Imaging Regional Metabolic Changes in the Ischemic Rat Heart In Vivo Using Hyperpolarized [1-13C]Pyruvate

Lauritzen, Mette Hauge; Magnusson, Peter; Laustsen, Christoffer; Asghar Butt, Sadia; Ardenkjær-Larsen, Jan Henrik; Søgaard, Lise Vejby; Paulson, Olaf B.; Åkeson, Per

Published in: International Journal of Tomography & Statistics

DOI: 10.18383/j.tom.2017.00008

Publication date: 2017

Document version: Publisher's PDF, also known as Version of record

Document license: CC BY-NC-ND

Citation for published version (APA):
Deformation mapping of aortic aneurysm
It’s simple: You need to glean more information from every *in vivo* experiment, giving you maximum data to enhance and broaden your biological understanding. And that’s what’s behind our multimodality imaging solutions. Coregistration is enabled through a combination of comprehensive optical, PET, and microCT platforms, plus a unique portfolio of bioluminescent cell lines and fluorescent agents. Together, these technologies offer a more complete visual assessment and understanding of disease treatment and progression over time, along with improved outcomes and treatment efficacy. Because when modalities combine, you see disease from *every* angle.

For research use only. Not for use in diagnostic procedures.

Visit our booth at AACR (#1631) to learn more.  
www.perkinelmer.com/Multimodality-Imaging
Tomography® is published quarterly. Tomography publishes basic (technical and pre-clinical) and clinical scientific articles which involve the advancement of imaging technologies. Tomography encompasses studies that use single or multiple imaging modalities including for example CT, US, PET, SPECT, MR and hyperpolarization technologies, as well as optical modalities (i.e. bioluminescence, photoacoustic, endomicroscopy, fiber optic imaging and optical computed tomography) in basic sciences, engineering, preclinical and clinical medicine. Tomography also welcomes studies involving exploration and refinement of contrast mechanisms and imagederived metrics within and across modalities toward the development of novel imaging probes for image-based feedback and intervention. The use of imaging in biology and medicine provides unparalleled opportunities to noninvasively interrogate tissues to obtain real-time dynamic and quantitative information required for diagnosis and response to interventions and to follow evolving pathological conditions. As multi-modal studies and the complexities of imaging technologies themselves are ever increasing to provide advanced information to scientists and clinicians, Tomography provides a unique publication venue allowing investigators the opportunity to more precisely communicate integrated findings related to the diverse and heterogeneous features associated with underlying anatomical, physiological, functional, metabolic and molecular genetic activities of normal and diseased tissue. Thus Tomography publishes peer-reviewed articles which involve the broad use of imaging of any tissue and disease type including both preclinical and clinical investigations. In addition, hardware/software along with chemical and molecular probe advances are welcome as they are deemed to significantly contribute towards the long-term goal of improving the overall impact of imaging on scientific and clinical discovery. Tomography provides a comprehensive venue for integration of imaging modalities to address biologically important and clinically relevant questions in order to facilitate more rapid and seamless scientific and clinical advancements. Analysis of simultaneous multidimensional (e.g. content, space and time) and multivariate computational data extraction of features which support unique identification of different pathologic tissue phenotypes for advancing imaging is welcomed by Tomography. Types of articles include regular Research Articles, Brief Reports, Reviews, Image Reports and Consensus Papers sponsored from national meetings or government funding agencies.

This journal is available online and can be accessed at www.tomography.org. The journal is an Open Access journal and can be freely downloaded for personal use.

Publication Information: Tomography (ISSN: 2379-1381) is published quarterly by Grapho Publications, 3000 Green Road, #131281, Ann Arbor, MI 48105, USA. Periodicals postage paid at Ann Arbor, MI and additional mailing offices.

USA POSTMASTER: Send address changes to Tomography, Grapho Publications, Customer Service Department, 3000 Green Road, #131281, Ann Arbor, MI 48105, USA.

Advertising Information: Advertising and classified advertising orders and inquiries can be sent to: e-mail: info@tomography.org.

Guide for Authors: For a full and complete Guide for Authors, please go to: http://www.tomography.org.

Author Inquiries: For inquiries relating to the submission of articles (including electronic submission) please visit this journal’s homepage at http://www.tomography.org. Contact details for questions arising after acceptance of an article, including those relating to proofs, will be provided by the publisher.

The paper meets the requirements of ANSI/NISO Z39.48-1992 [Permanence of Paper].

Reprints: To order reprints for educational, commercial, or promotional use, contact Cadmus Reprints, P.O. Box 822942, Philadelphia, PA 19182-2942; E-mail: csjreprints@cadmus.com; Phone: 866-487-5625; Fax: 877-705-1373

Funding Body Agreements and Policies: Grapho Publications has established agreements and policies to allow authors whose articles appear in journals published by Grapho Publications, to comply with potential manuscript archiving requirements as specified as conditions of their grant awards. Tomography is an open access journal which complies with NIH funding guidelines and all articles can be archived by authors in PubMed Central http://www.nihms.nih.gov or other international data bases to be in compliance with granting agency requirements.

© Grapho Publications LLC. All rights reserved.

This journal and the individual contributions contained in it are protected under copyright by Grapho Publications, LLC, and the following terms and conditions apply to their use:

Permitted reuse of articles is determined by the Creative Commons Attribution - NonCommercial - NoDerivs (CC BY-NC-ND) license, which lets others distribute and copy the article for non-commercial purposes, and include it in a collective work (such as an anthology), as long as they credit the author(s) and provided they do not alter or modify the article.

Notice: No responsibility is assumed by the Publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made. Although all advertising material is expected to conform to ethical (medical) standards, inclusion in this publication does not constitute a guarantee or endorsement of the quality or value of such product or of the claims made of it by its manufacturer.

Orders, Claims, and Journal Inquiries: Please contact the Grapho Publications Customer Service Department, Grapho Publications, 3000 Green Road, #131281, Ann Arbor, MI 48105, USA; phone: (+1) 734-221-0126 e-mail: info@tomography.org.
Research Articles

Imaging Regional Metabolic Changes in the Ischemic Rat Heart In Vivo Using Hyperpolarized [1-13C]Pyruvate
Mette Hauge Lauritzen, Peter Magnusson, Christoffer Laustsen, Sadia Asghar Butt, Jan Henrik Ardenkjær-Larsen, Lise Vejby Segaard, Olaf B. Paulson, and Per Åkeson

Monitoring Radiation Treatment Effects in Glioblastoma: FLAIR Volume as Significant Predictor of Survival
Matthew D. Garrett, Ted K. Yanagihara, Randy Yeh, Guy M. McKhann, Michael B. Sisti, Jeffrey N. Bruce, Sameer A. Sheth, Adam M. Sonabend, and Tony J. C. Wang

Correcting Nonpathological Variation in Longitudinal Parametric Response Maps of CT Scans in COPD Subjects: SPIROMICS
Antonio Fernández-Baldera, Charles R. Hatt, Susan Murray, Eric A. Hoffman, Ella A. Kazerooni, Fernando J. Martinez, MeiLan K. Han, and Craig J. Galbán

Can Hyperpolarized 13C-Urea be Used to Assess Glomerular Filtration Rate? A Retrospective Study
Christian Østergaard Mariager, Per Mose Nielsen, Haiyun Qi, Marie Schroeder, Lotte Bonde Bertelsen, and Christoffer Laustsen

High-Resolution MR Imaging of Muscular Fat Fraction—Comparison of Three T2-Based Methods and Chemical Shift-Encoded Imaging
Lena Trinh, Emelie Lind, Pernilla Peterson, Jonas Svensson, Lars E. Olsson, and Sven Månsson

Vascular Deformation Mapping (VDM) of Thoracic Aortic Enlargement in Aneurysmal Disease and Dissection
Nicholas S. Burris, Benjamin A. Hoff, Ella A. Kazerooni, and Brian D. Ross

Cover Illustration
Incidence of thoracic aortic aneurysm is increasing with approximately 3% of individuals older than 55 years of age. Imaging surveillance plays a central role in the management of asymptomatic patients with aortic disease. In the paper by Burris et al., Vascular Deformation Mapping (VDM) is presented as a novel imaging biomarker for quantification of change in aortic dimensions.
The Analysis & Collaboration Platform for Data-Intensive Research

Make your data & algorithms accessible, shareable, & reproducible.

**INCREASED UTILIZATION**
Easy to implement on-premise or in the cloud. Easy to use for all researchers.

**FAST LANE TO DISCOVERY**
Free up your time to do research, not IT.

**FUNDED TO PUBLISHED**
Simplify your processes from receiving funding to publishing your reproducible data.

Request a demo at flywheel.io

@Flywheel_io
We evaluated the use of hyperpolarized 13C magnetic resonance imaging (MRI) in an open-chest rat model of myocardial infarction to image regional changes in myocardial metabolism. In total, 10 rats were examined before and after 30 minutes of occlusion of the left anterior descending coronary artery using hyperpolarized [1-13C]pyruvate. Cardiac metabolic images of [1-13C]pyruvate and its metabolites [1-13C]lactate, [1-13C]alanine, and [13C]bicarbonate were obtained before and after ischemia. Significant reduction in the [1-13C]alanine and [1-13C]lactate signals were observed in the ischemic region post ischemia. The severity of the ischemic insult was verified by increased blood levels of troponin I and by using late contrast-enhanced MRI that showed enhanced signal in the ischemic region. This study shows that hyperpolarized MRI can be used to image regional metabolic changes in the in vivo rat heart in an open-chest model of ischemia reperfusion. Hyperpolarized MRI enables new possibilities for evaluating changes in cardiac metabolism noninvasively and in real time, which potentially could be used for research to evaluate new treatments and metabolic interventions for myocardial ischemia and to apply knowledge to future application of the technique in humans.
anhydride. The production of [13C]bicarbonate reflects the mitochondrial status, whereas the production of [1-13C]lactate and [1-13C]alanine reflects the general metabolic state of the myocytes, according to the location and activity of the enzymes in the cell. The metabolic consequences of myocardial ischemia have previously been shown in animal models, primarily in isolated perfused rat hearts (2, 7, 9) or in pigs (3, 6, 10). Recently, the technique has been applied in the in vivo rat heart in a complicated closed survival model of acute myocardial ischemia (11). In the present paper, we show an alternative approach to image the metabolic effects of regional myocardial ischemia in the in vivo rat heart, and demonstrate how challenges connected with its use in the small rapidly beating heart can be circumvented.

METHODS

Animals

All experiments were approved by The Danish Animal Experiments Inspectorate (License number: 2007/561-1350). In total, 12 male Sprague–Dawley rats (Taconic Europe, Denmark) weighing 250–350 g were examined in this study; 10 were 12 male Sprague–Dawley rats (Taconic Europe, Denmark) were sham-operated and used only for evaluation of troponin I blood levels. All rats were given water and standard rat chow ad libitum. The preischemic scan was performed 1 or 2 days before the ischemic insult was performed to reduce the time the rats were anesthetized, which potentially could affect our results. The preischemic scan was performed as described below but without the surgery. Rats were anesthetized with 4% isoflurane. The rats were intubated and connected to a small animal ventilator (SAR-830/P, IITC Life Science, CA). The ventilator was supplied with 1.75 L/min atmospheric air with additional 0.25 L/min oxygen mixed with 1.6%–2% isoflurane to maintain light anesthesia. Respiration was kept at 72 breaths/min. pCO2 was monitored on a NPB-75MAX Capnograph (Nellcor Puritan Bennett Inc, USA) connected to the ventilator. A catheter was introduced into the tail vein for intravenous administration of the hyperpolarized [1-13C]pyruvate solution. Another catheter was introduced in the left femoral artery for blood collection. During scanning, the animals were placed on a heating pad, and temperature, electrocardiogram (ECG), and expiration gases were monitored (body temperature: 37.0–38.0°C, expiration CO2: 3.5–4.0 kPa).

Ischemic Heart Model

Myocardial ischemia was induced by a previously described open–chest technique (12, 13). The rats were anesthetized, intubated, and connected to a small animal ventilator as described above. For additional pain relief, 0.05 μg/g buprenorphine (Temgesic, Reckitt Benckiser, Søborg, Denmark) was given subcutaneously 15–30 minutes before surgery. After a left thoracotomy and a pericardiotomy, the left anterior descending (LAD) artery was occluded by placing a ligature around a little plastic tube placed on top of the branch. Ischemia was verified visually by bleaching and blue-coloring of the myocardium distal to the ligature. The ligature was placed to achieve an ischemic area covering ~1/2 of the anterior wall of the left ventricle including the apex. The LAD coronary artery was occluded for 30 minutes, resulting in severe ischemia and infarction. Ischemia was followed by reperfusion achieved by releasing the tension of the ligature. A clear change in color from blue to red confirmed reperfusion. If reperfusion was not achieved, the rats were excluded from the study. For evaluation of tissue damage by the cardiac-specific biomarker troponin I, blood samples were drawn before ischemia and 1 and 2 hours after reperfusion in 6 of the 10 ischemic rats (owing to problems with the blood clotting in the catheter in the remaining 4). The level of troponin I was analyzed on an AQT90 Flex (Radiometer, Denmark). The 2 sham–operated animals underwent the surgical procedure including placing the ligature, but without occluding the artery.

Hyperpolarization

Here, 20 μL (~26 mg) of [1-13C]pyruvic acid (Sigma Aldrich, Germany) with 15mM trityl radical OX063 (Oxford Instruments, UK) and 1.5mM Dotarem (Guerbet, France) was loaded into a polarizer (HyperSense, Oxford Instruments, UK). The sample was dissolved in a neutralizing buffer (80 mM TRIS, 100 mg/L EDTA, 50 mM NaCl, 80mM NaOH), achieving a final concentration of 80 mM [1-13C]pyruvate (pH 7.0–8.0, temperature ~30°C, isotonic).

MRI

A 4.7 T preclinical MRI and magnetic resonance spectroscopy system (Agilent, Santa Clara, CA) at the Danish Research Centre for Magnetic Resonance was used for the magnetic resonance (MR) experiments. The rats were placed supine in a 13C/1H radiofrequency volume coil, and either a 13C circular receive surface coil or a 13C 4-channel receive array coil was placed over the heart (all coils from RAPID Biomedical GmbH, Germany). The volume coil had an inner diameter of 72 mm. The surface coil had a diameter of 20 mm, and it was sensitive to a depth of ~15 mm into the animal. The array coil consisted of 4 long elements of length 42.5 mm, each element parallel to each other covering the entire chest wall, and was sensitive to a depth of ~20 mm. The coil sensitivity profiles of the surface coil and the array coil in the transversal image plane are compared in Figure 1, A and B. B0 shimming was performed, and anatomical long-axis proton MR images were acquired for spatial localization of the heart using a cardiac- and respiratory-gated cine pulse sequence (repetition time [TR] = 195 milliseconds; echo time [TE] = 3 milliseconds; field of view = 60 × 120 mm2; section thickness = 2 mm; matrix size = 128 × 256, number of cardiac phases = 8). The position of the coils was verified by an external marker (oil pellet) placed on the top of the surface coil. On the array coil, the oil pellet was placed on one side of the coil and the animal was carefully placed so that the most sensitive part of the coil was closest to the heart.

