol* 2017;27:4571–80 CrossRef Medline
19. Abele TA, Lee SL, Twickler DM. **MR imaging quantitative analysis of fetal Chiari II malformations and associated open neural tube defects: balanced SSFP versus half-Fourier RARE and interobserver reliability.** *J Magn Reson Imaging* 2013;38:786–93 CrossRef Medline
20. Tsai T, Bookstein FL, Levey E, et al. **Chiari-II malformation: a biometric analysis.** *Eur J Pediatr Surg* 2002;12(Suppl 1):S12–18 Medline
21. Osuagwu FC, Lazareff JA, Rahman S, et al. **Chiari I anatomy after ventriculoperitoneal shunting: posterior fossa volumetric evaluation with MRI.** *Childs Nerv Syst* 2006;22:1451–56 CrossRef Medline
22. Grant RA, Heuer GG, Carrión GM, et al. **Morphometric analysis of posterior fossa after in utero myelomeningocele repair.** *J Neurosurg Pediatr* 2011;7:362–68 CrossRef Medline
23. Chen SC, Simon EM, Haselgrove JC, et al. **Fetal posterior fossa volume: assessment with MR imaging.** *Radiology* 2006;238:997–1003 CrossRef Medline
24. Garel C. **Methodology.** In: Carel C, Delezoide AL, Delezoide V, eds. *MRI of the Fetal Brain*. Berlin: Springer-Verlag; 2004
25. Saleem SN. **Fetal MRI: an approach to practice—a review.** *J Adv Res* 2014;5:507–23 CrossRef Medline
26. Garel C. **Fetal cerebral biometry: normal parenchymal findings and ventricular size.** *Eur Radiol* 2005;15:809–13 CrossRef Medline
27. Tilea B, Alberti C, Adamsbaum C, et al. **Cerebral biometry in fetal magnetic resonance imaging: new reference data.** *Ultrasound Obstet Gynecol* 2009;33:173–81 CrossRef Medline
28. D'Addario V, Pinto V, Del Bianco A, et al. **The clivus-supraocciput angle: a useful measurement to evaluate the shape and size of the fetal posterior fossa and to diagnose Chiari II malformation.** *Ultrasound Obstet Gynecol* 2002;18:146–49 Medline
29. Doel T, Shakir DI, Pratt R, et al. **GIFT-Cloud: A data sharing and**

collaboration platform for medical imaging research. *Comput Methods Programs Biomed* 2017;139:181–90 CrossRef Medline
30. Koo TK, Li MY. **A guideline of selecting and reporting intraclass correlation coefficients for reliability research.** *J Chirop Med* 2016; 15:155–63 CrossRef Medline
31. Cicchetti DV. **Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology.** *Psychological Assessment* 1994;6:284–90 CrossRef
32. Stevenson KL. **Chiari type II malformation: past, present, and future.** *Neurosurg Focus* 2004;16:E5 Medline
33. Tulipan N, Wellons JC 3rd, Thom EA, et al; MOMS Investigators. **Prenatal surgery for myelomeningocele and the need for cerebrospinal fluid shunt placement.** *J Neurosurg Pediatr* 2015;16:613–20 CrossRef Medline
34. Danzer E, Johnson MP, Bebbington M, et al. **Fetal head biometry assessed by fetal magnetic resonance imaging following in utero myelomeningocele repair.** *Fetal Diagn Ther* 2007;22:1–6 CrossRef Medline
35. Limperopoulos C, Tworetzky W, McElhinney DB, et al. **Brain volume and metabolism in fetuses with congenital heart disease: evaluation with quantitative magnetic resonance imaging and spectroscopy.** *Circulation* 2010;121:26–33 CrossRef Medline
36. McLone DG, Knepper PA. **The cause of Chiari II malformation: a unified theory.** *Pediatr Neurosci* 1989;15:1–12 CrossRef Medline
37. Nagaraj UD, Bierbrauer KS, Zhang B, et al. **Hindbrain herniation in Chiari II malformation on fetal and postnatal MRI.** *AJNR Am J Neuroradiol* 2017;38:1031–36 CrossRef Medline
38. Woitek R, Prayer D, Weber M, et al. **Fetal diffusion tensor quantification of brain stem pathology in Chiari II malformation.** *Eur Radiol* 2016;26:1274–83 CrossRef Medline
39. Nagaraj UD, Peiro JL, Bierbrauer KS, et al. **Evaluation of subependymal gray matter heterotopias on fetal MRI.** *AJNR Am J Neuroradiol* 2016;37:720–75 CrossRef Medline

"Ears of the Lynx" MRI Sign Is Associated with SPG11 and SPG15 Hereditary Spastic Paraplegia

B. Pascual, S.T. de Bot, M.R. Daniels, M.C. França Jr, C. Toro, M. Riverol, P. Hedera, M.T. Bassi, N. Bresolin, B.P. van de Warrenburg, B. Kremer, J. Nicolai, P. Charles, J. Xu, S. Singh, N.J. Patronas, S.H. Fung, M.D. Gregory, and J.C. Masdeu

ABSTRACT

BACKGROUND AND PURPOSE: The "ears of the lynx" MR imaging sign has been described in case reports of hereditary spastic paraplegia with a thin corpus callosum, mostly associated with mutations in the *spatacsin vesicle trafficking associated* gene, causing Spastic Paraplegia type 11 (SPG11). This sign corresponds to long T1 and T2 values in the forceps minor of the corpus callosum, which appears hyperintense on FLAIR and hypointense on T1-weighted images. Our purpose was to determine the sensitivity and specificity of the ears of the lynx MR imaging sign for genetic cases compared with common potential mimics.

MATERIALS AND METHODS: Four independent raters, blinded to the diagnosis, determined whether the ears of the lynx sign was present in each of a set of 204 single anonymized FLAIR and T1-weighted MR images from 34 patients with causal mutations associated with SPG11 or Spastic Paraplegia type 15 (SPG15). 34 healthy controls, and 34 patients with multiple sclerosis.

RESULTS: The interrater reliability for FLAIR images was substantial (Cohen κ, 0.66–0.77). For these images, the sensitivity of the ears of the lynx sign across raters ranged from 78.8 to 97.0 and the specificity ranged from 90.9 to 100. The accuracy of the sign, measured by area under the receiver operating characteristic curve, ranged from very good (87.1) to excellent (93.9).

CONCLUSIONS: The ears of the lynx sign on FLAIR MR imaging is highly specific for the most common genetic subtypes of hereditary spastic paraplegia with a thin corpus callosum. When this sign is present, there is a high likelihood of a genetic mutation, particularly associated with SPG11 or SPG15, even in the absence of a family history.

ABBREVIATIONS: AUC = area under the curve; HSP-TCC = hereditary spastic paraplegia with a thin corpus callosum; ROC = receiver operating characteristic curve

Autosomal recessive hereditary spastic paraplegia with a thin corpus callosum (HSP-TCC for short) presents clinically as a progressive spastic paraplegia, usually beginning during infancy or puberty, and is often associated with cognitive impairment, occasionally antedating the onset of paraparesis.[1-6] Both symptoms worsen slowly for decades. Mutations in the *spatacsin vesicle trafficking associated* (SPG11) gene, coding for spatacsin, are most commonly identified in these patients.[1-6] Mutations in the gene encoding spastizin (*zinc finger fyve domain-containing protein 26, ZFYVE26*, causing SPG15) or other genes are rarer.[2,4,5,7] The disorder is transmitted as an autosomal recessive trait; thus, it is common for patients to present without a family history of the disease.[4,6] Given the clinical presentation, these patients are most often studied with MR imaging to detect disorders affecting the brain and, particularly, the pyramidal tract.[8] MR imaging shows a thin corpus callosum, but this sign is not unique to HSP-TCC,[9-17]

Received July 5, 2018; accepted after revision October 30.