Hyperpolarized 13C MRI

Via the tail vein catheter, 1.0 mL of isotopic hyperpolarized [1-13C]pyruvate solution was injected over 7–10 seconds; 7 s after the end of injection, an ECG and respiratory-gated section-selective chemical shift imaging sequence was started (section thickness = 5 mm, flip angle = 10°, circular spiral k-space 12 ×
12 trajectory matrix, TR = 69 milliseconds, TE = 1.86 milliseconds, field of view (25 × 25 mm²). The chemical shift imaging sequence was acquired from the same long-axis section through the heart as the one used for the proton cine imaging.

Data Analysis

The [1-13C]pyruvate, [1-13C]lactate, [1-13C]alanine, and [13C]bicarbonate signals were quantified and mapped using in-house written MATLAB program (The MathWorks, Natick, MA). [1-13C]Pyruvate hydrate was also detected in the 13C spectrum, but because it is not a direct metabolic product of pyruvate, it was not included in the quantification. The postprocessing quantification included apodization and zero-filling of the spatial dimensions to a matrix size of 32 × 32. The data were quantified in magnitude mode. Spectral analysis was performed by simple integration at predefined frequency offsets of the metabolites relative to the frequency of the pyruvate peak. The chemical shift imaging data were presented as metabolic maps, which were registered to the corresponding first cardiac-phase cine image (proton MRI).

The effect of ischemia on the cardiac metabolism was quantified and illustrated using a region of interest analysis using the bullseye analogy (14). The myocardium was divided in 4 regions: apex (ischemic area in animals with lesion), anterior wall (close to the coil), posterior wall (remote from the coil), and left ventricular lumen (Figure 2). Hyperpolarized [1-13C]alanine, [1-13C]lactate, and [13C]bicarbonate measured in all 4 regions and normalized to (divided by) the [1-13C]pyruvate value for the corresponding compartment, were used for comparison between healthy (preischemia) and diseased (postischemia) hearts.

Late Enhancement MRI

In 7 of the rats, the infarct area was also verified by comparing the metabolic maps visually with proton late Gd enhancement images. ECG-gated inversion recovery gradient-echo MR images were obtained 10–20 minutes after the injection of 0.3 mmol/kg Gd-based contrast agent (Dotarem, Guerbet). The inversion time was adjusted individually for each rat (350–500 milliseconds) to obtain the best contrast between the Gd-enhanced tissue and the surrounding healthy tissue (TR = 600 milliseconds; TE = 10 milliseconds; field of view = 60 × 120 mm²; section thickness = 2 mm; matrix size = 192 × 512).

Statistics

A 2-way repeated ANOVA was performed to assess the differences between pre- and postischemic areas, and multiple comparisons were performed using Bonferroni post hoc correction. Myocardial damage marker troponin I was assessed by a paired t test assuming equal standard deviation, between baseline and 1 hour in the ischemia group (owing to unequal distribution of data points and insufficient sham animals). Normality was as-
sessed with quantile–quantile plots. A value of $P \leq .05$ was considered statistically significant. Statistical analysis was performed using GraphPad Prism (GraphPad Software, Inc., La Jolla, CA).

**RESULTS**

**Hyperpolarized $^{13}$C MRI of the Healthy Heart (Preischemia)**

Hyperpolarized signal from $[1-^{13}$C$]$pyruvate, $[1-^{13}$C$]$lactate, $[1-^{13}$C$]$alanine, and $[^{13}$C$]$bicarbonate was detected in the hearts of all animals before ischemia. As expected, signal from $[1-^{13}$C$]$pyruvate was primarily detected from blood inside the ventricle (Figure 3, column 1). Signals from $[1-^{13}$C$]$lactate, $[1-^{13}$C$]$alanine, and $[^{13}$C$]$bicarbonate were confined to the myocardium (Figure 3, columns 2–4). Because of the coil sensitivity profile, signals were primarily visual in the anterior wall of the myocardium closest to the coil. In some cases, lactate signal was observed in the injured chest wall and in the blood inside the ventricle, presumably because of the operative procedure (Figure 3, column 2).

**Effects of Ischemia**

From the metabolic maps, we can observe that $[1-^{13}$C$]$pyruvate almost exclusively is seen in the ventricle and $[1-^{13}$C$]$lactate is seen both in the ventricle and the myocardium, whereas the $[1-^{13}$C$]$alanine and $[1-^{13}$C$]$bicarbonate are seen exclusively in the myocardium.

Apparent changes are seen in the raw lactate, alanine, and bicarbonate images before and after ischemia (Figure 3, columns 2–4). No apparent differences are observed in the pyruvate images (Figure 3, column 1). A decrease in the $[1-^{13}$C$]$alanine signal is observed in the ischemic region (distal, apex region) of the anterior wall after ischemia compared with before ischemia (Figure 3, column 3). An overall lower $[1-^{13}$C$]$bicarbonate signal is observed after ischemia. A statistical significant effect of ischemia was observed in all the metabolite-to-pyruvate ratios when comparing the interaction between time (before and after ischemia) and intraregional differences (regions of interest). The multiple comparison test showed a statistical significant difference between before and after ischemia, specifically in the apex for $[1-^{13}$C$]$alanine:$[1-^{13}$C$]$pyruvate ($P < .0001$); $[1-^{13}$C$]$lactate:
[1-13C]pyruvate ($P = 0.0001$); and the [1-13C]bicarbonate:[1-13C]pyruvate ($P = 0.002$) (Figure 4A–C). Comparing the individual metabolites against each other is interesting because it could indicate a shift from one metabolic pathway to another. However, no effect of ischemia could be observed in the [1-13C]alanine:[1-13C]pyruvate, [1-13C]lactate:[1-13C]pyruvate, and [1-13C]bicarbonate:[1-13C]pyruvate (C). Right: [1-13C]alanine:[13C]bicarbonate (D), [1-13C]lactate:[1-13C]alanine (E), and [1-13C]lactate:[13C]bicarbonate (F). Significant differences are indicated with an asterisk. Mean and Standard Error of Mean are presented in the figure.

Severity of Ischemia
Elevated signal in the late-enhancement images corresponded to the area, which showed reduced 13C signal in the [1-13C]alanine maps (Figure 3, column 5). The troponin I blood level peaked 1 hour after reperfusion (41.0 ± 13.2 ng/mL, mean ± standard deviation, $P = 0.0002$) and remained elevated 2–3 hours after reperfusion (Figure 5). In the 2 sham-operated rats, the 1-hour value was in the normal range, 0.13 and 0.93 ng/mL. Further, as mentioned in the Methods, ischemia was verified visually by observing bleaching and blue coloring of the myocardium distal to the ligature, and reperfusion was confirmed by a change in color of the myocardium from blue to red.

DISCUSSION
Main Finding
The present study shows how hyperpolarized MRI can be used to image regional metabolic changes in the in vivo rat heart following 30 minutes of ischemia. The observed decrease in [1-13C]alanine, [1-13C]bicarbonate, and [1-13C]lactate signal in the reperfused ischemic area likely reflects an overall depression of...
the cellular metabolism following 30 minutes of severe ischemia and cell damage due to low supply of oxygen and nutrients to the area. The metabolic maps show that changes in [1-13C]alanine signal are a good marker for regional cellular damage because the signal is well confined to the myocardium. There is less disturbing contribution of [1-13C]alanine from the blood or surrounding tissue, which can be the case for [1-13C]lactate, which is produced in the injured tissue and transported out of the cells into the blood stream. Lactate dehydrogenase has previously been used as a serum biomarker of myocardial infarction in the clinic, so enzyme leak may also contribute to the reduced regional lactate and alanine signal (15). The severity of the ischemic insult is supported by observed increase in blood levels of troponin I, which is released to the blood because of myocardial damage (16), as well as the late enhanced Gd signal observed specifically in the ischemic region.

Our results support the recent findings by Oh-Ici, et al. (11). In this study ischemia was induced by a closed chest technique in rats. The LAD was occluded for 15 minutes and the [1-13C]pyruvate metabolism evaluated after 3 minutes, 30 minutes, 1 hour and 1 week respectively. In the current study, we used an open-chest model, and the effect of 30 minutes of ischemia was evaluated 2 hours after reperfusion. Interestingly the study by Oh-Ici, et al. shows an immediate change following ischemia by a tendency toward a higher mean lactate:bicarbonate ratio from 3 minutes to 1 week later. Higher mean lactate:bicarbonate ratio indicates alterations in the balance between the aerobic and anaerobic glycolysis and thus deviates from the findings in the present study. We speculate that this discrepancy might originate from the difference in induction and duration of ischemia (15 minutes vs 30 minutes in our study), as a doubling time of a total acute ischemia rat heart without collaterals is quite substantial and expected to result in more severe metabolic changes. However, both studies show that hyperpolarized [1-13C]pyruvate is useful for the evaluation of metabolic changes following cardiac ischemia in rats.

**Perfusion**

As mentioned above, perfusion may be an interesting confounding factor to the metabolic changes we observe. Low perfusion to the ischemic area would result in lower/slower delivery of [1-13C]pyruvate to the ischemic area, which would result in a lower metabolic conversion of pyruvate to it metabolic products as we observe. In the present study, reperfusion was ensured visually in all rats before scanning. Rats in whom reperfusion was not observed were immediately excluded from the study. This is an advantage over the closed-chest model where reperfusion not can be verified visually. However, to account for minor hypoperfusion after reperfusion, the signal from [1-13C]alanine, [1-13C]lactate, and [13C]bicarbonate was normalized to the [1-13C]pyruvate signal. A reduced blood flow would decrease the supply of oxygen and inhibit cellular oxygen-dependent metabolism (bicarbonate reduction), whereas this metabolic shift is believed to increase the lactate conversion (Pasteur effect), and thus, the maintained balance between the anaerobic and the aerobic pathways (lactate-to-bicarbonate) suggests that the origin of the metabolic alterations is likely stemming from a reduced nutrient uptake in the apex region than from an oxygen-dependent altered metabolic conversion.

**Bicarbonate**

The overall low [13C]bicarbonate signal postischemia could be a consequence of a general, low cardiac pyruvate dehydrogenase activity in the rat heart because of a metabolism shift toward the use of free fatty acids for energy production (17) and/or reduced anaerobic glucose metabolism. It is known that glucose and free fatty acids compete as substrates for energy production in the heart, and especially during fasting when fatty acid utilization increases and glucose metabolism and pyruvate dehydrogenase activity decreases (18, 19). The postischemic condition could mimic a “fasted” situation because the rat has been anesthetized for a long time before scanning. Alternatively, the overall low [13C]bicarbonate signal may be caused by a reduced carbonic anhydrase activity owing to, for example, ischemic stress that would increase the clearance of the rapidly diffusing (13) CO2 by the blood and result in a reduced [13C]bicarbonate signal assuming that the end metabolic product is CO2 and not H2CO3 (20). Acidosis and increased flow would only, to a minor degree, reduce the [13C]bicarbonate signal. The glycolytic metabolism might fluctuate because of fluctuations in the glucose metabolism and thereby result in an unstable [13C]bicarbonate signal. A way to avoid fluctuations and keep a constant high plasma glucose levels could be to supply the animal with an oral glucose load or a direct infusion of glucose during the MR acquisition, which previously has shown to ensure a high myocardial glucose metabolism and a resulting high hyperpolarized [13C]bicarbonate signal (10, 11, 21). Looking back, imaging of the sham-operated rats could have been useful to further evaluate the effects of anesthesia on the metabolic signals.

**Coils**

Two different coils were used, and data from both coils were included in the present study. A sensitivity profile of both coils is shown in Figure 1. The reason for using 2 coils was to evaluate their sensitivity and ensure coverage of the heart. Coverage of
the entire heart in the anterior–posterior direction (from coil surface into the animal) was not achieved by any of the coils. A strong signal was detected from the anterior wall closest to the coils, but low $^{13}$C signal was detected from the posterior myocardial wall using both coils. Thus, signal differences were only assessed in the anterior wall, from a long-axis section. The length of the array coil elements is 42.5 mm in the head to tail direction, which is sufficiently long to cover the anterior wall. Because central positioning of the heart in the coil always was ensured, and because the heart was scanned both before and after ischemia with the same coil, the effects from coil sensitivity profiles on the evaluation from the anterior wall are considered to be negligible.

Concluding Remarks

This study shows that regional metabolic changes following severe myocardial ischemia can be imaged in the in vivo rat heart by means of hyperpolarized $[1-^{13}$C]pyruvate. Signal from both $[1-^{13}$C]pyruvate, $[1-^{13}$C]lactate, $[1-^{13}$C]alanine, and $[^{13}$C]bicarbonate could be detected before ischemia, and localized metabolic images of the anterior part of the rat myocardium using a surface coil, a 4-channel array coil, and a simple chemical shift imaging sequence could be produced. The decrease in the $[1-^{13}$C]alanine and $[1-^{13}$C]lactate signals in the ischemic region after 30 minutes of occlusion of the LAD coronary artery suggests that the ischemic insult is severe, affecting not only the sensitive mitochondrial metabolism but also the cytosolic metabolism. It most likely accounts for an overall depression of the cellular metabolism due to nutrient starvation, which is in agreement with the results of previous cardiac studies in pigs by use of hyperpolarized $[1-^{13}$C]pyruvate (3, 6, 10). The metabolic finding was supported by the concomitant increase in the blood levels of troponin I and enhanced Gd signal using late enhancement MRI.

This study and previous studies using hyperpolarized $^{13}$C MRI suggest that the technique has the potential to advance basic knowledge and improve diagnosis and prognosis of cardiac diseases. A successful translation of hyperpolarized $^{13}$C MRI into the clinic will require an extensive understanding of the biological systems examined, which can be exploited in animal models such as the one presented here. The first hyperpolarized $[1-^{13}$C]pyruvate MRI of a human heart was recently performed (22), proving the feasibility of the technique to acquire high signal-to-noise ratio metabolic images in human hearts. However, further technical advancements, such as increased polarization, coils with better sensitivity, and improved MR sequences to detect low $^{13}$C signals, are still needed. In this study, $[1-^{13}$C]pyruvate was used. In addition, other hyperpolarized substrates may offer alternative ways to study cardiac metabolism or specific chemical pathways affected by disease.

ACKNOWLEDGMENTS

Authors’ Contributions: MHL carried out the animal experiments, including the heart surgery (LAD occlusion), and performed the MR scanning, collected the data, and drafted the manuscript; PM programmed and optimized the MR pulse sequences for the hyperpolarized studies and optimized the late enhancement imaging, constructed hyperpolarized image analysis tools, participated in the image analysis, and helped to draft the manuscript; CL constructed the hyperpolarized image analyzing tool, participated in data analysis, and helped to draft the manuscript; SAB helped with the hyperpolarized MRI scanning; JHAL helped coordinate the study, optimize the hyperpolarized MRI scanning and collect the data, interpret the results, and draft the manuscript; LVS coordinated the study and helped design the hyperpolarized $^{13}$C MRI experiments, helped collect the data and draft the manuscript; OBP helped interpret the results and draft the manuscript; PA conceived and designed the study, helped coordinate the study, and helped to draft the manuscript.

REFERENCES


Monitoring Radiation Treatment Effects in Glioblastoma: FLAIR Volume as Significant Predictor of Survival

Matthew D. Garrett¹, Ted K. Yanagihara¹, Randy Yeh², Guy M. McKhann³,⁴, Michael B. Sisti³,⁴, Jeffrey N. Bruce³,⁴, Sameer A. Sheth³,⁴, Adam M. Sonabend³,⁴, and Tony J. C. Wang¹,⁴

¹Departments of Radiation Oncology; ²Radiology; and ³Neurological Surgery, Columbia University Medical Center, New York, NY; and ⁴Herbert Irving Comprehensive Cancer Center, Columbia University Medical Center, New York, NY

Corresponding Author:
Tony J. C. Wang, MD
Department of Radiation Oncology, Columbia University Medical Center, 622 West 168th Street, BNH B011, New York, NY 10032; E-mail: tjw2117@cumc.columbia.edu

Key Words: glioblastoma, radiation, FLAIR, MRI, brain cancer

INTRODUCTION
Glioblastoma (GBM) is the most common primary malignant neoplasm of the central nervous system with an annual incidence in the United States of 4–5 per 100,000 (1). The standard of care includes a maximal safe surgical resection and adjuvant radiotherapy with concurrent and continued adjuvant temozolomide chemotherapy (2). Despite aggressive trimodality therapy, patients invariably recur with a median interval to disease progression of ~8 months and a 5-year overall survival (OS) rate of <10% (3).

After completing radiation therapy, patients are monitored closely, typically with serial magnetic resonance imaging (MRI) and routine clinical assessment (4). Disease progression is diagnosed radiographically and/or in the context of clinical symptoms or declining performance status. The MacDonald criteria, introduced in 1990, provided an objective methodology for tumor response assessment using changes in tumor area derived from maximal bidimensional measurements of enhancing regions on T1-weighted imaging with gadolinium. In addition, corticosteroid use and performance status were also considered in these criteria (5). In 2010, the Response Assessment in Neuro-Oncology Working Group (RANO) proposed new criteria for response assessment following chemoradiotherapy (CRT) to address the limitations of the MacDonald criteria, including the problems of pseudoprogression and nonenhancing tumor progression (6). Therefore, the criteria for disease progression differ depending on the time interval from initial treatment, with 12 weeks as a discriminator to account for pseudoprogression. In addition, nonenhancing tumor shown by increased T2-weighted Fluid-Attenuated Inversion Recovery (FLAIR) signal was included in the evaluation for disease progression.

In addition to post-treatment follow-up, FLAIR signal is an integral component of radiation therapy (RT) planning for GBM. Clinical target volumes in radiation treatment include T1 and...
T2/FLAIR hyperintensities along with the resection cavity and a 2-cm margin (7). Given the importance placed upon FLAIR signal in radiation treatment and in post-treatment surveillance imaging, many studies have attempted to correlate FLAIR volume with important clinical variables such as OS and progression-free survival (PFS) in the posttreatment setting. Results of these studies have varied, and despite RANO recommendations encouraging its inclusion, the significance and prognostic value of FLAIR signal have not yet been clearly shown in tumor response assessment of GBM (8–12).

The purpose of the current study was to test the hypothesis that the volume of hyperintense FLAIR signal correlates with meaningful clinical outcomes, particularly OS and PFS. Earlier studies have relied on ellipsoid and spheroid estimates of FLAIR volume, generating crude approximations of tumor burden (13). The RANO report itself, in 2010, admitted a technological shortcoming in reference to more objective quantifications of FLAIR volumes (6). In contrast to these previous limitations, our method of manual segmentation followed by computerized volumetry simultaneously reduces user bias and improves reproducibility (14, 15).

**METHODODOLOGY**
All analyses were performed in compliance with our institutional review board guidelines, and consent was waived based on the retrospective nature of the study. We searched our departmental database for patients treated between January 2011 and February 2014 with standard-of-care therapy at the Columbia University Medical Center for a new diagnosis of histologically confirmed GBM. To be included, patients must have undergone T2/FLAIR MRI acquisition within 2 weeks before adjuvant treatment was initiated, as well as between 60 and 180 days after conclusion of RT. Because distinguishing between progression and pseudoprogression using MRI is not possible, and RANO criteria recommend a waiting period posttreatment before changes in management can be made (6), our study set a minimum wait period of 60 days after conclusion of RT. All imaging sequences were reviewed manually and were excluded if any data were of poor quality (ie, motion-degraded) by a user who was blinded to the patient’s clinical data, purpose, or results of this study. Patients with multiple GBM foci were similarly excluded. For patients with multiple FLAIR sequences within the 2- to 6-month window posttreatment, the earliest available images were used.

The radiation protocol for 28 patients consisted of fractionated linear accelerator-based irradiation Monday through Friday for a period of 6 weeks, with patients receiving 2 Gy per fraction for a total of 60 Gy. All patients received concomitant temozolomide (75 mg/m²/d, 7 d/wk) during radiotherapy followed by adjuvant temozolomide for up to 6 cycles (150–200 mg/m²/d for 5-day duration, each 28-day cycle). Regions of FLAIR hyperintensity were contoured by a single experimenter adhering to The Radiation Therapy Oncology Group guidelines and blinded to the OS results. Functional MRI of the Brain Software Library (FSL) was used to perform segmentation, and the volumes were calculated with MATLAB software.

A hypofractionated protocol, consisting of 2.67 Gy per fraction for total dose of 40 Gy over 3 weeks, was prescribed to 12 patients. The hypofractionated protocol was reserved for elderly patients given its shorter treatment period; in accordance with other published data, an earlier study evaluating outcomes of patients at our own institution yielded no significant difference in OS between standard and hypofractionated treatment regimens when adjusted for age and concomitant chemotherapy (16).

OS was computed from the start of RT to date of death either known by Columbia University Medical Center or from internet searches of publicly available death data; OS was defined as all-cause mortality. Any patient lost to follow-up or for whom follow-up indicated that death was imminent but a specific date could not be identified was censored at the date of last confirmed contact; in this manner, 11/40 patients were censored. PFS was computed from the start of RT to the date of primary disease progression confirmed by an independent neuroradiologist and immediately followed by a change in management, typically bevacizumab per standard-of-care guidelines; 4/40

### Table 1. Patient Characteristics

<table>
<thead>
<tr>
<th>Total</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>16 (40.0%)</td>
</tr>
<tr>
<td>Male</td>
<td>24 (60.0%)</td>
</tr>
<tr>
<td><strong>Median age at diagnosis</strong></td>
<td>62.5 (range 16–85)</td>
</tr>
<tr>
<td><strong>Alive at study time</strong></td>
<td>1 (2.5%)</td>
</tr>
<tr>
<td><strong>Median overall survival</strong></td>
<td>457 days (range 119–1372)</td>
</tr>
<tr>
<td><strong>Median progression-free survival</strong></td>
<td>176 days (range 42–835)</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>28 (70%)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>9 (22.5%)</td>
</tr>
<tr>
<td>Asian</td>
<td>2 (5%)</td>
</tr>
<tr>
<td>Black</td>
<td>1 (2.5%)</td>
</tr>
<tr>
<td><strong>Surgery</strong></td>
<td></td>
</tr>
<tr>
<td>Biopsy only</td>
<td>2 (5%)</td>
</tr>
<tr>
<td>Subtotal resection</td>
<td>33 (82.5%)</td>
</tr>
<tr>
<td>Total resection</td>
<td>5 (12.5%)</td>
</tr>
<tr>
<td><strong>Radiotherapy dose</strong></td>
<td></td>
</tr>
<tr>
<td>60 Gy</td>
<td>28 (70%)</td>
</tr>
<tr>
<td>40 Gy</td>
<td>12 (30%)</td>
</tr>
<tr>
<td><strong>Median FLAIR volumes</strong></td>
<td></td>
</tr>
<tr>
<td>PreRT</td>
<td>35.957 cm³ (range 3.150–144.629)</td>
</tr>
<tr>
<td>PostRT</td>
<td>29.143 cm³ (range 0.297–175.641)</td>
</tr>
<tr>
<td><strong>MGMT status</strong></td>
<td></td>
</tr>
<tr>
<td>Methylated</td>
<td>10 (25%)</td>
</tr>
<tr>
<td>Unmethylated</td>
<td>16 (40%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>14 (35%)</td>
</tr>
<tr>
<td><strong>IDH-1 status</strong></td>
<td></td>
</tr>
<tr>
<td>Mutated</td>
<td>3 (7.5%)</td>
</tr>
<tr>
<td>Unmutated</td>
<td>36 (90%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1 (2.5%)</td>
</tr>
</tbody>
</table>
patients were censored because death occurred before confirmed progression of disease.

Statistical analysis of volumetric parameters and other relevant parameters used Pearson correlation and Cox proportional hazards model. Statistical significance was \( P \leq .05 \), and all statistical calculations and graphs were performed using SPSS version 24.

RESULTS

In total, 40 patients (females, 16; males, 24; median age, 62.5 years at diagnosis) were included in the final analysis (Table 1). On the basis of the operative report, surgical procedures were defined as biopsy, subtotal resection, or total resection. The median time period between surgery and treatment onset was 29 days. At the time of this study, 1/40 patients remained alive. The median OS of the cohort is 457 days, and the median PFS is 176 days.

Four ethnicities (Latino, Asian, black and white) were represented in the patient population, with whites and Latinos comprising 92.5% of the cohort. An earlier study at our institution yielded no significance in OS between these four racial groups after treatment for GBM (17).

\( O^6 \)-methylguanine DNA methyltransferase (MGMT) promoter methylation status was known in 26 patients; 38% of this group were methylated. IDH-1 mutation status was known in 39 patients; all but 3 in this group were wild-type.

For the 80 pre- and posttreatment FLAIR sequences analyzed, the median pre-RT volume was 35.975 cm\(^3\), and the median post-RT volume was 29.143 cm\(^3\). Pearson bivariate correlation was performed for all patients; for censored patients, the date of last visit was used as a survival endpoint. Figure 1 displays the results of the regression analyses for OS. There was no significant correlation between the volume of FLAIR signal pretreatment and OS \( (P = .956, R = -0.009) \). A statistically significant correlation was found between the volume of post-treatment FLAIR signal and OS \( (P = .048, R = -0.315) \); the negative slope of the regression line indicates worse OS with increased post-RT FLAIR volume in the 2- to 6-month posttreat-
Figure 2. Pretreatment FLAIR volume was not correlated with progression-free survival (PFS) (A). Posttreatment FLAIR volume was significantly correlated with PFS and with a moderate negative Pearson coefficient (B). Patient age at diagnosis was not correlated with PFS (C). Karnofsky performance status (KPS) was not correlated with PFS (D).

Figure 3. Worst survival: representative patient who had the lowest survival in the cohort (119 days). The patient had a moderate degree of FLAIR hyperintensity before beginning chemoradiation (A) that was dramatically increased at 71 days after treatment (B). Best survival: representative patient who had the best survival in the cohort (1372 days). The patient had a dominant FLAIR signal prior to beginning chemoradiation (C) that nearly resolved at the 96-day time point (D).
The negative slope of the regression line indicates worse PFS with increased post-RT FLAIR volume in the 2- to 6-month posttreatment window. Neither Pearson correlation of patient age at the time of diagnosis nor KPS was significant with PFS ($P = .299$, $R = 0.168$ and $P = .246$, $R = 0.188$, respectively) (Figure 3).

Univariable and multivariable analyses were performed with Cox proportional hazards model, with the latter using age at diagnosis, pre-RT FLAIR volume, and post-RT FLAIR volume as covariates (Tables 2 and 3). Post-RT FLAIR volume reached statistical significance for OS in both univariable and multivariable analyses ($P = .017$ and $P = .043$, respectively), with hazard ratios of 1.010 and 1.018, respectively. Post-RT FLAIR volume also reached statistical significance for PFS in both univariable and multivariable analysis ($P < .001$ and $P < .001$, respectively), with hazard ratios of 1.008 and 1.016, respectively. Age at diagnosis reached statistical significance for OS in both univariable and multivariable analyses ($P = .013$ and $P = .024$, respectively), although not for PFS.

Pre-RT FLAIR volume on multivariable analysis and the percent change in FLAIR volume from pre- to post-RT on univariable analysis showed a significant association with PFS ($P = .042$ and $P = .001$, respectively), although not with OS. Patient gender and radiotherapy dose (60 Gy vs 40 Gy) were also not associated with OS or PFS on univariable analysis, although the latter showed a trend toward association with OS ($P = .075$).

In total, 26 patients had known MGMT promoter status; there was no significant association between MGMT and OS/PFS within this cohort. Patients with IDH-1 mutations showed a trend toward increased OS ($P = .096$), although not significant. Although extent of tumor resection is a well-established predictor for OS, no significant association between extent of resection and OS/PFS was determined within this patient cohort (18).

**DISCUSSION**

Hyperintensity on FLAIR sequences is presumed to represent edema due to microscopic cancer infiltration (19), and regions of FLAIR signal are considered as essential elements of treatment planning. Many studies have sought to correlate patient outcomes and tumor response on the basis of FLAIR signal alone, and the results have been mixed. Despite the importance of FLAIR signal in radiotherapy planning, in modern clinical practice, performance status and patient-reported symptoms are the most indicative of neurological decline and disease progression. Our study sought to evaluate the prognostic value of FLAIR signal in the context of GBM in the pre- and posttreatment setting. Our results also showed a statistically significant nega-
tive correlation between FLAIR volume 2–6 months after CRT and both OS and PFS.

The historical practice of geometric approximations of tumor burden has been increasingly vilified in recent years. Iliadis et al. evaluated preoperative FLAIR volumes for survival prognosis in 50 patients with GBM and compared geometric approximations with computerized volumetry. The conclusion was that geometric models tend to overestimate FLAIR volume and should not be used (13). Zhang et al. proposed a pretreatment volume ratio of FLAIR to gross tumor burden as a significant predictor of survival, but listed its volume acquisition method (ie, ellipsoid approximation) as a limitation in accuracy (9). In a seminal study assessing reliability, precision, and ease of use of computerized volumetry among doctors, students, and volunteers, Huber et al. determined that changes in FLAIR volume in particular had the fewest precision errors among an already reliable set of tested parameters (14). The results of our study mirror those of an Israeli study published in 2016, in which pretreatment FLAIR volume was not associated with OS, whereas 3-month post-treatment volumes were significantly correlated with OS (20). We posit that there are inconsistencies in the findings of other studies owing to variations in analytic technique and reliance, by some, on geometric models to estimate FLAIR volume. Although these approaches avoid subjectivity, our analysis was blinded to clinical data and relied on a standard FLAIR contouring technique used for RT planning.

Our study showed no correlation between pre- and post-treatment FLAIR volumes. This lack of coherence could be explained by the acute effects of both ionizing radiation and chemotherapy leading to alterations in the volume and exact composition of FLAIR hyperintense regions. It is for these reasons that pseudoprogression confounds analyses in the immediate post-treatment setting. We accounted for potential radiographic pseudoprogression by setting a minimum waiting period of 60 days following CRT, a time frame similar to the 12 weeks proposed by the RANO criteria. Again, disagreements in the published literature may be explained by the time points analyzed and the influence of pseudoprogression on the results.