From the Departments of Neurology (B.P., M.R.D., J.C.M.), Biostatistics (J.X.), and Radiology (S.S., S.H.F.), Houston Methodist Research Institute, Houston, Texas; Department of Neurology (S.T.d.B.), Leiden University Medical Centre, Leiden, the Netherlands; Department of Neurology (M.C.F.), University of Campinas, Campinas, Brazil; National Institutes of Health Intramural Research Program (C.T., N.J.P., M.D.G.), Bethesda, Maryland; Department of Neurology (M.R.), Clínica Universidad de Navarra, Pamplona, Spain; Department of Neurology (P.H.), Vanderbilt University Medical Center, Nashville, Tennessee; Laboratory of Molecular Biology (M.T.B.), Scientific Institute Istituto di Ricovero e Cura a Carattere Scientifico E. Medea, Bosisio Parini, Lecco, Italy; Department of Neuroscience and Mental Health (N.B.), University Hospital Policlinico Ca'Granda, University of Milan, Milan, Italy; Department of Neurology (B.P.v.d.W.), Donders Institute for Brain, Cognition, and Behaviour, Radboud University Medical Center, Nijmegen, the Netherlands; Department of Neurology (B.K.), University Medical Center Groningen, Groningen, the Netherlands; Department of Neurology (J.N.), Maastricht University Medical Centre, Maastricht, the Netherlands; and Department of Genetics (P.C.), Hôpital Pitié-Salpêtrière, Paris, France.

This study was partially funded by the Chao, Graham, Harrison, and Nantz Funds of the Houston Methodist Foundation. Portions of the data studied were obtained under protocol OHSRP 12231 and were supported (in part) by the National Institutes of Mental Health Intramural Research Program.

This work does not necessarily reflect the views of the US government or the National Institutes of Health.

Please address correspondence to Belen Pascual, PhD, Nantz National Alzheimer Center, Stanley H. Appel Department of Neurology, 6560 Fannin St., Scurlock Tower, 8th Floor, Houston, TX 77030, email: bpascual@houstonmethodist.org

Indicates article with supplemental on-line table.

http://dx.doi.org/10.3174/ajnr.A5935

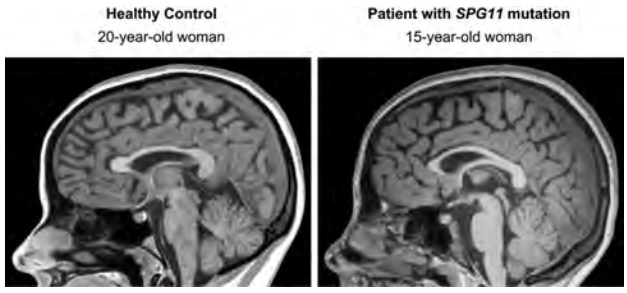


FIG 1. Corpus callosum size similar to normal in a patient with HSP-TCC. The patient had mild cognitive impairment and a mild spastic paraparesis. She had SPG 11 pathogenetic mutations. The ear of the lynx sign in this patient is shown in Fig 2.

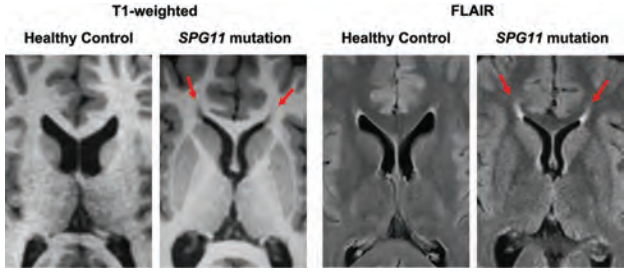


FIG 2. Ears of the lynx on MR imaging. Axial images across the anterior forceps of the corpus callosum. Note an abnormality in the region of the forceps minor of the corpus callosum, corresponding to the genu fibers, which appear dark on T1-weighted and bright on FLAIR images (*arrows*). Midline sagittal images from the same individuals are seen in Fig 1.

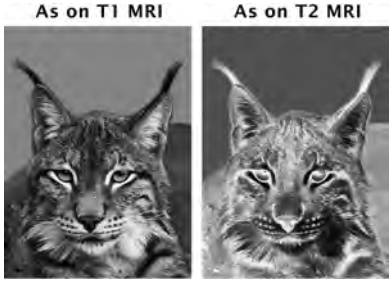


FIG 3. Ears of the lynx. Shown are the gray-scale and corresponding negative images of the head of a lynx. The hair tufts at the tip of the ears resemble the MR imaging finding described here. Modified with permission from an original photo taken by Aleksandar Vasic.

and there are patients with HSP-TCC caused by SPG11 mutations in whom the corpus callosum has an apparently normal thickness (Fig 1).

Another MR imaging feature may be helpful to lead to the diagnosis. A characteristic abnormality affecting the region of the forceps minor of the corpus callosum[18,19] has been described as the "ears of the lynx" sign.[20] The forceps minor of the corpus callosum, corresponding to the genu fibers, has prolonged T1 and T2 values. As a result, this region appears bright on T2-weighted and dark on T1-weighted images (Fig 2). On axial sections, the abnormality bears a remarkable resemblance to the ears of a lynx, with the areas of abnormal signal reminiscent of the tufts of hair crowning the tips of the ears of this animal (Fig 3).

If specific, the presence of this MR imaging finding in a seemingly sporadic case could be very useful to guide genetic testing or help in interpreting genetic findings. Thus, separating patients with genetic mutations from those who may show a similar radiologic finding caused by more common etiologies, such as gliosis at the calloso-caudate angle of the frontal horn of the lateral ventricles[21] or multiple sclerosis, is much more relevant in terms of imaging than helping to differentiate among rare genetic variants. If a genetic disorder is suggested by the radiologic picture, a detailed genetic study should be performed to determine the exact genetic etiology. Therefore, our objective was to determine the sensitivity and specificity of this sign for patients with known genetic mutations associated with HSP-TCC versus potentially similar common patterns in this age group, such as gliosis at the calloso-caudate angle or demyelinating disease. Thus, we compared the MRIs of patients with HSP-TCC who were found by genetic testing to have causal mutations associated with SPG11 or SPG15 with those of healthy controls and of patients with multiple sclerosis, a common cause of white matter changes in an age group similar to the HSP-TCC sample.

Table 1: Demographics

	Diagnosis		
	SPG11 (n = 31) or SPG15 (n = 3) Mutations	MS	Healthy Control
Number	34	34	34
Sex, female/male	16:18	16:18	16:18
Mean age (SD) at MRI (yr)	24.6 (7.8)	25.0 (6.8)	24.6 (7.5)

MATERIALS AND METHODS

All procedures were approved by the Human Studies Committees of the institutions involved, including the National Institutes of Health Office of Human Subjects Research for the healthy controls and patients with MS. In all cases, only retrospective, anonymized information was used for the study; therefore, individual written informed consent was waived. Patients with HSP-TCC had been recruited between 2002 and 2014 at 9 institutions, listed in the authors' affiliations. Nine of these patients had been included in a study[22] reporting on the clinical course of genetic variants of HSP-TCC, not on the radiologic sign described in this article. Another 5 patients had been included in a study[23] reporting diffusion tensor imaging changes in white matter and brain volume in HSP-TCC, but this study did not discuss the radiologic sign studied here. Finally, another 4 patients with HSP-TCC had been included in the original report of the ears of the lynx sign,[20] but its sensitivity or specificity was neither studied nor described in that article. Patients with MS and healthy controls had participated in other studies at the National Institutes of Health, but we had access to only their anonymized MRIs, obtained from 2002 to 2013, which were selected to match the HSPTCC sample in age and sex. All patients with HSP-TCC had identified pathogenic mutations, 31 associated with SPG11 and 3 with SPG15. Each of the 3 groups (HSP-TCC, MS, healthy controls) consisted of 34 subjects, split almost evenly across sexes, with a mean age at MR imaging of approximately 25 years (Table 1).