Our study also determined a significant correlation between OS and both patient age at diagnosis and KPS; these findings are in congruence with well-established prognostic thought and are nonmodifiable risk factors. They do, therefore, lend further credence to the validity of this patient cohort within the context of the larger patient population.

| Table 3. Cox Proportional Hazards Model to Assess Effects on PFS |
|--------------------------|--------------------------|--------------------------|--------------------------|
| **Covariate** | **Univariable Analysis** | **Multivariable Analysis** |
|               | **Coefficient (B)** | **HR [expB]** | **95% CI** | **P Value** | **Coefficient (B)** | **HR [expB]** | **95% CI** | **P Value** |
| Age | 0.008 | 1.008 | 0.988–1.028 | .431 | 0.002 | 1.002 | 0.983–1.021 | .849 |
| KPS | −0.017 | 0.983 | 0.956–1.012 | .254 | N/A | N/A | N/A | N/A |
| Gender | | | | | | | | |
| Female | 0.113 | 1.120 | 0.565–2.220 | .746 | N/A | N/A | N/A | N/A |
| FLAIR volume | | | | | | | | |
| Pre-RT | 0.001 | 1.001 | 0.990–1.013 | .811 | −0.012 | 0.988 | 0.976–0.999 | .042 |
| Post-RT | 0.020 | 1.020 | 1.011–1.029 | <.001 | 0.026 | 1.026 | 1.015–1.038 | <.001 |
| % Change | 0.001 | 1.001 | 1.001–1.002 | .001 | N/A | N/A | N/A | N/A |
| MGMT status | | | | | | | | |
| Unmethylated | −0.170 | 0.844 | 0.367–1.941 | .690 | N/A | N/A | N/A | N/A |
| IDH-1 status | | | | | | | | |
| Unmutated | 0.189 | 1.208 | 0.287–5.084 | .796 | N/A | N/A | N/A | N/A |
| RT dose | | | | | | | | |
| 40 Gy | 0.261 | 1.298 | 0.619–2.719 | .490 | N/A | N/A | N/A | N/A |
| 60 Gy | (reference) | (reference) | (reference) | (reference) | (reference) | (reference) | (reference) | (reference) |
| Extent of resection | | | | | | | | |
| Biopsy only | −0.523 | 0.593 | 0.108–3.257 | .633 | N/A | N/A | N/A | N/A |
| Subtotal | 0.171 | 1.186 | 0.409–3.444 | .754 | N/A | N/A | N/A | N/A |

Note: Data in bold represent statistical significance (P < .05).
Abbreviations: PFS, progression-free survival; KPS, Karnofsky performance status; HR, hazard ratio; CI, confidence interval

2016. FLAIR as a Significant Predictor of Survival in Patients With GBM

136
Although selection bias is a common risk in retrospective studies, we minimized this risk by including all patients with GBM treated at our institution over the aforementioned time period who had high-quality pre- and posttreatment FLAIR sequences; the latter occurring 60–180 days after CRT. Although our study was limited by the small patient population, the findings are made robust by inclusion of only high-quality imaging data and selection of time points that limit the effects of pseudoprogression. As with any study evaluating nonenhancing tumor using T2/FLAIR signal, our study was also limited by inherent difficulties and inaccuracy of determining the exact extent of nonenhancing tumor, as surrounding peritumoral edema and postradiation white matter changes have similar signal on FLAIR sequences (21). Another limitation is the omission of concomitant steroid usage during and after treatment; these data were not available for the majority of patients. Steroids, such as dexamethasone, inhibit both phospholipase A2, as well as histamine release from mast cells, leading to reduced edema and therefore FLAIR signal (22, 23). It is thus warranted that future studies control for steroid usage and possibly include dosage as a covariate in multivariable analyses. Incidentally, no patient received antiangiogenic agents such as bevacizumab during the periods in which imaging data were collected.

In conclusion, post-treatment FLAIR volume is a significant predictor of OS and PFS in patients with GBM, as determined with multiple statistical methods. Future models should consider this specific time point in the development of more refined prognostic tools.

Conflict of Interest: The authors have no conflicts of interest to declare.

Disclosures: No disclosures to report.

REFERENCES

Correcting Nonpathological Variation in Longitudinal Parametric Response Maps of CT Scans in COPD Subjects: SPIROMICS

Antonio Fernández-Baldera¹, Charles R. Hatt², Susan Murray³, Eric A. Hoffman⁴, Ella A. Kazerooni¹, Fernando J. Martínez⁵, MeiLan K. Han⁶, and Craig J. Galbán¹

¹Department of Radiology, University of Michigan, Ann Arbor, MI; ²Imbio, LLC, Minneapolis, MN; ³Department of Public Health, University of Michigan, Ann Arbor, MI; ⁴Departments of Radiology and Biomedical Engineering, University of Iowa, IA; ⁵Department of Medicine, Cornell University, NY; and ⁶Department of Internal Medicine, University of Michigan, Ann Arbor, MI

Corresponding Author:
Craig J. Galbán, PhD
Department of Radiology, University of Michigan, BSRB, Room A506, 109 Zina Pitcher Place, Ann Arbor, MI 48109-2200; E-mail: cgalban@med.umich.edu

Key Words: COPD, parametric response mapping, longitudinal, computed tomography

Abstract: Small airways disease (SAD) is one of the leading causes of airflow limitations in patients diagnosed with chronic obstructive pulmonary disease (COPD). Parametric response mapping (PRM) of computed tomography (CT) scans allows for the quantification of this previously invisible COPD component. Although PRM is being investigated as a diagnostic tool for COPD, variability in the longitudinal measurements of SAD by PRM has been reported. Here, we show a method for correcting longitudinal PRM data because of nonpathological variations in serial CT scans. In this study, serial whole-lung high-resolution CT scans over a 30-day interval were obtained from 90 subjects with and without COPD accrued as part of SPIROMICS. It was assumed in all subjects that the COPD did not progress between examinations. CT scans were acquired at inspiration and expiration, spatially aligned to a single geometric frame, and analyzed using PRM. By modeling variability in longitudinal CT scans, our method could identify, at the voxel-level, shifts in PRM classification over the 30-day interval. In the absence of any correction, PRM generated serial percent volumes of functional SAD with differences as high as 15%. Applying the correction strategy significantly mitigated this effect with differences /H110111%. At the voxel-level, significant differences were found between baseline PRM classifications and the follow-up map computed with and without correction (P < .01 over GOLD). This strategy of accounting for nonpathological sources of variability in longitudinal PRM may improve the quantification of COPD phenotypes transitioning with disease progression.

Introduction

Chronic obstructive pulmonary disease (COPD) is a leading cause of morbidity, mortality, and burden on the world’s health and financial systems (1, 2). Advances in the clinical management of patients with COPD have led to an improved understanding of the multitude COPD phenotypes. It has been postulated that a spectrum of pathological processes may result in unique progression patterns among these patients. Extensive research has been devoted toward identifying surrogate biomarkers of disease progression with a strong emphasis on noninvasive imaging techniques and analytical approaches (3).

Parametric response mapping (PRM) is an analytical approach that, when applied to spatially aligned high-resolution computed tomography (HRCT) scans, allows both visualization and quantification of lung parenchyma affected by small airways disease (SAD), even when only emphysema is visibly observed (4). This technique quantifies a previously occult component of COPD and can be applied to retrospective HRCT data. Included in various NIH-funded clinical trials on COPD (5, 6), PRM of functional SAD (fSAD) has been demonstrated as an independent indicator of clinically relevant outcome measures (7). More recent studies have identified PRM as a surrogate of spirometric decline in COPD (7) and also a means for identifying and monitoring the onset of bronchiolitis obliterans syndrome in bone marrow and lung transplant recipients (8–10). In a preliminary study, PRM was evaluated as a marker for monitoring change in disease classification (ie, normal, fSAD, and emphysema) from subjects accrued as part of SPIROMICS (5). In this study, “voxel-based tracking,” a method for evaluating longitudinal changes in PRM classification at the voxel level, has been used. Although this approach when applied to PRM shows promise at providing local disease progression, variability in
Hounsfield unit (HU) values from uncontrollable sources (eg, scanner noise, patient breathing level, and image registration) may result in shifts in voxel PRM classification that are not related to alterations in disease state (6).

Various studies have demonstrated the efficacy of PRM as a diagnostic and prognostic indicator of decline in pulmonary function and COPD severity. Nevertheless, the use of PRM to monitor COPD progression has shown a high sensitivity of voxel classification to HU variability between longitudinal CT examinations resulting in erroneous results. The purpose of this study was to present a strategy to mitigate the effects of nonpathological HU variability on voxel classification for analyzing COPD progression using PRM.

**METHODOLOGY**

**Study Population**

All clinical procedures were conducted under an institutional review board-approved protocol, and all subjects involved provided written informed consent. In total, 90 subjects (age range at baseline, 40–80 years), with paired volumetric inspiratory and expiratory HRCT scans and clinical examinations at a 30-day interval, were prospectively accrued as part of the Repeatability and Replicate Substudy of SPIROMICS (11). Subjects evaluated included smokers with a smoking history of ≥20 pack-years and GOLD (Global Initiative for Chronic Obstructive Lung Disease) scores across the scale including 0, 1, 2, 3 and 4 (11-13) and never-smokers (79 smokers and 11 never-smokers, respectively). Postbronchodilator forced expiratory volume at 1 second (FEV1) was determined by spirometry at each time point. In our set of subjects, exclusion criteria included intolerance to bronchodilators, body mass index (BMI) >40 kg/m² at baseline, presence of non-COPD obstructive lung disease, diagnosis of unstable cardiovascular disease, lung surgery, or presence of metal in the chest that might affect chest CT interpretation. Further, 13 subjects from this cohort have been previously used to define thresholds that indicate disease-provoked changes in PRM metrics (5).

**Computed Tomography**

Whole-lung volumetric multidetector HRCT scans were acquired for all 90 subjects using the SPIROMICS imaging protocol (13). The current of 120 kVP was adjusted to meet the CT dose index volume targets for inspiration and expiration by making use of 3 settings—large (BMI > 30 kg/m²), medium (BMI, 20–30 kg/m²), and small (BMI < 20 kg/m²)—with vendor-specific reconstruction kernels (Standard, B, B35, FC03) (11). In this study, HRCT data reconstructed using the “GE standard” kernel were analyzed. Quantitative HRCT data were presented in HU values, in which stability of CT measurements for each scanner was monitored on a monthly basis by use of the COPDGene phantom (14). For reference, ambient air and water attenuation values should be −1000 and 0 HU, respectively. Of the 90 subjects, rescanning using a different scanner was conducted among 11 subjects and that using a different field of view (ΔFOV > 5%) from their original scan was conducted among 22 subjects. To reduce scanner noise, a 3³ median filter was applied to all CT scans before processing and analysis.

**Parametric Response Mapping**

Lung segmentation and image registration to a single geometric frame (ie, baseline expiration CT scan) were performed on all paired CT data using Lung Density Analysis (LDA) software (Imbio, LLC, Minneapolis, MN). Classification of individual voxels was performed using in-house algorithms developed using MATLAB version 2015b (MathWorks, Inc, Natick, MA). Details on the PRM analysis have been previously reported (4). The nomenclature of these measures for normal lung parenchyma, fSAD, and emphysema includes PRMNormal, PRMfSAD, and PRMEmph, respectively. Additional details are provided in the online supplemental Methods.

**Correction Strategy for Longitudinal PRM**

After establishing that the difference in HU between interval examinations has a quasi-normal distribution (online supplemental Methods), we determined the variance using the serial inspiration and expiration voxel data. This approach is analogous to previous works on voxel-to-voxel therapeutic response assessment in cancer (15, 16). These data were plotted on a Cartesian coordinate system with the x and y axes denoted as the baseline and follow-up, respectively. Using principal component analysis, the data were transformed to the axes of primary and secondary variance (principal and secondary eigenvectors, respectively). Next, a linear fit along the principal eigenvector was performed with the subsequent residuals mapped into the second

---

**Figure 1.** Workflow summarizing the correction of erroneous PRM classifications.

---

**Correction in Longitudinal PRM**

TOMOGRAPHY.ORG | VOLUME 3 NUMBER 3 | SEPTEMBER 2017 139
Correction in Longitudinal PRM

**Table 1. Subject Characteristics**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Never-Smokers</th>
<th>GOLD 0</th>
<th>GOLD 1</th>
<th>GOLD 2</th>
<th>GOLD 3</th>
<th>GOLD 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>11</td>
<td>17</td>
<td>15</td>
<td>18</td>
<td>20</td>
<td>9</td>
</tr>
<tr>
<td>Gender [M/F]</td>
<td>4/7</td>
<td>10/7</td>
<td>13/2</td>
<td>13/5</td>
<td>13/7</td>
<td>5/4</td>
</tr>
<tr>
<td>Age [years]</td>
<td>55 (7)</td>
<td>55 (8)</td>
<td>67 (8)</td>
<td>64 (8)</td>
<td>64 (9)</td>
<td>61 (9)</td>
</tr>
<tr>
<td>Height [cm]</td>
<td>170 (12)</td>
<td>169 (9)</td>
<td>172 (6)</td>
<td>172 (10)</td>
<td>172 (10)</td>
<td>169 (13)</td>
</tr>
<tr>
<td>Weight [kg]</td>
<td>79 (20)</td>
<td>82 (20)</td>
<td>79 (15)</td>
<td>84 (18)</td>
<td>81 (24)</td>
<td>78 (20)</td>
</tr>
<tr>
<td>BMI [kg/cm²]</td>
<td>27 (4)</td>
<td>29 (5)</td>
<td>27 (5)</td>
<td>28 (6)</td>
<td>27 (6)</td>
<td>27 (5)</td>
</tr>
<tr>
<td>FEV1 [% predicted]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 days</td>
<td>110 (7)</td>
<td>98 (12)</td>
<td>92 (10)</td>
<td>65 (9)</td>
<td>41 (6)</td>
<td>26 (4)</td>
</tr>
<tr>
<td>30 days</td>
<td>107 (9)</td>
<td>92 (12)</td>
<td>86 (13)</td>
<td>58 (12)</td>
<td>36 (8)</td>
<td>24 (8)</td>
</tr>
<tr>
<td>Exp volume [L]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 days</td>
<td>2.39 (0.61)</td>
<td>2.81 (0.82)</td>
<td>3.36 (1.08)</td>
<td>3.87 (1.11)</td>
<td>4.57 (1.22)</td>
<td>5.09 (0.99)</td>
</tr>
<tr>
<td>30 days</td>
<td>2.49 (0.67)</td>
<td>2.68 (0.55)</td>
<td>3.17 (0.71)</td>
<td>3.81 (0.85)</td>
<td>4.53 (1.00)</td>
<td>5.11 (0.97)</td>
</tr>
<tr>
<td>Ins volume [L]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 days</td>
<td>5.93 (1.68)</td>
<td>5.59 (0.95)</td>
<td>5.96 (1.34)</td>
<td>6.09 (1.45)</td>
<td>6.31 (1.39)</td>
<td>6.53 (1.31)</td>
</tr>
<tr>
<td>30 days</td>
<td>5.76 (1.61)</td>
<td>5.49 (1.05)</td>
<td>5.88 (1.24)</td>
<td>6.28 (1.32)</td>
<td>6.38 (1.50)</td>
<td>6.55 (1.33)</td>
</tr>
</tbody>
</table>

Subject characteristics by GOLD stage. Values are mean (standard deviation) (except for gender). Characteristics not temporally disaggregated were recorded at 0 day.

Abbreviations: BMI, body mass index; FEV1, forced expiratory volume in 1 second (percentage of predicted); Ins and Exp, inspiration and expiration, respectively.

**RESULTS**

**Subject Characteristics**

Study cohort population characteristics are provided in Table 1, and PRM results are displayed in Table 2. No significant differences in BMI, height, and weight were observed between strata. Never-smokers were found to be significantly younger than GOLD 1 subjects. As expected, lung volumes, FEV1, and PRM metrics at both interval examinations were found to be dependent on GOLD. No significant correlations were observed between FEV1 and PRM classifications within the stratum and at individual examinations. In addition, no significant relationships were obtained between subject gender and GOLD status. Finally, change in PRMNormal, PRMPSAD, and PRMEmph for each GOLD stratum was found to be insignificant over the 30-day interval (all cases with \( P > .07 \)). The 95% confidence intervals in changes in PRM metrics over the 30-day interval are presented in the online supplemental Results.

**Statistical Analysis**

Differences in subject age, height, weight, BMI, FEV1 (percent predicted), lung volumes, and percent lung volumes of PRM metrics at both interval examinations were assessed using the 1-way ANOVA controlled for multiple comparisons (Bonferroni post hoc test). Association between gender population and GOLD was assessed using the log likelihood ratio test. Temporal changes in lung volumes, relative volume of PRM metrics, and FEV1 were assessed using the Wilcoxon signed rank test. Differences in the percent agreement, ie, sum of voxels with like-PRM class normalized to total lung voxels, between uncorrected and corrected PRM at follow-up were also analyzed using the Wilcoxon signed rank test. Differences in percent agreement were also analyzed over GOLD stratum using the Kruskal–Wallis test. Percent agreement was computed using MATLAB version 2015b (MathWorks, Inc.). Statistical analyses were conducted with SPSS version 2.1 (IBM, Armonk, NY). All the results were considered statistically significant at the .05 level.

eigenvector axis. Residuals were used to calculate the 95% confidence interval of the fit (95% CI). The value of the confidence interval was transformed back to the original image space and defined as the index of measurement variability (IMV). This procedure was performed among all subjects and CT breath-hold examinations at both inspiration and expiration. To account for HU dependence on IMV, a cumulative exponential model was applied to all voxels in the CT data:

\[
IMV(x) = V[1 - e^{-(x/R + 1000)}] + \delta,
\]

where \( V = A - \delta \), \( A \) is the maximum amplitude of IMV, \( \delta \) is the minimum amplitude of IMV (±30 HU), \( R \) is a rate constant, and \( x \) is the voxel HU value. Model derivation and calculation of parameter values are provided in the online supplemental Methods. This functional form of IMV was incorporated into the correction strategy to account for variations in voxel variance.

Figure 1 shows an illustration of the correction strategy. In step 1, we calculated the difference and average maps, ie, \(|\Delta HU|\) and \(\langle HU\rangle\), respectively, between serial examinations. In step 2, we applied the following logical statement to each voxel of the baseline scan: if \(|\Delta HU| < \langle HU\rangle\), then voxel = 1, or else voxel = 0. This binary map was multiplied to the baseline scan, whereas the inverse was multiplied to the follow-up scan. In step 3, a composite follow-up CT scan was generated by summing the masked baseline and follow-up scans from step 2. Finally, in step 4, we calculated PRM (4). All data processing was performed using MATLAB version 2015b (MathWorks, Inc.).
All subjects’ generated histograms for ΔExp and ΔIns were similar to normal distributions. Of the 90 subjects, 90% were found to have histograms with a t-location scale probability distribution, whereas in the remaining 10%, histograms were generated with a logistic distribution. These functions, like Student $t$, are symmetric about its mean value with a leptokurtic shape. We found that the means of the fitted distribution functions over all 90 subjects and breath-holds were negligibly different from the expected value of 0.

Our procedure for applying a linear regression to the principal component analysis-transformed serial CT data at inspiration and expiration was found to generate consistent results irrespective of GOLD status or ventilation. To illustrate this point, Figure 2 presents voxel HU scatter plots at inspiration and expiration of subjects diagnosed with GOLD 1 and GOLD 3 COPD. Although the location of the voxel distribution varied between the 2 cases, the regression fits (red lines in Figure 2) generated consistent slopes and Y-intercepts (inserts in Figure 2). This consistency was also verified over the entire population (online supplemental Results). As expected, IMV was found to decrease with increasing COPD severity (Figure 2). As the disease progresses, lung parenchymal density approaches ambient air.

To address limitations in HU variability at low density, IMV was modeled as a function of HU. We assumed that the serial pair

---

**Table 2. PRM Results**

<table>
<thead>
<tr>
<th>PRM</th>
<th>Never-Smokers</th>
<th>GOLD 0</th>
<th>GOLD 1</th>
<th>GOLD 2</th>
<th>GOLD 3</th>
<th>GOLD 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 days</td>
<td>98.3 (3.1)</td>
<td>91.2 (17.1)</td>
<td>81.3 (17.9)</td>
<td>69.3 (22.6)</td>
<td>53.6 (24.0)</td>
<td>36.1 (12.9)</td>
</tr>
<tr>
<td>30 days</td>
<td>97.5 (3.1)</td>
<td>93.9 (10.5)</td>
<td>85.4 (10.8)</td>
<td>70.0 (18.8)</td>
<td>52.0 (23.5)</td>
<td>36.1 (15.5)</td>
</tr>
<tr>
<td>fSAD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 days</td>
<td>1.4 (2.7)</td>
<td>8.3 (16.7)</td>
<td>16.0 (15.4)</td>
<td>23.0 (19.0)</td>
<td>33.6 (13.9)</td>
<td>46.7 (7.8)</td>
</tr>
<tr>
<td>30 days</td>
<td>2.3 (3.1)</td>
<td>5.7 (10.2)</td>
<td>12.0 (7.7)</td>
<td>22.4 (13.8)</td>
<td>35.3 (14.4)</td>
<td>46.5 (8.1)</td>
</tr>
<tr>
<td>Emphysema</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 days</td>
<td>0.1 (0.1)</td>
<td>0.18 (0.4)</td>
<td>1.96 (4.2)</td>
<td>6.84 (13.8)</td>
<td>12.26 (13.3)</td>
<td>16.78 (10.7)</td>
</tr>
<tr>
<td>30 days</td>
<td>0.04 (0.03)</td>
<td>0.13 (0.3)</td>
<td>1.74 (3.6)</td>
<td>6.80 (13.8)</td>
<td>12.02 (12.8)</td>
<td>16.97 (10.5)</td>
</tr>
</tbody>
</table>

Percent lung volumes of PRM$^{\text{Normal}}$, PRM$^{f\text{SAD}}$, and PRM$^{\text{Emphysema}}$ over GOLD stages. Values are presented as mean (standard deviation). Spatial alignment of data was not corrected for insufficient ventilation between time points.

Abbreviation: fSAD, functional small airways disease.

---

**Figure 2.** Density scatter plots of voxels with interval HU values acquired at expiration (left) and inspiration (right) are presented for representative GOLD 1 (top) and GOLD 3 (bottom) subjects. The regression lines (red line) and 95% confidence intervals (index of measurement variability [IMV]; black lines) from the fit of the data transformed using the principal component analysis are included in the plots. Values of the slope and Y-intercept of the fit are shown at the upper left corner of each plot, jointly with the value of the IMV.
of modes was an adequate approximation of the centroid (center of mass) for the density distribution observed in the scatter plots. The dependence of IMV on HU values is particularly clear in Figure 3, in which IMV values show a nonlinear drop in value with decreasing HU average of modes. To strengthen the fit of IMV model to the data, expiration (blue markers in Figure 3) and inspiration (red markers in Figure 3) data were pooled. The optimal parameters obtained for the IMV model were \( V = 66.5 \) and \( r = -0.014 \) with a goodness of fit of NRMSE = 0.87 as defined as the normalized root mean squared error (online supplemental Results).

**Application of Correction Approach on Follow-up PRM**

Figure 4 shows representative sections from subjects diagnosed as GOLD 1 and 3 COPD at baseline; these are the same subjects

---

**Figure 3.** Presented is a plot of IMV as a function of the Hounsfield unit (HU) modes, with values averaged over 0 and 30 days, for each patient. A nearly sigmoid-shaped model was nonlinearly fitted to the data from subjects with relative changes in lung volume below 10%. Data for serial inspiration and expiration CT scans were pooled for this analysis. The optimal model for the presented data was: \( \text{IMV}(x) = 66.5 \left[ 1 - e^{-0.014(x+1000)} \right] + 30. \)
shown in Figure 2. In the first case, PRM$_{\text{fSAD}}$ was found to drop by 2.8% at follow-up (Figure 4, top row fourth column) with no correction strategy implemented. Processing the serial CT scans using the strategy outlined in Figure 1 generated a corrected PRM$_{\text{fSAD}}$ that resulted in a drop from baseline of only 0.3% (Figure 4, top row fifth column). Consequent to the low relative volume of emphysema as determined by PRM, negligible benefits were observed when correcting the follow-up PRM. Nevertheless, the effect of our correction is evident locally near the apex of the left lung for subject GOLD 1 (Figure 4 top row black arrow fourth and fifth columns). The subject diagnosed with GOLD 3 COPD showed a decrease in emphysematous lung voxels, going from 28.7% at 0-day to 26.4% at the 30-day CT scan, yielding a short-term drop of 2.3% in emphysema (Figure 4, bottom row fourth column). Applying our correction scheme on the follow-up PRM resulted in a difference in PRM$_{\text{Emph}}$ of only 0.3%. A similar result was observed for PRM$_{\text{fSAD}}$, with a change of 2.8% for the original PRM$_{\text{fSAD}}$ and 0.4% for the corrected PRM$_{\text{fSAD}}$. Upon closer inspection of the bottom right lung, PRM voxel classification (w/o correction) at follow-up varies substantially from the baseline PRM for this subject (Figure 4 bottom row blue arrows).

The distributions of ΔPRM$_{\text{Norm}}$, ΔPRM$_{\text{fSAD}}$, and ΔPRM$_{\text{Emph}}$ separated by GOLD, for both uncorrected and corrected models, are presented in Figure 5. As expected, never-smokers and GOLD 0 subjects showed prevalence of voxels classified as PRM$_{\text{Norm}}$ with an interquartile range (IQR) of <2.4% for ΔPRM$_{\text{Norm}}$ and ΔPRM$_{\text{fSAD}}$ in both never-smokers and GOLD 0. As the disease severity increased toward more fSAD and emphysema, erroneous shifts in PRM classifications were more prevalent. We observed maximum classification variability on GOLD 2 and GOLD 3 subjects, yielding an IQR of >8.4% for ΔPRM$_{\text{Norm}}$ and ΔPRM$_{\text{fSAD}}$ in both never-smokers and GOLD 0. As the disease severity increased toward more fSAD and emphysema, erroneous shifts in PRM classifications were more prevalent. We observed maximum classification variability on GOLD 2 and GOLD 3 subjects, yielding an IQR of >8.4% for ΔPRM$_{\text{Norm}}$ and ΔPRM$_{\text{fSAD}}$. In the case of GOLD 3, we found the largest PRM$_{\text{Emph}}$ mismatch with IQR = 2.3%. Applying our correction strategy...
To assess the agreement in PRM classification, overall percent agreement scores were determined between baseline PRM and the 30-day CT scans either with or without corrections of PRM voxel classification. Figure 6 displays the distribution of the percentage agreement (%agreement) scores for follow-up PRM results over GOLD stages. We found significant differences between the uncorrected and corrected models, both globally ($P < .0001$) and within GOLD ($P < .01$ in all cases). Uncorrected follow-up PRM values were found to generate %agreement that significantly varied with increasing GOLD ($P < .0001$). This trend was not observed for %agreement using the corrected follow-up PRM (Figure 6).

**DISCUSSION**

We propose a strategy that addresses voxel-level measurement variability in serially aligned inspiratory–expiratory paired CT scans that affect voxel classification by PRM. CT data acquired at inspiration and expiration at a 30-day interval were used to quantify the variability of HU values owing to system noise and alignment imperfections during postprocessing. After verifying that the change in voxel HU values from baseline and at follow-up preserves the properties of a normal distribution, we defined an IMV that adjusts as a function of HU measurements. Our correction strategy allows voxel-level classification shifts to occur only when changes in voxel HU values exceed IMV($<\text{HU}>$). We found that voxel-level agreement between uncorrected follow-up and baseline PRM data worsened with disease severity. Our strategy of accounting for system noise diminished this trend in agreement between longitudinal PRM data.

Even with detailed spatial information present in CT imaging, the standard approach for assessing and monitoring disease by this modality has often been limited to calculating whole-lung or large tissue (ie, lobes) measures. Typically based on summary statistics (ie, mean), such as mean airway wall measurements (17) or percent air trapping using only an expiratory CT scan acquisition, variability in longitudinal CT scans has led to erroneous conclusions (18). For example, recently, Smith et al. (19) showed in a comparison of spatially matched airway segments of nonsmokers with a COPD population, that COPD subjects were found to have thinner walls, contrary to the data relying on a more lumped approach. The same whole-lung approach has also been applied to spatially aligned data, such as PRM, where individual classifications are presented as percent-ages of the entire lung volume (4, 20). In response to the findings from McDonough et al. where they postulated that SAD is an intermediate step toward emphysema (21); Boes et al proposed an approach for assessing voxel-level changes in PRM classification, which was referred to as “voxel-based tracking” (5). One-year interval paired CT data from SPIROMICS were spatially aligned to a single geometric frame, in this case the baseline expiration CT scan. As described in this study, each voxel in the lung parenchyma consisted on 2 temporally resolved PRM classifications. In a subject diagnosed with GOLD 2 COPD and found to have a decline in absolute FEV1 from 2.34 L to 2.12 L over 1 year, 48% of all voxels identified by PRM as emphysema at follow-up were fSAD at baseline. The ability to monitor COPD at the voxel-level could provide clinicians with a means to monitor local progression. Although promising, the effect of measurement variability on PRM classification was evident in the whole-lung measures analyzed in that study, where PRM$^{\text{Emph}}$ was found to decrease in a small number of subjects over the 1-year period (5), similar to our finding in Figure 4 for the GOLD 3 subject. Our observations concluded that fSAD measurements from uncorrected follow-up PRM, on average, showed moderate agreement to baseline PRM (Figure 5), even within the relatively short time frame of 30 days. The deterioration in PRM$^{\text{fSAD}}$ agreement with increasing GOLD status is attributed to the large number of voxels with HU values near the $-950$ HU and $-856$ HU thresholds for GOLD 1-3 subjects. Even a small deviation in the HU value between longitudinal CT scans ($<\text{IMV}$) could result in a shift in PRM classification. For example, a voxel may easily shift over $-856$ HU from normal to/from fSAD (Norm ↔ fSAD) or over $-950$ HU from fSAD to/from Emph (fSAD ↔ Emph). Through the proposed approach, we have provided a strategy to mitigate erroneous shifts in voxel-level PRM classification when analyzing longitudinal data by PRM.