Anonymized MRIs were used for all groups. One axial T2 FLAIR and one T1-weighted image across the anterior forceps of the corpus callosum (Fig 2) of each patient with HSP-TCC and each control (healthy or MS) were randomized for presentation to

the evaluating raters. Thus, each rater read 204 images blindly. Single images were presented on a computer screen containing the image to be evaluated and buttons to indicate the presence or absence of the ears of the lynx sign and the quality of the image. Raters included 3 US board-certified radiologists with Certificate of Additional Qualification in Neuroradiology, with an average experience of 22 years of practice (range, 11–32 years), and a board-certified neurologist with 12 years of experience in imaging. Before reading the images blindly, the raters were trained to recognize the ears of the lynx with 4 images positive for it and 4 images negative for it, including both T2 and T1 studies. They were not given information on the diseases of the patients but were only shown the morphology of the ears of the lynx sign. For the unknown images, the raters were asked to make a forced choice, deciding whether the image contained the ears of the lynx sign. In addition, they had to define whether each image was of good or poor quality.

Table 2: Interrater reliability (Cohen κ)

	Cohen κ for FLAIR (95% CI)	Cohen κ for T1-Weighted (95% CI)
Rater 1 vs 2	0.74 (0.60–0.88)	0.51 (0.32–0.71)
Rater 2 vs 3	0.74 (0.60–0.88)	0.42 (0.24–0.63)
Rater 3 vs 4	0.77 (0.63–0.91)	0.75 (0.54–0.96)
Rater 1 vs 3	0.66 (0.51–0.82)	0.59 (0.38–0.80)
Rater 2 vs 4	0.71 (0.56–0.86)	0.40 (0.20–0.60)
Rater 1 vs 4	0.73 (0.59–0.87)	0.62 (0.42–0.83)
All raters	0.72 (0.61–0.82)	0.53 (0.35–0.69)

Statistical Analysis

This is a retrospective case-control study. Given the absence of previous studies on the sensitivity and specificity of the ears of the lynx sign, we arrived at the sample size by following guidelines for studies with an unknown effect size.[24] The Cohen κ statistic was used to calculate interrater reliability. The degree of agreement was interpreted on the basis of the κ coefficients as follows: 0–0.2, poor; 0.21–0.4, fair; 0.41–0.6, moderate; 0.61–0.8, substantial; 0.81–1, almost perfect.[25] Sensitivity, specificity, and the 95% CIs for T1-weighted and FLAIR images were calculated. The area under the curve (AUC) was obtained from the logistic regression model. The logistic regression model with cluster, which allowed interrater correlation, was also used to calculate the effect of the rated quality of images on the agreement of the rating with the true allocation of each image. All analyses were performed with STATA version 15 (StataCorp, College Station, Texas). Statistical significance was defined as a 2-tailed $P < .05$ for all tests.

Table 3: ROC results with sensitivity and specificity for FLAIR and T1-weighted images

	ROC Results		
	AUC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Sensitivity and specificity for FLAIR images			
Rater 1	93.9 (89.4–98.5)	97.0 (84.2–99.9)	90.9 (81.3–96.6)
Rater 2	87.1 (79.7–94.5)	81.8 (64.5–93.0)	92.4 (83.2–97.5)
Rater 3	87.1 (79.6–94.6)	78.8 (61.1–91.0)	95.5 (87.3–99.1)
Rater 4	89.4 (82.3–96.5)	78.8 (61.1–91.0)	100 (94.6–100)
Sensitivity and specificity for T1-weighted images			
Rater 1	84.8 (76.1–93.5)	71.4 (51.3–86.8)	98.2 (90.4–100)
Rater 2	81.3 (72.0–90.5)	75.0 (55.1–89.3)	87.5 (75.9–94.8)
Rater 3	68.8 (59.4–78.1)	39.3 (21.5–59.4)	98.2 (90.4–100)
Rater 4	69.6 (60.4–78.9)	39.3 (21.5–59.4)	100 (93.6–100)

RESULTS

The interrater reliability for FLAIR images was substantial, ranging from 0.66 to 0.77. For T1-weighted images, it was moderate to substantial, ranging from 0.51 to 0.75 (Table 2). In the T2 FLAIR images, the ears of the lynx sign sensitivity across raters ranged from 78.8 to 97.0 and the specificity ranged from 90.9 to 100 (Table 3). On T1-weighted images, the sign was not detected as often, with a sensitivity ranging from 39.3 to 75, but the specificity was still high (87.5–100) (Table 3). The receiver operating characteristic curves (ROCs) showed better

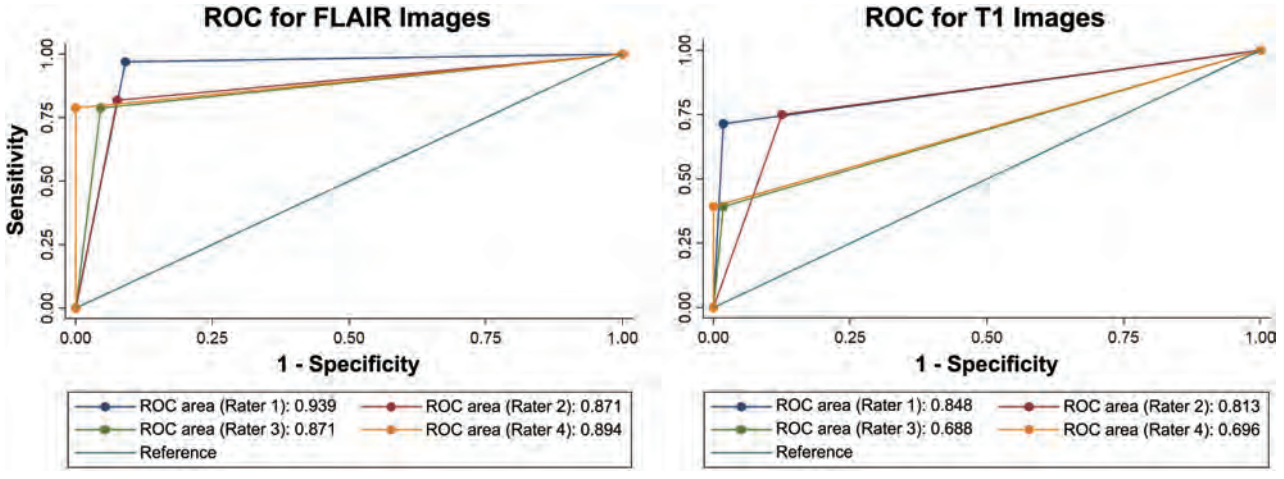


FIG 4. ROCs for T1-weighted and FLAIR images. The receiver operating characteristic curves show better discrimination for FLAIR than for T1-weighted images. For FLAIR images, the area under the ROC curve showed that the ears of the lynx sign performed in the very good-to-excellent range for a diagnostic test.

discrimination for FLAIR than for T1-weighted images (Fig 4). For FLAIR images, the area under the ROC curve showed that the ears of the lynx sign performed in the very good-to-excellent (87.1–93.9) range for a diagnostic test. The quality of the images did not make a significant difference between the blinded rating and the true outcome ($P = .29$) for FLAIR images, but it affected the agreement between the rating and outcome ($P < .001$) for T1-weighted images (On-line Table).