Limitations in our approach deserve further attention. Although CT scans from 90 subjects were available for development of the approach, the composition of the population significantly varied across GOLD (Table 1). Quantitative CT values are highly dependent on scanner vendor, scanner type, acquisition parameters (eg, kV, mA, and FOV) and reconstruction kernels. In our study population, a subset of subjects underwent serial CT scans on different scanner types and acquired at different FOVs, with both limitations producing negligible differences in %agreement (details in online supplemental Results). In addition, we did not correct for inadequate ventilation at serial CT examinations (discussion in online supplemental Results). The effect of HU variability, consequent of CT acquisition, processing, and inadequate ventilation, on PRM quantification, has been previously reported, and techniques for alleviating their effect has been discussed (6). Although this approach does not address all errors associated with evaluating serial
CT scans, our strategy for correcting PRM classification shifts is highly adaptable, allowing additional techniques that resolve more specific sources of error to be easily integrated in our workflow.

Consequent to the impact of COPD to health systems worldwide, extensive research is being devoted to the development and evaluation of novel biomarkers. PRM has been shown in multiple studies to serve as an objective and quantitative measure of disease. Large-scale multicenter observational studies such as COPDGene and SPIROMICS provided temporally resolved HRT to evaluate metrics, such as PRM, for monitoring COPD progression and, ideally, for therapeutic response assessment. Our methodology for correcting shifts in PRM classification due to variability in longitudinal HRT scans may improve the clinical management of patients through more accurate monitoring of COPD subtypes.

Supplemental Materials
Supplemental Appendix: http://dx.doi.org/10.18383/j.tom.2017.00013.sup.01

ACKNOWLEDGMENTS
The authors would like to acknowledge the SPIROMICS investigators for providing the computed tomography scans and data used in this study. The Subpopulations and Intermediate Outcomes in COPD Study [SPIROMICS] was supported by contracts from the NIH/NHBI (HHSN26820090013C, HHSN26820090014C, HHSN26820090015C, HHSN26820090016C, HHSN26820090017C, HHSN26820090018C, HHSN26820090019C, HHSN26820090020C), which were supplemented by contributions made through the Foundation for the NIH from AstraZeneca; Bellerophon Therapeutics; Boehringer-Ingelheim Pharmaceuticals, Inc; Chiesi Farmaceutici S.p.A; Forest Research Institute, Inc; GSK; Grifols Therapeutics, Inc; Ikarion, Inc; Nycemed GmbH; Takeda Pharmaceutical Company; Novartis Pharmaceuticals Corporation; Regeneron Pharmaceuticals, Inc; and Sanofi. This work was also supported by the US National Institutes of Health research grants R01HL122438 and R44HL118837.

Disclosure: No disclosures to report.

Conflict of Interest: None reported.

REFERENCES
Can Hyperpolarized $^{13}$C-Urea be Used to Assess Glomerular Filtration Rate? A Retrospective Study

Christian Østergaard Mariager, Per Mose Nielsen, Haiyun Qi, Marie Schroeder, Lotte Bonde Bertelsen, and Christoffer Laustsen

Department of Clinical Medicine, MR Research Centre, Aarhus University, Aarhus, Denmark

Corresponding Author: Christoffer Laustsen, PhD
Palle Juul-Jensens Boulevard
99, 8200 Aarhus N, Denmark;
E-mail: cl@clin.au.dk

Key Words: MRI, hyperpolarization, GFR
Abbreviations: Glomerular filtration rate (GFR), magnetic resonance (MR), single-kidney GFR (skGFR), dynamic contrast-enhanced (DCE), arterial input function (AIF), Baumann–Rudin (BR), ischemia-reperfusion (I/R), magnetic resonance imaging (MRI), renal blood flow (RBF)

INTRODUCTION

Glomerular filtration rate (GFR) measures are essential to the daily care of patients, as either an estimate or an exact quantifiable measure (1). GFR is often estimated by the serum creatinine levels or creatinine clearance, derived from both blood and urine samples. Creatinine estimation is a relative insensitive marker of GFR owing to the GFR-dependent tubular secretion of creatinine (2). Inulin clearance is considered to be the most reproducible, quantitative index of renal function, as it is not reabsorbed and thus transported freely to the urine. However, the specificity is lacking in both methods, as the total GFR can overshadow alterations in single kidney function or even in intrarenal differences (1).

Nuclear medicine-based techniques remain the reference method for quantification of the single-kidney GFR (skGFR) (1); widespread application of these, however, has been limited by the ionizing radiation associated with the examination. Several magnetic resonance (MR)-based methods have emerged as alternative methods to quantify skGFR. Contrast-based methods, such as dynamic contrast-enhanced (DCE) MR, have been used to generate GFR analytical models in both experimental disease and in humans (3-6). Although the methods in general show great promise, the clinical translation is lacking. This may be largely because of the lack of general consensus on model standardization, a direct consequence of the complex system in question and the obtainable signal-to-noise ratio in MR.

Recently, an alternative method for high-signal, contrast-enhanced MR has been introduced. By means of hyperpolarization, contrast-enhanced MR can be enhanced by 4 orders of magnitude. In this technique, the hyperpolarized tracer itself is the origin of the signal, thereby overcoming some of the challenges associated with traditional MR contrast agents. The novel technique of hyperpolarized MR has shown applicability in a broad range of biological applications including cancer, cardiovascular, brain, liver and kidney research (7-9), with the primary goal to interrogate organ-specific metabolic substrate selection associated with various disease states (10, 11). The technique enables the...

In addition to metabolic imaging, several artificial or endogenous tracers have been developed for angiographic and perfusion imaging (19, 18, 5, 20-22). A tracer of particular interest is $^{13}$C-urea and $[^{13}$C,$^{15}$N$_2$]urea, which is an essential osmolite associated with renal function (20). $[^{13}$C,$^{15}$N$_2$]urea possesses particular optimal properties for hyperpolarization, as the $^{15}$N reduces the relaxation loss and increases the T$_2$ at low magnetic fields (23). Urea is vital for the kidneys’ ability to concentrate urine, thereby preventing loss of water and essential nutrients (24). Urine concentration is directly determined by GFR, and thus, the intrarenal dependency of urea distribution in conjunction with renal function has previously been investigated in rodents and in porcine models with hyperpolarized $^{13}$C-urea. This enables assessments of perfusion, osmolality gradients, and relaxation alterations under various functional and disease conditions (20, 23, 25-29).

Here we combine a simple, model-free analysis of renal hemodynamics and a simple, nonarterial input function (arterial input function [AIF]) GFR model, the so-called Baumann–Rudin (BR) model, on data describing the kinetics of hyperpolarized $[^{13}$C,$^{15}$N$_2$]urea handling in the rodent kidney. This indirect model of GFR assumes 2 distinct compartments—cortex and medulla; the cortex and medulla predominately contain blood and urine, respectively (Figure 1). The model assumes a unidirectional transport of contrast from the cortical space to the medullary/pelvic region. Our aim in performing this retrospective data analysis was to determine if hyperpolarized $^{13}$C-urea could be used to estimate GFR in the rodent kidney.

**METHODOLOGY**

The data presented here were derived through retrospective analysis of hyperpolarized $[^{13}$C,$^{15}$N]urea imaging data acquired previously from the kidneys of control rats ($n = 5$), early diabetes rats ($n = 6$) (26), and ischemia-reperfusion (I/R) rats 24 hours after reperfusion ($n = 6$) (30). Originally, the data were analyzed to determine renal perfusion (Figure 2); here, new analytical tools were applied to extract a putative GFR based on hyperpolarized MR data (our calculated GFR will be referred to here as hGFR). To verify the findings in the previously acquired data, 9 additional animals were examined. The additional examinations include accurate hyperpolarized $[^{13}$C,$^{15}$N]-urea T$_1$ relaxation estimation using a single pulsed global NMR experiment ($n = 4$) and a gold-standard inulin clearance GFR estimation accompanying a single-kidney DCE magnetic resonance imaging study.
imaging (MRI) GFR estimation with the BR model and the model-free perfusion model (31).

Animal Handling
Experimental details have been described previously in the original perfusion imaging publications (26, 30). To summarize, $^{[13C,15N]}$urea imaging was performed on similar conditioned female Wistar rats (220 g) 2 weeks after streptozotocin treatment (55 mg/kg) to induce a prediabetic nephropathy model (early signs of renal dysfunction, increased oxygen consumption). In similar conditioned female Wistar rats (220 g), $^{[13C,15N]}$urea imaging was performed 24 hours after reperfusion following severe I/R injury in the left kidney (60 minutes of ischemia) (30). Additional similar conditioned 9 female Wistar rats (220 g) were anesthetized with inactin (120 mg/kg subcutaneously) for evaluation of $^{13C}$ global $T_1$, DCE skGFR, and inulin clearance.

Fluorescein Isothiocyanate-Inulin Clearance
GFR was determined using an intravenous bolus injection of fluorescein isothiocyanate (FITC)-inulin. A solution of 1.5% FITC-inulin was prepared and dialyzed (membrane molecular weight cutoff: 1000). Before injection, the FITC-inulin solution was filtered through a 0.2-μm syringe filter for sterilization. Animals were given an injection of 2 μL/kg. Further, 100 μL blood samples were collected at 1, 3, 5, 10, 15, 35, 55, and 75 minutes. Hereafter the collected plasma fractions were isolated. During the experiments, FITC-inulin were protected from light and kept on ice. Samples were diluted in 1:10 in a HEPES buffer (pH 7.4) and measured in duplicate on a 384-well plate. The original FITC-inulin solution was diluted in 1:100. Analyses were performed on a PHERAstar FS micro plate reader (Em/Ex 485 nm/520 nm; BM Labtech, Birkerod, Denmark). FITC-inulin clearance was analyzed with a noncompartmental pharmacokinetic model (32-34).

Imaging
In both studies, a 2D fully balanced steady-state sequence with (repetition time/echo time/field of view/spectral width/matrix/section thickness) of 4.8 milliseconds/2.4 milliseconds/60 × 60 mm$^2$/20 kHz/32 × 32/10 mm, separated by 3 seconds was used to allow perfusion assessment of the renal hemodynamics (20). The experiments were performed on a 9.4 T (Agilent, Palo Alto, California) horizontal preclinical MRI system, equipped with a $^1$H/$^{13C}$ Litz coil (Doty Scientific, Columbia, South Carolina) for transmission and reception. $^1$H DCE-MRI was performed with similar experimental setup, with a standard gradient spoiled echo sequence with fat suppression (repetition time/echo time/field of view/spectral width/matrix/section thickness of 14 milliseconds/1.8 milliseconds/60 × 60 mm$^2$/50 kHz/128 × 128/2 mm) covering both kidneys with a temporal resolution of 1.75 seconds. Hyperpolarized $^{[13C,15N]}$-urea $T_1$ relaxation estimation was performed with a dynamic series of nonselective spectroscopic acquisitions (repetition time/spectral width/flip angle of 2 seconds/20 kHz/10°).

Hyperpolarization
In both studies, a clinically ready 5 T SPINLAB polarizer was used (35). The samples was prepared by adding a mixed ratio of 200 μL of $^{[13C,15N]}$urea (Sigma-Aldrich, Brøndby, Denmark), glycerol (Sigma-Aldrich, Brøndby, Denmark), and AH111501 (GE Healthcare, Brøndby, Denmark) (6.4 M concentration; 0.30: 0.68:0.02) to a fluid path and placing it in the 5 T SPINLab polarizer (GE Healthcare, Brøndby, Denmark) for more than 2 hours to achieve a reproducible polarization of $>30\%$. The sample was subsequently rapidly dissolved and transferred to the rats already placed in a 9.4 T preclinical MR scanner, with an injection volume of ~1.0 mL (26, 30).

Data Analysis
Renal blood flow (RBF) was estimated by using the model-free formulation by Johansson et al. (5), in which the area-under-the-curve (AUC) ratio between the AIF and the cortical tissue curve is defined as follows (in mL/min per mL cortical tissue):

$$RBF = \frac{\sum AUC_{cortex}}{\Delta t \sum AIF}$$

(1)

where $\Delta t$ represents the interimage delay (here 3 seconds). A correction for the plasma hematocrit, assumed to be 0.45, was used, which is similar to that used by Johansson et al. (5). Before fitting, the signal was smoothed with a lowess filter in the temporal dimension and corrected for $T_1$ relaxation with a $T_1$ relaxation time using a single exponential correction of 24 seconds (global $^{[13C,15N]}$urea $T_1$ relaxation time as measured experimentally; see Results). GFR was estimated by calculating the kinetic rate ($K_{cl}$) of appearance of the signal in the medulla/pelvic region (36).

$$\frac{dC(t)_{medulla}}{dt} = K_{cl} C(t)_{cortex}$$

(2)

The upslope of the curve showing $[^{13C,15N}]$urea signal in the medulla was estimated (Figure 3), in MATLAB (The MathWorks, Inc., Natick, Massachusetts), between the initial point of the cortical slope and the peak of the medullary slope (gray area in Figure 3). GFR (estimated $K_{cl}$) was then obtained by dividing the medullary slope with the mean renal cortex concentration during the upslope period (37). The GFR was expressed in milliliter per minute, to allow for comparison with previously reported values for skGFR and total GFR (38). We assumed a cortical and medullary tissue density of 1 for the conversion of the perfusion and GFR values.

Statistics
Normality was assessed with quantile–quantile plots. A $P$-value $<.05$ was considered statistically significant. Statistical analysis was performed using GraphPad Prism (GraphPad Software, Inc., La Jolla, California). A 2-way paired ANOVA (left and right kidney paired) was used for statistical analysis of the renal perfusion and the GFR estimations; a post hoc Sidak multiple comparisons correction was used when appropriate. An unpaired Student $t$ test was used for statistical analysis of the total GFR and the plasma creatinine concentration between the I/R group and the control group.

RESULTS
A $^{[13C,15N]}$urea $T_1$ relaxation was found to be 24.5 ± 4 seconds ($n = 4$, in vivo at 9.4 T), allowing $T_1$ correction of the hemodynamic acquisitions. A significant renal blood flow variation was
observed among the 3 groups (P = .018), with a tendency toward an increased RBF in the I/R group.

No significant group difference was observed between control $^1$H DCE-derived RBF and control hyperpolarized $[^{13}C,^{15}N]$-urea RBF estimations ($P = .23$) (Table 1). A significant variation in hGFR was observed among the individual kidneys ($P = .02$), with a significant difference among the groups (interaction term group $\times$ kidney, $P = .02$) originating from a reduction to skGFR within the I/R group with a difference of $-1.6\ mL/min$ ($P = .005$) between each animal’s 2 kidneys. No difference was seen between the control group ($P = .99$) and the diabetes group ($P = .99$) or within these groups.

### Table 1. Hemodynamic and Physiological Parameters from $^1$H DCE and FITC-Inulin

<table>
<thead>
<tr>
<th></th>
<th>RBF Left Kidney (mL/min/mL Tissue)</th>
<th>RBF Right Kidney (mL/min/mL Tissue)</th>
<th>Body Weight (g)</th>
<th>Kidney Weight (g)</th>
<th>Cortical Weight (g)</th>
<th>Total GFR (DCE) (mL/min)</th>
<th>GFR (inulin) (mL/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.02</td>
<td>4.71</td>
<td>256</td>
<td>0.8</td>
<td>0.75</td>
<td>5</td>
<td>3.8</td>
</tr>
<tr>
<td>2</td>
<td>4.14</td>
<td>3.47</td>
<td>235</td>
<td>0.75</td>
<td>0.7</td>
<td>4.7</td>
<td>7.3</td>
</tr>
<tr>
<td>3</td>
<td>6.17</td>
<td>5.92</td>
<td>266</td>
<td>0.8</td>
<td>0.72</td>
<td>4.4</td>
<td>—</td>
</tr>
<tr>
<td>4</td>
<td>6.78</td>
<td>6.99</td>
<td>250</td>
<td>0.81</td>
<td>0.74</td>
<td>1.2</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>3.38</td>
<td>1.9</td>
<td>222</td>
<td>0.68</td>
<td>0.58</td>
<td>2.3</td>
<td>2.3</td>
</tr>
<tr>
<td>Mean ± SD</td>
<td>4.9 ± 1.3</td>
<td>4.6 ± 1.8</td>
<td>245.8 ± 15.6</td>
<td>0.77 ± 0.05</td>
<td>0.7 ± 0.06</td>
<td>3.5 ± 1.5</td>
<td>5.3 ± 2.4</td>
</tr>
</tbody>
</table>

Abbreviations: RBF, renal blood flow; GFR, glomerular filtration rate; DCE, dynamic contrast-enhanced.
A filtration fraction of ~40% was found in both early diabetic rats (44.8 ± 9.5%) and healthy controls (42.8 ± 13.7%), whereas the filtration fraction reduced in the I/R group to 25.1 ± 12.2%. To evaluate the accuracy of the GFR estimation, the MR-derived “total hGFR” (sum of both kidneys) was compared with the plasma creatinine concentration between the controls [normal animals (26) pre-I/R (30)] of 5.1 ± 2.6 μmol/L and the I/R group of 3.4 ± 1.4 μmol/L (30). The total hGFR was not significantly different between the control and the I/R group (unpaired t test, P = .28), albeit an increased plasma creatinine concentration associated with I/R 24 hours after reperfusion was observed (unpaired t test, P = .0045) (Figure 4). Furthermore, no statistical significant difference was found between 3H GFR and inulin measurements (paired t test, P = .33). In addition, no statistical difference was observed between any combinations of total GFR estimation (1-way ANOVA, P = .4) (Figure 4E).

**DISCUSSION**

The main finding in this study is the proof-of-concept that first-pass hyperpolarized [13C,15N]urea transport can be used to estimate GFR. GFR findings by use of hyperpolarized MR showed good agreement with gold-standard inulin clearance values and with 3H DCE GFR values found under similar conditions and in congruence with previously reported values in similarly conditioned female rats, measured via the gold-standard inulin clearance (32) and with DCE-MRI (38).

Further, our estimated hGFRs compare well (1–3 mL/min lower) with human values, which are known to be on the order of 120 mL/min/70 kg, following conversion to rodent values by the allometric relationship described by Rhodin et al. (39). The allometric GFR was estimated by using the exponent range 2/3–3/4 (1.6 –2.6 mL/min) (39). As previously reported (30), rats subjected to 60 min of ischemia and 24 h of reperfusion showed unaltered urine output. These data support our findings of a substantial reduction to single-kidney hGFR, whereas the total hGFR was only slightly reduced (indicated by increased plasma creatinine concentration and largely maintained hGFR, owing to residual function of the contralateral kidney). Interestingly, the data presented here suggest a compensatory effect on RBF in the contralateral kidney not subjected to I/R, although this trend did not reach significance (Figure 4). It is important to note that the inverse correlation between plasma creatinine concentration and I/R damage was not apparent in the paired experiments from the original study (30). This might be explained by the additional animals included here in the control group (n = 11) (5 from the study by Qi et al., 26, and 6 pre-IR from the study by Nielsen et al., 30) compared with the I/R group (n = 6).

Several limitations are apparent in this study. First, GFR estimation requires high spatiotemporal resolution. The interimage delay of 3 seconds reduced the accuracy of the time curve estimation and thus the fitting of the upslope of the medulla time curve. A similar effect is observed in 3H GFR methods (40), and it can be largely solved by increasing the temporal resolution when acquiring the hyperpolarized images. However, the available signal must be taken into consideration, as it is limited by the hyperpolarized radio frequency (RF) signal depletion,
with a lower effective $T_1$ seen in the imaging section ($T_{1\text{eff}} = 19 \pm 3$ seconds, estimated from the bSSFP images) owing to the imaging acquisition. It is difficult to compensate for the RF depletion (estimated to be 67% in these experiments; 41), as the imaging section is replenished by flowing spins into the imaging section. This is particularly important for the bolus differentiation-perfusion assessment (5), potentially reducing the accuracy of the assessment, as the acquisition did not fully saturate the signal between images. The spatial resolution of $1.9 \times 1.9 \text{ mm}^2$ is a limiting factor as well, as shown in Figure 3. Here, the medullary signal was contaminated by the cortical signal owing to partial volume effects, thereby reducing the accuracy of the method. Furthermore, it should be stressed that because of the significant reabsorption of urea, it is likely that that urea estimate GFR is apparent by nature. Furthermore, the retrospective use of data and comparison with other methods (DCE and inulin) in additional groups is a limitation of the study. Further studies are needed to fully determine the observed correlation between true GFR and the estimated hGFR.

We selected the BR model because of its ease of implementation, its lack of reliance on AIF sampling, and the need for estimating only the upslope of the signal, removing the need to sample beyond the $T_1$ relaxation decay. More complex models often depend on rapid and accurate sampling of the AIF and thus are particularly sensitive to appropriate placement of the imaging section and partial volume effects of the intense signal from blood. The imaging section is typically a 1-cm mean intensity profile slab (permitted by the lack of background signal in $^{13}\text{C}$ MR), containing kidneys, aorta, and vena cava. However, the simplicity of the BR model also presents limitations, namely, it assumes 2 distinct separated volumes (blood and urine), when in reality, both the cortical and medullary compartments contain blood and urine (3, 36). In future, advanced AIF sampling schemes may enable the use of more sophisticated models to improve the hGFR estimation. Importantly, the current knowledge on the relaxation behavior of the hyperpolarized $[^{13}\text{C},^{15}\text{N}_2]$urea tracer ($T_1 = 24$ seconds at 9.4 T) limits the correction to a global $T_1$ correction. In future, more appropriate relaxation models, both $T_1$ and $T_2$, could be incorporated that take compartmentalized relaxation properties into account (23, 28, 42, 43). Interestingly Reed et al. (23) have shown intrarenal compartmentalized $T_2$ relaxation behavior at 3 T contrary to the reported $T_2$ relaxation times at 9.4 T, finding only 1 $T_2$ component (28). This implies that this would be particularly important at 3 T and supported by using the novel $T_1$ and $T_2$ mapping sequences (23) for accurate GFR assessment.

Finally, the use of hyperpolarized $[^{13}\text{C},^{15}\text{N}_2]$urea could potentially give rise to variations in GFR estimation, as 50% of urea is reabsorbed (24). Thus, alternative molecules such as creatinine, which are reabsorbed to a lesser degree, could potentially improve the hyperpolarized MR-based GFR estimation. Although, while typical GFR estimations are performed with free filtered-tracers (32, 2), the reabsorption of urea could represent a potential advantage over existing methods by allowing simultaneous estimation of the reabsorption (23, 25, 26, 20, 30, 29, 27). Furthermore, it has been demonstrated that several perfusion tracers can be hyperpolarized and imaged simultaneously (44), allowing more detailed knowledge on the filtration and reabsorption by combining biomarkers with different hemodynamic profiles (21, 45).

CONCLUSION

In conclusion, this study shows that hyperpolarized MR is a promising method for functional imaging of the kidneys. The study found that the estimated $^{13}\text{C}$-urea GFR was in good agreement with GFR calculated from inulin clearance and DCE MRI, as well as plasma creatinine measurements and literature findings. Future work to optimize MR data acquisition schemes and to quantitatively evaluate this approach is warranted.


High-Resolution MR Imaging of Muscular Fat Fraction—Comparison of Three T2-Based Methods and Chemical Shift-Encoded Imaging

Lena Trinh¹, Emelie Lind¹,², Pernilla Peterson¹, Jonas Svensson¹,³, Lars E. Olsson¹, and Sven Månsson¹

¹Medical Radiation Physics, Department of Translational Medicine, Lund University, Skåne University Hospital, Malmö, Sweden; ²Department of Medical Radiation Physics, Lund University, Lund, Sweden; ³Medical Imaging and Physiology, Skåne University Hospital, Lund, Sweden

Corresponding Author:
Lena Trinh, MSc
Medical Radiation Physics, Inga Marie Nilssons gata 49, Skåne University Hospital, SE-205 02 Malmö, Sweden; E-mail: lena.trinh@med.lu.se

Key Words: fat quantification, Bayesian probability theory, non-linear least squares, T₂, high-resolution imaging

Chemical shift-encoded imaging (CSEI) is the most common magnetic resonance imaging fat–water separation method. However, when high spatial resolution fat fraction (FF) images are desired, CSEI might be challenging owing to the increased interecho spacing. Here, 3 T₂-based methods have been assessed as alternative methods for obtaining high-resolution FF images. Images from the calf of 10 healthy volunteers were acquired; FF maps were then estimated using 3 T₂-based methods (2- and 3-parameter nonlinear least squares fit and a Bayesian probability method) and CSEI for reference. In addition, simulations were conducted to characterize the performance of various methods. Here, all T₂-based methods resulted in qualitatively improved high-resolution FF images compared with high-resolution CSEI. The 2-parameter fit showed best quantitative agreement to low-resolution CSEI, even at low FF. The estimated T₂-values of fat and water, and the estimated muscle FF of the calf, agreed well with previously published data. In conclusion, T₂-based methods can provide improved high-resolution FF images of the calf compared with the CSEI method.

INTRODUCTION
Chemical shift-encoded imaging (CSEI) is a common quantitative magnetic resonance imaging (MRI) method for fat–water separation and measurement of fat content in numerous body parts, such as the liver and skeletal muscles (1–5). In skeletal muscles, fatty infiltration has been related to, for example, insulin resistance and various neuromuscular diseases (6–11). The location of fat accumulation within the muscle has also been shown to be important (6), as some muscle groups are more likely to accumulate fat (12). Depending on the muscle group involvement, the outcome of some neuromuscular diseases can show a large variability (11, 13). In addition, different neuromuscular diseases show different fat infiltration patterns of the muscle groups. By detecting these patterns, it might be easier to identify a specific disease (11, 14). To enable and simplify the distinction between the different muscle groups, and between inter- and intramuscular fat, high-resolution fat fraction (FF) images are desirable. CSEI is a validated method for fat quantification purposes (4, 15), and it has previously been used for skeletal muscle applications (1, 2, 5).

Previously, fat quantification methods based on differences in fat and water T₂ (16) rather than chemical shifts have been suggested for applications in skeletal muscles (17, 18). With T₂-based methods, there is a possibility of obtaining information on FF and T₂ relaxation times simultaneously (18). This would offer more information about the status of the disease, as a change in muscle T₂-relaxation time has been shown to reflect the activity and progress of neuromuscular diseases (13, 19), complementing the information about the fat infiltration degree that primarily serves as a severity indicator (14). Moreover, there are several challenges associated with the CSEI technique, particularly when high resolution is required, which may be addressed by using T₂-based methods. For example, increasing the resolution increases the minimal achievable interecho time which may have a negative impact on the CSEI fat quantification accuracy (20). In addition, it is common that fat/water swaps are present in FF images when using CSEI.

To obtain both the amplitudes and the T₂-relaxation times of the fat, as well as the water component of the signal, a nonlinear least squares (NLLS) fitting method is commonly used (21, 22). However, NLLS has known problems with estimating the parameters correctly when 1 component is considerably larger than the other (23, 24). As a consequence, it may be difficult to measure low FFs using NLLS. In such cases, a fitting method based on Bayesian probability theory could be an alternative, as it has also been shown to be more robust against noise.
compared with NLLS (25). Bayesian fitting models have been proposed for other MRI applications such as intra-voxel incoherent motion imaging (26) and myelin water fraction estimations (27), but have, to the best of our knowledge, not yet been evaluated for fat quantification purposes.

The aim of this study was to examine the accuracy and noise performance of 3 different T2-based fat quantification methods, using high-resolution MRI for low FFs in healthy volunteers, and compare it with CSEI. The first T2-based method uses fixed T2-relaxation times of both water and fat as described by Kan et al. (17). The second method uses only a fixed fat T2-relaxation time to study the possibility of obtaining a simultaneous T2 map of water. The third method is based on Bayesian probability theory as described by Barbieri et al. (26), in which neither relaxation time is fixed. In this study, the muscular FF was measured in the calf of healthy volunteers, and simulations were made to study possible biases in the estimation of the FF using the T2-based methods.

**METHODOLOGY**

**Subjects**

In total, 10 healthy volunteers, 3 males (mean age, 28 years; range 25–30 years) and 7 females (mean age, 28 years; range, 24–32 years), were recruited and scanned with the approval from the regional ethical board. Informed consent was obtained from all volunteers.

**MRI Data Acquisition**

All measurements were acquired using a 3 T scanner (MAGNETOM Trio, Siemens Healthineers, Erlangen, Germany) and a 6-element body matrix coil. All data were obtained at 2 matrix sizes, 128 × 128 and 512 × 512, keeping the field of view constant at 280 × 280 mm² and thus acquiring data at low- and high-spatial resolution. A single 6-mm transversal slice was collected for each acquisition, centered at the widest part of the left calf of each volunteer.

A multi-echo gradient echo (MGRE) sequence with 6 echoes was used for the CSEI method. To avoid T1 bias, a long repetition time (TR = 500 milliseconds) and a small flip angle (12°) were used. By estimating the number of signal averages as a function of interecho spacing, echo times (TEs) were chosen to obtain as small interecho spacing for the highest number of signal averages value as possible. The bandwidth (BW) was then set as low as possible without affecting the interecho time. In this way, a minimal interecho time with a high noise performance was ensured. With the first TE set to the shortest possible, the following parameters were used: TE1/ΔTE = 1.11/1.56 milliseconds (low resolution), TE1/ΔTE = 2.57/3.92 milliseconds (high resolution), BW = 1776 Hz/px (low resolution), and BW = 651 Hz/px (high resolution). Acquiring 1 average, the scanning times were 1 minute 6 seconds (low resolution) and 4 minutes 18 seconds (high resolution). From 1 subject, additional high-resolution CSEI images were collected with 2, 3, and 9 averages that had scanning times of 8 minutes 34 seconds, 12 minutes 50 seconds, and 38 minutes 29 seconds, respectively.