DISCUSSION

Our study indicates that the ears of the lynx sign on axial FLAIR MR imaging is associated with HSP-TCC caused by genetic mutations characteristic of SPG11 and SPG15. We did not explore its association with other genetic mutations because this radiologic finding is probably not sufficient to make the genetic diagnosis; a genetic study is required. However, our study shows that the ears of the lynx sign is helpful to suggest a genetic disorder in a patient with an apparently sporadic spastic paraparesis or cognitive impairment. This is important because responsible mutations have been sought particularly in families with several affected members,[3,6,26] though pathogenetic mutations often have sporadic presentations.[4,6] As an imaging endophenotype, the ears of the lynx sign seems useful to suggest that genetic testing should be performed in patients who apparently have a sporadic rather than hereditary disorder or in patients with an atypical clinical picture. Our study determined that the ears of the lynx sign, particularly on axial FLAIR images, is highly associated with these mutations. In addition, when Whole Exome Sequencing finds genetic variants with potential pathogenicity that is nonetheless difficult to assess, the presence of the ears of the lynx sign on MR imaging would support the pathogenicity of the variants.

An important question is whether this MR imaging sign is present before the onset of clear-cut, defining clinical findings. Our study cannot answer this question. Only 1 patient was studied at age 11, when she had a low intelligence quotient but had not yet developed a spastic paraparesis, which started at 15 years of age. Already at 11 years of age, the ears of the lynx sign was present on her MR imaging. We hope that the description of the reliability of this sign may encourage its search at younger ages.

We studied mostly patients with mutations in the SPG11 gene, on chromosome 15. Of the 34 patients with HSP-TCC, only 3 had mutations in another gene, causing SPG15; their MR imaging also contained the ears of the lynx sign. Therefore, this sign is not pathognomonic for SPG11 mutations but can probably be found with mutations that affect the corpus callosum in a manner similar to the SPG11 mutations. We are aware of a number of reports mentioning the ears of the lynx sign in patients with genetically determined HSP-TCC,[7,22,27-31] but we are also aware of a single report of a somewhat similar finding in a patient with a possible Marchiafava-Bignami syndrome.[32] However, in this case, the imaging finding was difficult to interpret in the published image; furthermore, that patient had a thin corpus callosum and no genetic testing. Given the range of phenotypic presentations of these genetic disorders, an atypical HSP-TCC could not be ruled out confidently in that case. The main potential mimic of the ears of the lynx sign is an area of gliosis observed in many healthy individuals at the calloso-caudate angle of the frontal horn of the lateral ventricles (Fig 2).[21] However, this finding differs in morphology—it has a rounded capping, rather than flaming shape—and is smaller than the ears of the lynx. This normal finding, present in many of our MRIs with normal findings, probably led the raters to misclassify some cases, but this number was very small and null for some of the raters, as indicated in the Results section.

The raters were not asked to separate the series of images by diagnosis: They were completely unaware of the 3 diagnostic groups. They were only asked whether the ears of the lynx finding was present on an image. Therefore, other features of the image, such as other hyperintense lesions in multiple sclerosis scans, were supposed to be irrelevant to their readings.

Our study had a main limitation. It is difficult to rule out an effect of the presence of a thin corpus callosum on the decision of the raters. To minimize this issue, we did not reveal to the raters the disease entities to which the images corresponded. HSP-TCC is uncommon, and on postreading, only 2 of the raters indicated that they had some exposure to this disorder. One had read previously about 3 cases; the other, none. These 2 raters did not have better accuracy than the other 2. Because the images were presented in a sequence, we were able to evaluate the learning effects. Had the raters used a different feature in the image, namely the thin corpus callosum, to aid in accuracy, a learning effect would be expected. We did not detect better accuracy for images read later in the series.

CONCLUSIONS

The ears of the lynx MR imaging sign suggests the presence of a genetic mutation, likely characteristic of SPG11 or SPG15, in patients with cognitive impairment or a spastic paraparesis. In addition, the presence of this MR imaging finding in a subject identified as having a genotype associated with HSP-TCC suggests that the mutation is pathogenic.

ACKNOWLEDGMENT

Aleksandar Vasic obtained the original photo in Fig 3 and gave permission to use it.

Disclosures: Marcondes C. França Jr—*UNRELATED*: Fundação de Amparo à Pesquisa do Estado de São Paulo grant 2013/01766–7, *Comments*: The Fundação de Amparo à Pesquisa do Estado de São Paulo is a government agency in Brazil that funded genetic testing and MRI acquisitions.* Bart P. van de Warrenburg—*UNRELATED*: *Grants/Grants Pending*: Hersenstichting, ZonMw, Bioblast Pharma*; *Royalties*: Reed Elsevier*; *Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed*: International Parkinson and Movement Disorders Society. Berry Kremer—*UNRELATED*: *Board Membership*: European Huntington's Disease Network/Child Health and Development Institute, *Comments*: travel cost reimbursements; *Expert Testimony*: Veduma Medical Expertise, Zaltbommel, the Netherlands; *Comments*: various book chapters; *Stock/Stock Options:* Fresenius, Germany. Joost Nicolai—*UNRELATED*: *Payment for Lectures Including Service on Speakers Bureaus*: SEPION, the Netherlands, *Comments*: Money was paid for organizing an epilepsy course twice a year for residents; *Comments*: Fee for several chapters in neurology books (Dutch). Joseph C. Masdeu—*UNRELATED*: *Board Membership*: GE Healthcare; *Expert testimony:* National Collegiate Athletic Association; *Grants/Grants Pending*: Avanir Pharmaceuticals, Acadia, Biogen, Eli Lilly, AbbVie, Novartis, National Institutes of Health, *Comments*: clinical trials*; *Payment for Lectures Including Service on Speakers Bureaus*: Eli Lilly, *Comments*: amyloid imaging; *Royalties*: Walters Kluwer, *Comments*: localization in Neurology; *Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed*: International Collaboration with Siemens.* Michael D. Gregory—*UNRELATED*: *Employment*: National Institutes of Mental Health, Inova Medical Group. *Money paid to the institution.

REFERENCES

1. Chrestian N, Dupré N, Gan-Or Z, et al. **Clinical and genetic study of hereditary spastic paraplegia in Canada.** *Neurol Genet* 2017;3:e122 CrossRef Medline
2. Goizet C, Boukhris A, Maltete D, et al. **SPG15 is the second most common cause of hereditary spastic paraplegia with thin corpus callosum.** *Neurology* 2009;73:1111–19 CrossRef Medline
3. Hehr U, Bauer P, Winner B, et al. **Long-term course and mutational spectrum of spatacsin-linked spastic paraplegia.** *Ann Neurol* 2007; 62:656–65 CrossRef Medline
4. Kara E, Tucci A, Manzoni C, et al. **Genetic and phenotypic characterization of complex hereditary spastic paraplegia.** *Brain* 2016;139: 1904–18 CrossRef Medline
5. Pensato V, Castellotti B, Gellera C, et al. **Overlapping phenotypes in complex spastic paraplegias SPG11, SPG15, SPG35 and SPG48.** *Brain* 2014;137:1907–20 CrossRef Medline
6. Stevanin G, Azzedine H, Denora P, et al; SPATAX consortium. **Mutations in SPG11 are frequent in autosomal recessive spastic paraplegia with thin corpus callosum, cognitive decline and lower motor neuron degeneration.** *Brain* 2008;131:772–84 CrossRef Medline
7. Estrada-Cuzcano A, Martin S, Chamova T, et al. **Loss-of-function mutations in the ATP13A2/PARK9 gene cause complicated hereditary spastic paraplegia (SPG78).** *Brain* 2017;140:287–305 CrossRef Medline
8. Eichler F, Ratai E, Carroll JJ, et al. **Inherited or acquired metabolic disorders.** *Handb Clin Neurol* 2016;135:603–36 CrossRef Medline
9. Abdel-Hamid MS, Issa MY, Otaify GA, et al. **PGAP3-related hyperphosphatasia with mental retardation syndrome: report of 10 new patients and a homozygous founder mutation.** *Clin Genet* 2018;93: 84–91 CrossRef Medline
10. Barmherzig R, Bullivant G, Cordeiro D, et al. **A new patient with intermediate severe Salla disease with hypomyelination: a literature review for Salla disease.** *Pediatr Neurol* 2017;74:87–91.e82 CrossRef Medline
11. Depienne C, Nava C, Keren B, et al. **Genetic and phenotypic dissection of 1q43q44 microdeletion syndrome and neurodevelopmental phenotypes associated with mutations in ZBTB18 and HNRNPU.** *Hum Genet* 2017;136:463–79 CrossRef Medline
12. Flex E, Niceta M, Cecchetti S, et al. **Biallelic mutations in TBCD, encoding the tubulin folding cofactor D, perturb microtubule dynamics and cause early-onset encephalopathy.** *Am J Hum Genet* 2016;99:962–73 CrossRef Medline
13. Kurata H, Terashima H, Nakashima M, et al. **Characterization of SPATA5-related encephalopathy in early childhood.** *Clin Genet* 2016;90:437–44 CrossRef Medline
14. Lamers IJC, Reijnders MR, Venselaar H, et al. **Recurrent de novo mutations disturbing the GTP/GDP binding pocket of RAB11B cause intellectual disability and a distinctive brain phenotype.** *Am J Hum Genet* 2017;101:824–32 CrossRef Medline
15. Meng L, Donti T, Xia F, et al. **Homozygous variants in pyrroline-5-carboxylate reductase 2 (PYCR2) in patients with progressive microcephaly and hypomyelinating leukodystrophy.** *Am J Med Genet A* 2017;173:460–70 CrossRef Medline
16. Rosti RO, Dikoglu E, Zaki MS, et al. **Extending the mutation spectrum for Galloway-Mowat syndrome to include homozygous missense mutations in the WDR73 gene.** *Am J Med Genet A* 2016;170A: 992–98 CrossRef Medline
17. Sidhu M, Brady L, Tarnopolsky M, et al. **Clinical manifestations associated with the N-Terminal-Acetyltransferase NAA10 gene mutation in a girl: Ogden syndrome.** *Pediatr Neurol* 2017;76:82–85 CrossRef Medline
18. Kraus MF, Susmaras T, Caughlin BP, et al. **White matter integrity and cognition in chronic traumatic brain injury: a diffusion tensor imaging study.** *Brain* 2007;130:2508–19 CrossRef Medline
19. Mori S, Wakana S, van Zijl PC, et al. *MRI Atlas of Human White Matter*. Amsterdam: Elsevier; 2005, see: https://www.elsevier.com/books/mri-atlas-of-human-white-matter/mori/978-0-444-51741-8
20. Riverol M, Samaranch L, Pascual B, et al. **Forceps minor region signal abnormality "ears of the lynx": an early MRI finding in spastic paraparesis with thin corpus callosum and mutations in the spatacsin gene (SPG11) on chromosome 15.** *J Neuroimaging* 2009;19:52–60 CrossRef Medline
21. Sze G, De Armond SJ, Brant-Zawadzki M, et al. **Foci of MRI signal (pseudo lesions) anterior to the frontal horns: histologic correlations of a normal finding.** *AJR Am J Roentgenol* 1986;147:331–37 CrossRef Medline
22. de Bot ST, Burggraaff RC, Herkert JC, et al. **Rapidly deteriorating course in Dutch hereditary spastic paraplegia type 11 patients.** *Eur J Hum Genet* 2013;21:1312–15 CrossRef Medline
23. França MC Jr, D'Abreu A, Maurer-Morelli CV, et al. **Prospective neuroimaging study in hereditary spastic paraplegia with thin corpus callosum.** *Mov Disord* 2007;22:1556–62 CrossRef Medline
24. Hertzog MA. **Considerations in determining sample size for pilot studies.** *Res Nurs Health* 2008;31:180–91 CrossRef Medline
25. Landis JR, Koch GG. **The measurement of observer agreement for categorical data.** *Biometrics* 1977;33:159–74 CrossRef Medline
26. Stevanin G, Santorelli FM, Azzedine H, et al. **Mutations in SPG11, encoding spatacsin, are a major cause of spastic paraplegia with thin corpus callosum.** *Nat Genet* 2007;39:366–72 CrossRef Medline
27. Faber I, Servelhere KR, Martinez ARM, et al. **Clinical features and management of hereditary spastic paraplegia.** *Arq Neuropsiquiatr* 2014;72:219–26 CrossRef Medline
28. Renvoisé B, Chang J, Singh R, et al. **Lysosomal abnormalities in hereditary spastic paraplegia types SPG15 and SPG11.** *Ann Clin Transl Neurol* 2014;1:379–89 CrossRef Medline
29. Giannoccaro MP, Liguori R, Arnoldi A, et al. **Atypical late-onset hereditary spastic paraplegia with thin corpus callosum due to novel compound heterozygous mutations in the SPG11 gene.** *J Neurol* 2014;261:1825–27 CrossRef Medline
30. Fraidakis MJ, Brunetti M, Blackstone C, et al. **Novel compound heterozygous spatacsin mutations in a Greek kindred with hereditary spastic paraplegia SPG11 and dementia.** *Neurodegener Dis* 2016; 16:373–81 CrossRef Medline
31. Chakrabarty S, Vijayakumar N, Radhakrishnan K, et al. **Spastizin mutation in hereditary spastic paraplegia with thin corpus callosum.** *J Neurol* 2016;263:2130–32 CrossRef Medline
32. Pacheco FT, Rego MM, do Rego JI, et al. "**Ears of the lynx" sign in a Marchiafava-Bignami patient: structural basis and fiber-tracking DTI contribution to the understanding of this imaging abnormality.** *J Neuroimaging* 2014;24:205–07 CrossRef Medline

Celebrating 35 Years of the AJNR

January 1984 edition

Cerebral NMR Imaging: Early Results with a 0.12 T Resistive System

Robert A. Zimmerman[1]
Larissa T. Bilaniuk[1]
Herbert I. Goldberg[1]
Robert I. Grossman[1]
Richard S. Levine[1]
Roberta Lynch[1]
William Edelstein[2]
Paul Bottomley[2]
Rowland Redington[2]

Over a 6-month period, 157 patients, 85 of whom had central nervous system tumors, were examined on a prototype 0.12 T resistive nuclear magnetic resonance (NMR) imaging unit. All of the patients had computed tomography (CT), which was used as a standard to which the NMR findings were compared. Studies were done primarily by saturation-recovery technique with short repetition times. The signal intensity with saturation-recovery technique did not allow differentiation among most tumor types. Location, extent, and morphology helped to some extent in attempts at differentiation. In the multiplanar mode, NMR compared favorably to CT with regard to lesion detection. Limited early experience suggests that NMR also may detect some lesions when the CT is negative and may detect additional lesions when one or more are present. The NMR examination was well tolerated by selected patients.

In November 1982, a General Electric (GE) 0.12 T (5.1 MHz), developmental, resistive nuclear magnetic resonance (NMR) proton imaging system was installed at the Hospital of the University of Pennsylvania. This allowed the study of 157 patients for disease processes involving the brain, face, or upper cervical region over the next 6 months. During that time, software improvements were made, resulting in ability to obtain more sections in less time and simultaneous sectioning of a given volume of interest in coronal and sagittal planes, in addition to the transverse plane. We present the initial experience with this unit.

Subjects and Methods

Between November 1982 and May 1983, 11 volunteers and 157 selected patients were examined by NMR in order to evaluate the brain in most cases and in several instances the face, neck, and upper cervical spinal cord. Thirty-one patients were age 20 or younger. The

This article appears in the December 1983 issue of AJR and the January/February 1984 issue of AJNR.

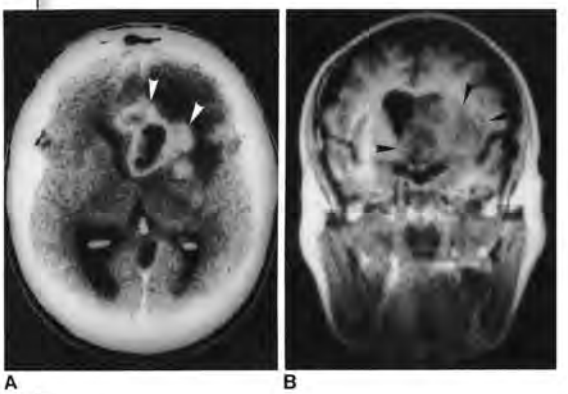
Received August 4, 1983; accepted August 19, 1983.

Presented in part at the annual meeting of the American Roentgen Ray Society, Atlanta, April 1983.

[1]Department of Radiology, Hospital of the University of Pennsylvania, 3400 Spruce St., Philadelphia, PA 19104. Address reprint requests to R. A. Zimmerman.

[2]General Electric Research and Development Center, Schenectady, NY 12301

AJNR 5:1–7, January/February 1984
0195–6108/84/0501–0001 $00.00
© American Roentgen Ray Society



A B

Magnetic Resonance Imaging of the Cervical Spine: Technical and Clinical Observations

Michael T. Modic[1]
Meredith A. Weinstein[1]
William Pavlicek[1]
Francis Bourmphrey[2]
Daniel Starnes[1]
Paul M. Duchesneau[1]

Seventy-two patients were examined to determine the clinical potential for magnetic resonance imaging (MRI) of the spine. MRI using different pulse sequences was compared with plain radiography, high-resolution computed tomography, and myelography. There were 35 normal patients; pathologic conditions studied included canal stenosis, herniated disk, metastatic tumor, neurofibroma, trauma, Chiari malformation, syringomyelia, arteriovenous malformation, and rheumatoid arthritis. MRI provided sharply defined anatomic delineation and tissue characterization. It was diagnostic in syringomyelia and Chiari malformation and was useful in the evaluation of trauma and spinal canal block from any cause. MRI was sensitive to degenerative disk disease and infection. The spin-echo technique, with three pulse sequence variations, seems very promising. A short echo time (TE) produces the best signal-to-noise ratio and spatial resolution. Lengthening the TE enhances differentiation of various tissues by their signal intensity, while the combined increase of TE and recovery time (TR) produces selective enhancement of the cerebrospinal fluid signal intensity.

Early clinical experience with magnetic resonance imaging (MRI) has demonstrated its potential in the evaluation of the spine and foramen magnum [1–5]. With appropriate pulse-sequence technique, the spinal cord, brainstem, cerebrospinal fluid (CSF), and extradural structures such as the intervertebral disk have been depicted without the use of intrathecal contrast or ionizing radiation. We evaluated different magnetic resonance pulse sequences and compared the images with myelograms, CT scans, and plain radiographs of the cervical spine.

Subjects and Methods

Seventy-two patients formed the subject group for this study. Thirty-seven patients were studied for known or suspected disease of the cervical spine and foramen magnum. Thirty-five patients were studied as normal controls.

This article appears in the December 1983 issue of AJR and the January/February 1984 issue of AJNR.

Received July 19, 1983; accepted August 1, 1983.

Presented in part at the annual meeting of the American Roentgen Ray Society, Atlanta, April 1983.

[1]Department of Radiology, Cleveland Clinic Foundation, 9500 Euclid Ave., Cleveland, OH 44106. Address reprint requests to M. T. Modic.

[2]Department of Orthopedic Surgery, Cleveland Clinic Foundation, Cleveland, OH 44106

AJNR 5:15–22, January/February 1984
0195–6108/84/0501–0015 $00.00
© American Roentgen Ray Society

A B C

T1-Hyperintense Plaques on Intracranial-versus-Extracranial Vessel Wall MRI

We read with great interest the recent article by Zhu et al[1] regarding the association between T1-hyperintensity of basilar artery stenosis as detected by intracranial vessel wall MR imaging, presumably intraplaque hemorrhage (IPH), and cerebral infarction in the basilar artery territory. T1-hyperintense plaques as a marker of IPH have been validated in the extracranial carotid artery using carotid endarterectomy specimens and have been studied extensively in various populations. In contrast, data on similar lesions in intracranial arteries are rather limited. The study by Zhu et al highlighted the clinical relevance of T1-hyperintensity in intracranial atherosclerotic disease, which should call for further investigations of its etiology and pathophysiology, especially considering the similarities and differences compared with T1-hyperintensity in carotid artery disease.

In the study by Zhu et al[1] and virtually all previous studies on intracranial IPH,[2,3] imaging criteria used for detecting IPH were adopted from carotid MR imaging studies. Radiologic-pathologic correlation studies validating the empiric signal intensity threshold adopted from carotid MR imaging for detecting intracranial IPH are currently lacking. This issue is understandable because it is usually difficult to assess intracranial artery specimens soon after in vivo imaging. However, caution must be exercised because images are generated by different T1-weighted sequences that may produce different T1 contrasts. Although not discussed by Zhu et al, the use of different sequences with different T1 contrasts may be another potential explanation for the vast differences in the prevalence of IPH in previous studies.[1,3] In some previous studies, it is unclear whether fat suppression was implemented. In theory, plaques with a high content of lipids may also appear hyperintense on T1-weighted MR imaging. Studies developing and validating imaging criteria for detecting intracranial IPH are warranted.

Notably, the prevalence of IPH in the basilar artery was not different between the low-grade and high-grade stenosis groups,[1] whereas the prevalence of IPH in the carotid artery often increased with luminal stenosis.[4] One may explain this finding by a selection bias because the low-grade stenosis group may have included more symptomatic patients. However, the degree of stenosis was not associated with symptom status in the study.[1] Alternatively, this difference between the carotid and basilar arteries may indicate different remodeling mechanisms that exist between different arterial beds. Whether wall area measurements were different between basilar plaques with and without IPH is unknown but a question worth asking.

We congratulate Zhu et al[1] for reporting these important data and findings on a relatively unexplored topic. Notably, the association of T1-hyperintensity with cerebral infarction appeared to be even stronger than that of postcontrast enhancement, though most previous studies on intracranial vessel wall MR imaging have only focused on the latter. Certainly, more questions can be raised than answered from the study by Zhu et al, and future efforts are needed to further understand the etiology and clinical significance of T1-hyperintensity on intracranial vessel wall imaging.

Disclosures: Jie Sun—*UNRELATED*: *Grants/Grants Pending*: American Heart Association, *Comments*: I am the principle investigator and have received a grant from the American Heart Association to conduct imaging research on carotid plaque. **Money paid to the institution of Washington.

REFERENCES

1. Zhu C, Tian X, Degnan AJ, et al. **Clinical significance of intraplaque hemorrhage in low- and high-grade basilar artery stenosis on high-resolution MRI.** *AJNR Am J Neuroradiol* 2018;39:1286–92 CrossRef Medline
2. Xu WH, Li ML, Gao S, et al. **Middle cerebral artery intraplaque hemorrhage: prevalence and clinical relevance.** *Ann Neurol* 2012;71:195–98 CrossRef Medline
3. Yu JH, Kwak HS, Chung GH, et al. **Association of intraplaque hemorrhage and acute infarction in patients with basilar artery plaque.** *Stroke* 2015;46:2768–72 CrossRef Medline
4. Zhao X, Underhill HR, Zhao Q, et al. **Discriminating carotid atherosclerotic lesion severity by luminal stenosis and plaque burden: a comparison utilizing high-resolution magnetic resonance imaging at 3.0 Tesla.** *Stroke* 2011;42:347–53 CrossRef Medline

W. Yuan
Department of Radiology
Navy Qingdao No. 1 Sanatorium of People's Liberation Army
Qingdao, Shandong, China
J. Sun
Department of Radiology
University of Washington
Seattle, Washington

http://dx.doi.org/10.3174/ajnr.A5878

REPLY:

We thank Drs Yuan and Sun for their interest in our article[1] and their group's recent work on intraplaque hemorrhage (IPH) in carotid plaque.[2] While IPH is validated as a high-risk plaque feature in extracranial carotid artery plaque[3] due to facile histologic validation from endarterectomy specimens, our current understanding of the role of IPH in intracranial atherosclerotic plaque vulnerability assessment is limited. The characterization of intracranial plaque features by in vivo imaging remains challenging due to the small plaque size and impracticality of histologic validation. The definition of T1-weighted hyperintense signal suggestive of IPH as >150% signal relative to the nearby medial pterygoid muscles on precontrast T1-weighted imaging used in our study was adopted from previous carotid studies and supported by initial studies of IPH within the middle cerebral artery.[4]

Intracranial IPH is increasingly recognized as a harbinger of elevated stroke risk,[5] and our study adds to the literature concerning basilar artery plaque in showing an association with stroke risk independent of the degree of stenosis. We appreciate Drs Yuan and Sun's suggestions on the discussion of imaging sequences for improving the accuracy of intracranial IPH detection. The use of different TR/TE parameters or inversion pulses can alter image contrast, which may partially explain the heterogeneity of prior study results. Histologic validation is preferred to standardize imaging approaches; however, it has practical challenges in intracranial atherosclerotic plaque, which cannot be obtained in vivo. An alternative practical approach to standardize sequences for clinical detection of IPH could be to use phantoms with different T1 values; this method needs to be explored further. In addition, recent studies have increasingly used 3D high-resolution fast spin-echo sequences (sampling perfection with application-optimized contrasts by using different flip angle evolution [SPACE] sequence, Siemens, Erlangen, Germany; CUBE, GE Healthcare, Milwaukee, Wisconsin; or volume isotropic turbo spin-echo acquisition [VISTA], Philips Healthcare, Best, the Netherlands) for intracranial plaque imaging. IPH detection should be used with caution because the long echo-train (>30) induces considerable T2-weighting,[6] which alters contrast from traditional 2D T1-weighted fast spin-echo or gradient-echo sequences.

In our study, the degree of stenosis was not associated with stroke symptoms, and the prevalence of IPH was comparable in low- and high-grade stenoses. This finding may reflect IPH as an independent risk factor for stroke symptoms in basilar artery plaque, though the limited sample size in a single-center population could introduce sampling error. Nonetheless, these data highlight the importance of basilar artery IPH even in low-grade stenosis as potentially leading to stroke. High-risk plaque in intracranial arteries with <50% stenosis[7] has been recognized as a potential cause of cryptogenic stroke, and identifying these high-risk features, including IPH, may improve the management of these patients. Nonetheless, we agree with Drs Yuan and Sun that both IPH and plaque enhancement should be studied further in a prospective, longitudinal study to better understand potential imaging-related independent risk factors that may predict stroke.

http://dx.doi.org/10.3174/ajnr.A5895

REFERENCES

1. Zhu C, Tian X, Degnan AJ, et al. **Clinical significance of intraplaque hemorrhage in low- and high-grade basilar artery stenosis on high-resolution MRI.** *AJNR Am J Neuroradiol* 2018;39:1286–92 CrossRef Medline
2. Wang X, Sun J, Zhao X, et al; CARE-II study investigators. **Ipsilateral plaques display higher T1 signals than contralateral plaques in recently symptomatic patients with bilateral carotid intraplaque hemorrhage.** *Atherosclerosis* 2017;257:78–85 CrossRef Medline
3. Saam T, Hetterich H, Hoffmann V, et al. **Meta-analysis and systematic review of the predictive value of carotid plaque hemorrhage on cerebrovascular events by magnetic resonance imaging.** *J Am Coll Cardiol* 2013;62:1081–91 CrossRef Medline
4. Turan TN, Bonilha L, Morgan PS, et al. **Intraplaque hemorrhage in symptomatic intracranial atherosclerotic disease.** *J Neuroimaging* 2011;21:e159–61 CrossRef Medline
5. Yu JH, Kwak HS, Chung GH, et al. **Association of intraplaque hemorrhage and acute infarction in patients with basilar artery plaque.** *Stroke* 2015;46:2768–72 CrossRef Medline
6. Zhu C, Haraldsson H, Tian B, et al. **High resolution imaging of the intracranial vessel wall at 3 and 7 T using 3D fast spin echo MRI.** *MAGMA* 2016;29:559–70 CrossRef Medline
7. Cho HJ, Kim KH, Kim EJ, et al. **Clinical implications of basilar artery plaques in the pontine infarction with normal basilar angiogram: a high-resolution magnetic resonance imaging study.** *J Stroke Cerebrovasc Dis* 2018 Sep 12. [Epub ahead of print] CrossRef Medline

C. Zhu
Department of Radiology and Biomedical Imaging
University of California, San Francisco
San Francisco, California
X. Tian
Department of Radiology
Changhai Hospital
Shanghai, China
A.J. Degnan
Department of Radiology
Children's Hospital of Philadelphia
Philadelphia, Pennsylvania
J. Lu
Q. Liu
Department of Radiology
Changhai Hospital
Shanghai, China

Disproportionate International Contributions to Subspecialties of Neuroradiology in the *American Journal of Neuroradiology*

In a follow-up study to our article entitled, "Who's Contributing Most to American Neuroscience Journals: American or Foreign Authors?"[1] we looked at the various branches of neuroradiology to identify the countries that contribute the most published articles to the *American Journal of Neuroradiology* (*AJNR*). We had shown that contributions to the *AJNR* by foreign institutions have dramatically increased in recent decades.[1] We subsequently examined whether all areas of neuroradiology have been equally affected in the *AJNR*.

For this analysis, we assessed the country of the first author's institution in published articles in the past 30 *AJNR* issues (January 2016 to June 2018) to determine the contributions to Brain, Head & Neck (H&N), Spine, Neurointerventional Radiology (NIR), and Pediatrics (Ped) sections. In addition to calculating the percentage of US authorship, we determined which countries were contributing the most to each branch of neuroradiology.

We reviewed 793 articles. Overall, 350 (44.1%) articles were from American institutions, and 443 (55.9%) were from non-American institutions. The percentages of articles from US institutions were 44.4% in 2016, 47.2% in 2017, and 37.4% from January to June 2018. The percentage of US authorship was least in NIR (31.8%; $P < .001$) and Brain (41.5%; $P = .21$), while it was significantly higher than non-US authorship in Spine (60.7%; $P = .002$), followed by H&N (53.7%), and Ped (51.3%) (Figure).

The foreign countries that had the highest contributions in Brain, H&N, Spine, Ped, and NIR were Japan, Korea, Canada, Italy, and Germany, respectively (Table).

We concluded that contributions to the *AJNR* from non-US authors dominated in the NIR category, likely due to the more restrictive limitations of the FDA on new NIR devices compared with the more lenient oversight in Europe by their governmental bodies, such as the Medicines and Healthcare products Regulatory Agency in the United Kingdom and the French Organization for the Safety of Health Products. Asian countries may contribute more to H&N due to the increased prevalence of thyroid and nasopharyngeal carcinomas and Epstein–Barr Virus infections. Europeans and Canadians publish more than those in Asian countries in Pediatrics, in part from noteworthy Italian authorities in pediatric neuroradiology. The US prominence is foremost in the spine, and Canada dominates the non-US spine contributions. Overall, the contributions from non-US authors (55.9%) exceeded those originating from the United States (44.1%).

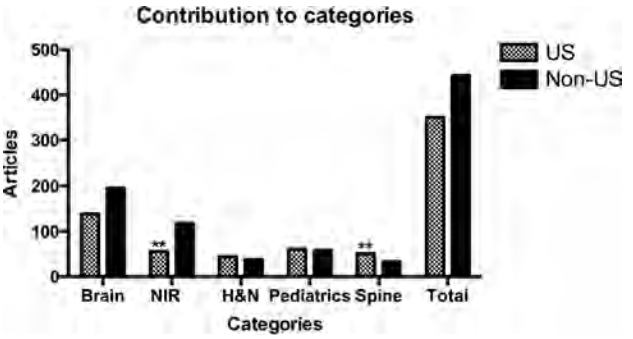


FIGURE. Contribution of US and Non-US countries to different subsections of the *AJNR*. Values are expressed as number of articles and were analyzed using the χ^2 independence test. *Double asterisks* show $P < .01$ compared with the other groups.

Disclosures: David M. Yousem—*UNRELATED*: expert testimony: medicolegal work; *Payment for Lectures Including Service on Speakers Bureaus:* American College of Radiology Education Center; *Royalties*: Elsevier, *Comments*: 5 books; *Travel/Accommodations/Meeting Expenses Unrelated to Activities Listed*: Radiological Society of North American Educator Career Award 2018.

REFERENCE

1. Charkhchi P, Mirbolouk M, Jalilian R, et al. **Who's contributing most to American neuroscience journals: American or foreign authors**? *AJNR Am J Neuroradiol* 2018;39:1001–07 CrossRef Medline

S. Emamzadehfard
Department of Radiology
V. Eslami
Department of Neurology
University of Texas Health Science Center
Houston, Texas
D.M. Yousem
S. Sahraian
Division of Neuroradiology
The Russell H. Morgan Department of Radiology and Radiologic Science
Johns Hopkins Medical Institution
Baltimore, Maryland

http://dx.doi.org/10.3174/ajnr.A5885

Top 2 foreign countries with highest contributions to subsections of *AJNR*

Specialty	Top Country Contributor	Top Country (No. of Articles) (% of Non-US)	No. 2 Country Contributor	No. 2 Country (No. of Articles) (% of Non-US)
Brain	Japan	28 (14.3%)	China	24 (12.2%)
H&N	Korea	15 (39.5%)	China	6 (15.8%)
NIR	Germany	22 (18.6%)	France	19 (16.1%)
Pediatrics	Italy	12 (20.6%)	Canada	8 (13.8%)
Spine	Canada	10 (30.3%)	Switzerland	4 (12.1%)

The authors regret that in the article "White Matter Changes Related to Subconcussive Impact Frequency during a Single Season of High School Football" (*AJNR Am J Neuroradiol* 2018;39:245–51; https://doi.org/10.3174/ajnr.A5489), the art and legend for Fig 3 did not match. A corrected figure with the original legend is reproduced below.

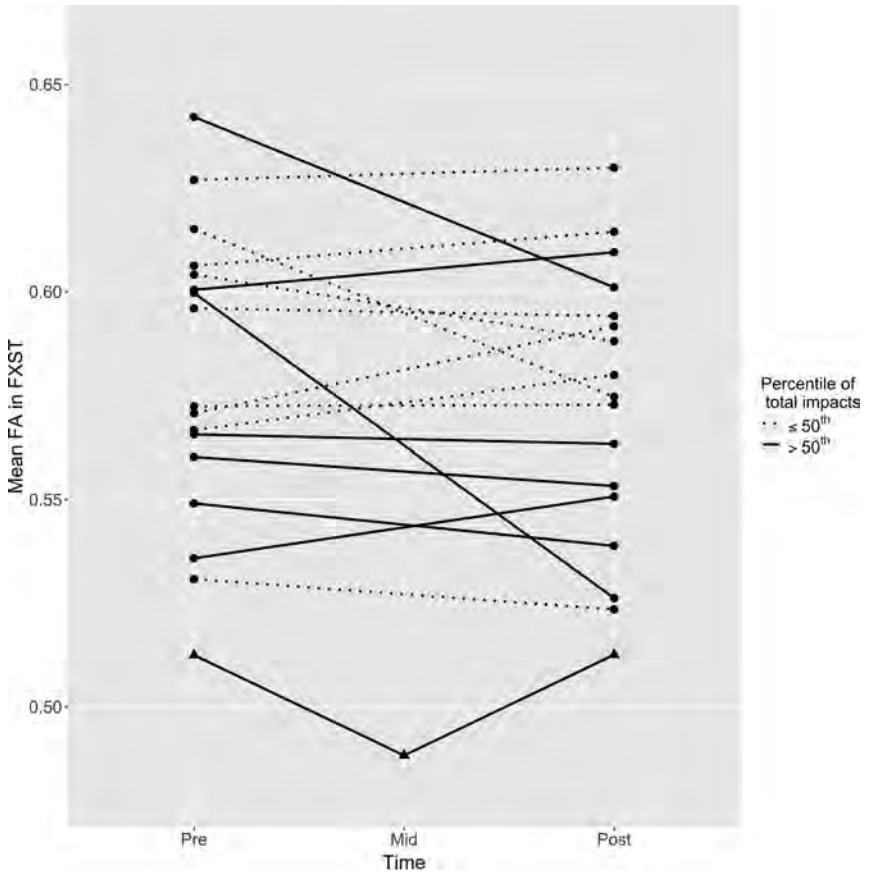


FIG 3. Scatterplot of changes in FXST FA for all subjects across the season. Individual player data are connected by *solid* or *dotted lines* based on the median split of the total number of impacts during the season, as depicted in the legend. The concussed subject is plotted with *triangles*.

http://dx.doi.org/10.3174/ajnr.A5907