For the T2-based methods, 32 multi-echo spin echo (MESE) images were acquired with a 180° refocusing pulse and the following settings: TR = 2000 milliseconds, ΔTE = 9.2 milliseconds (low and high resolution), BW = 425 Hz/px (low resolution), BW = 391 Hz/px (high resolution), and number of averages = 1. To avoid long acquisition times, parallel imaging (GRAPPA) was used with an acceleration factor of 2. The resulting scan times were 2 minutes 36 seconds (low resolution) and 9 minutes (high resolution).

**Fat/Water-Separation Methods**

The methods used in this study are summarized in Figure 1. All calculations were performed using MATLAB (r2017a, The MathWorks, Inc., Natick, MA).

**Chemical Shift-Encoded Imaging**

The FF was calculated using a complex and magnitude-based iterative multiecho water-fat separation algorithm (28), with a
Using 6 echoes (30), the FF was calculated using the following equation:

\[
\text{Fat fraction} = 100 \cdot \text{real} \left( \frac{F}{W + F} \right) \tag{1}
\]

where the complex-valued \( F \) and \( W \) are the estimated fat and water signals, respectively. \( T_1 \) bias was avoided by using a low flip angle acquisition.

**T2-Relaxation Time-Based Imaging**

Two of the \( T_2 \)-based methods use a fixed \( T_2 \)-relaxation time of fat \( (T_{2,F}) \), of which 1 uses a fixed \( T_2 \)-relaxation time of water \( (T_{2,W}) \) as well. To obtain these values, a monoexponential fit of the signal decay was carried out voxel by voxel, resulting in a \( T_2 \) map. For each volunteer, individual \( T_{2,F} \) and \( T_{2,W} \) values were then calculated as the mean value within corresponding regions of interest (ROIs), which were drawn in subcutaneous fat and muscle tissue, respectively (Figure 2). The ROI of fat was drawn to include as much of the subcutaneous fat as possible, avoiding visible blood vessels. In 1 volunteer, the subcutaneous fat layer was too thin for ROI definition. For this volunteer, the mean \( T_{2,F} \) of the rest of the volunteers was calculated and used instead. The muscle ROI was drawn in a small part of tibialis anterior without any visible fat to minimize fat bias in the estimation of \( T_{2,W} \). Echoes 2–16 were used for all estimations using MESE data. The first echo was excluded owing to stimulated echo effects present in all other echoes, whereas the last echoes were excluded to reduce noise bias.

**Two-Parameter Fit—Fixed \( T_{2,F} \) and \( T_{2,W} \).** Using the estimated \( T_{2,F} \) and \( T_{2,W} \) values from the monoexponential fit, the amplitudes of water \( W \) and fat \( F \) could be calculated by a simple linear regression, as described by Kan et al. (17). The signal model is given by using the following equation:

\[
S(t) = We^{-\frac{t}{T_{2,W}}} + Fe^{-\frac{t}{T_{2,F}}} \tag{2}
\]

where \( S \) is the measured signal at TE \( t \), and \( T_{2,F} \) and \( T_{2,W} \) are kept fixed.

**Three-Parameter Fit—Fixed \( T_{2,F} \).** Using the same signal model [equation (2)] as in the 2-parameter fit and fixed \( T_{2,F} \) value, \( T_{2,W} \), \( W \), and \( F \) were estimated using a trust region–based NLLS fitting algorithm.

**Bayesian Fitting Method.** An alternative to exponential fitting is using a Bayesian probability method (31). Here, all four parameters \((T_{2,W}, T_{2,F}, W, \text{and} \ F)\) are estimated simultaneously using the method described by Barbieri et al. (26) using the MATLAB function slicesample. The signal model is given by the following equation:

\[
S = S_0 (1 - f) e^{-\frac{t}{T_{2,W}}} + fe^{-\frac{t}{T_{2,F}}} \tag{3}
\]

where \( S_0 \) denotes the signal at \( t = 0 \) and \( f \) denotes the FF in the range \([0, 1]\). To obtain \( S_0 \), linear regression was performed on linearized data, \( \ln(S) \), in each voxel. However, owing to the biexponential form of the signal decay, \( \ln(S) \) is not linear. To compensate for this, \( \ln(S) \) was weighted by the signal amplitude \( S \), making the fit rely mostly on the earlier echoes of the signal. Water and fat amplitudes, \( W = S_0 (1 - f) \) and \( F = S_0 f \), respectively, were calculated before correcting for \( T_1 \) bias and calculating FF as described by equation (4).

**Fat Fraction Calculation and \( T_1 \)-Correction**

Owing to the long \( T_1 \)-relaxation time of muscle tissue and the desire to keep the acquisition times feasibly low, all the \( T_2 \)-based fat quantification methods described in the above sections were corrected for \( T_1 \)-relaxation bias. The \( T_1 \)-relaxation times \( T_{1,W} = 1420 \) milliseconds and \( T_{1,F} = 371 \) milliseconds (16) were used to correct the water and fat signal amplitudes according to \( F_{1corr} = F[1 - \exp(-TR/T_{1,F})] \) and \( W_{1corr} = W[1 - \exp(-TR/T_{1,W})] \), respectively. Hence, the FF can be described using the following equation:

\[
\text{Fat Fraction} = 100 \cdot \frac{F_{1corr}}{W_{1corr} + F_{1corr}} \tag{4}
\]

Because the MGRE data were collected with a low flip angle, no correction for \( T_1 \) bias was needed for CSEI.

**Data Analysis**

To compare the 4 methods, 3 ROIs were drawn in the calf muscles of all 10 volunteers following the outlines of tibialis anterior, soleus, and gastrocnemius (Figure 2). Small areas with fat–water swaps in the high-resolution FF images calculated with CSEI were excluded from the ROIs. If the fat–water swap extended over a large area covering most of the muscle such that no swap-free ROI could be defined, the entire muscle group was excluded from further analysis.

Mean signal-to-noise ratios (SNRs) of the collected MGRE and MESE magnitude images were calculated as \( \text{SNR} = 0.655 \cdot S/\sigma \) where 0.655 is due to the Rayleigh distribution of the noise in magnitude images (32) and \( \sigma \) is the standard deviation of the background noise. The SNR of both subcutaneous fat and muscle tissue was calculated. To calculate the standard deviation of the background noise of the MESE data, the ROIs were placed near the edge of the images where the \( g \)-factor was expected to be close to 1.

Wilcoxon signed-rank tests and Bland–Altman analysis were performed to compare the estimated FFs within the ROIs.
using the 2-parameter fit, 3-parameter fit, Bayesian fit, and high-resolution CSEI, to the FFs calculated with low-resolution CSEI.

Simulations
Simulations were conducted to investigate the effects of incorrect $T_2$-relaxation times, of incorrect signal amplitude, and of noise on the calculated FF. In all simulations, a biexponential model [equation (2)] was used to describe the signal decay using $T_{2,W} = 40$ milliseconds and $T_{2,F} = 160$ milliseconds as true $T_2$-relaxation times. Signals from 5 different FFs (2%, 5%, 10%, 30%, and 95%) were simulated, each with 20 echoes. The signal amplitude at $t = 0$ was set to 1.

To study the effect of inaccurate $T_2$-relaxation times, simulations were performed by using incorrect $T_{2,W}$ and $T_{2,F}$ in the 2-parameter fit method and incorrect $T_{2,F}$ in the 3-parameter fit method. $T_{2,W}$ was set to vary between 22 and 42 milliseconds and $T_{2,F}$ was set to vary between 70 and 260 milliseconds. No noise was added to the signal.

In the Bayesian fitting method, the effect of using an inaccurate $S_0$ value was studied by varying the $S_0$ value between 0.8 and 1.2. No noise was added, and each calculation was carried out 1000 times.

The effect of noise was studied in all 3 $T_2$-based methods by altering the SNR of the simulated signal. The true $T_2$-relaxation times were used to generate a noise-free signal. Complex Gaussian noise was then added to the signal before calculating the magnitude value. The effect was studied at 5 different SNR levels (20, 50, 150, 300, and 600), defined at $t = 0$. Each simulation was carried out 1000 times.

RESULTS
Volunteer Study
The estimated mean $T_2$-relaxation times and standard deviations of muscle (tibialis anterior) and fat (subcutaneous fat), using the monoexponential fit, the 3-parameter fit, and the Bayesian fit are presented in Figure 3. The 3-parameter fit estimated a lower value of $T_{2,W}$ compared with the monoexponential fit and the Bayesian fit. The estimated $T_{2,F}$ from all three methods were independent of matrix size.

Example FF images of all 4 methods can be seen in Figure 4. In contrast to the $T_2$-based methods, the high-resolution CSEI produced an FF image with a noise level that concealed the anatomy of the calf. Although all 3 $T_2$-based methods produced FF images in which the different muscles were distinguishable, the estimated FFs were different between the methods. Because the high-resolution CSEI images with a single average (Figure 4) had a low SNR, additional high-resolution MGRE images were acquired with more averages from 1 volunteer (data not shown). Although SNR naturally increased with the number of averages, the noise level was still obscuring the anatomy of the muscles when using 9 averages.
Scatter plots and Bland–Altman plots of the methods are presented in Figure 5. The linear regression parameters and corresponding confidence intervals are shown in Table 1. Owing to fat–water swaps in the estimated FF images using high-resolution CSEI, results from 3 volunteers were excluded. Compared with the low-resolution CSEI method as reference, the 2-parameter fit was able to estimate the muscle FF accurately, showing only a small overestimation of FFs >3%. High-resolution CSEI overestimated the lower FFs and underestimated the higher FFs of the muscles, whereas the 3-parameter fit consistently overestimated the FF. The Bayesian fitting method showed an underestimation that increased with the FF. In Table 1, the mean values and the standard deviations of the estimated FFs of gastrocnemius, soleus, and tibialis anterior and the corresponding P-values of all volunteers and image resolutions are presented. All calculated mean FFs obtained from the 3-parameter fit, at both high and low resolution, significantly (P < .05) overestimated the FFs obtained from the reference method in comparison with the 2-parameter fit in which no significant differences were found.

In Figure 6, the acquired signal decay of 3 voxels of low- and high-resolution MESE images and the fitted curves of the 3 $T_2$-methods are depicted. All 3 methods performed equal at high FF, whereas at lower FFs (~17% and 3%), the estimated signals differ. The 3-parameter fit results in a slower decaying signal compared with the other 2 methods, whereas the Bayesian method results in a faster decaying signal.

The mean SNRs of the single average MESE (second echo)/MGRE (first echo) images of all volunteers were 919/250 (low resolution, muscle), 2048/208 (low resolution, fat), 214/71 (high resolution, muscle), and 449/63 (high resolution, fat). Because the SNR varies over the MESE images owing to parallel imaging, these values represent SNR when the g-factor is close to 1.

Simulations
The simulated effect on the estimation of FF when using incorrect $T_{2,W}$ and $T_{2,F}$, respectively, is shown in Figure 7. In both cases, an underestimation of $T_{2,W}$-relaxation time resulted in an overestimation of the FF, whereas an overestimation of $T_{2,F}$-relaxation time resulted in an underestimation of FF. The 3-parameter fit was more sensitive to errors in $T_{2,F}$ compared with the 2-parameter fit. At higher FFs, it was more important that $T_{2,F}$ was estimated correctly, whereas a correct $T_{2,W}$ was more important at low FFs. For the Bayesian fit, the effect of using an incorrect $S_0$ value, as well as the standard deviation of the estimated FF, is shown in Figure 8. Using an underestimated $S_0$ value resulted in an underestimation of the FF, and an overestimated $S_0$ value resulted in an overestimated FF. However, the effect of using an overestimated $S_0$ value was greater than that using an underestimated one. Lower FFs (2%–5%) were less sensitive for incorrect $S_0$ compared with higher FFs (10%–30%). This can also be seen by looking at the standard deviation that was larger for higher FFs.

In Figure 9, the simulated effect of noise is shown as the difference between the estimated and true FFs and the standard deviation of the estimated FF of each of the 3 $T_2$-based methods. The accuracy of both the 2- and 3-parameter fits increased with SNR (except for FF = 95% using the 3-parameter fit). The 2-parameter fit was less sensitive to noise than the 3-parameter fit. The Bayesian fit was more affected by noise at higher SNR compared with the NLLS-based methods. As the SNR increased, the standard deviation decreased for all methods and FFs.

**DISCUSSION**
In this work, 3 $T_2$-based approaches (2-parameter fit, 3-parameter fit, and a Bayesian probability method) have been studied...
and compared with CSEI in terms of their capability to correctly estimate low FFs using high-resolution images. This was carried out through a study on healthy volunteers and by simulations. All T$_2$-based methods provided high-resolution FF images in which it was possible to delineate the different muscles of the calf. With respect to the estimated FFs, the 2-parameter fit showed best agreement to the reference method.

Even though the measured mean FF of high-resolution CSEI in vivo corresponded well with the low-resolution CSEI, the high noise level of the images made it impossible to differentiate any anatomy within the muscle fascia, and therefore, it was difficult to locate potential fat accumulations of the muscles. After acquiring 9 averages, thus increasing SNR, it was still not possible to separate the different muscles of the calf. Using an even larger number of averages to increase SNR further could increase the possibility of differentiating the different muscles groups, but this would result in infeasibly long acquisition times. SNR may also be increased by using a larger flip angle (33). This would, however, require a correction for the T$_1$-bias, including an estimation of the true flip angle map, and it would therefore introduce an additional source of error. As low-resolution CSEI was used as a reference method, this approach was thus not used. Another way to increase SNR is to acquire 3D MGRE data. In this study, we chose to collect 2D MGRE images for comparison with 2D MESE images. Because only 1 slice was used as a reference method, this approach was thus not possible, a 2D acquisition allowed for a longer TR and a larger flip angle compared with a 3D acquisition of the same total scan time.

### Table 1. Mean Estimated Fat Fraction and Standard Deviation Between Volunteers, Using All 4 Methods, in Gastrocnemius, Soleus, and Tibialis Anterior Muscles

<table>
<thead>
<tr>
<th></th>
<th>CSEI</th>
<th>2-Parameter Fit</th>
<th>3-Parameter Fit</th>
<th>Bayesian Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>128 × 128</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fat fraction (%)</td>
<td>3.39 ± 0.97</td>
<td>2.70 ± 1.95</td>
<td>5.13 ± 1.60</td>
<td>1.99 ± 0.59</td>
</tr>
<tr>
<td>Gastrocnemius</td>
<td>(P = 0.43)</td>
<td>(P = 9.1 · 10$^{-3}$)</td>
<td>(P = 1.3 · 10$^{-3}$)</td>
<td></td>
</tr>
<tr>
<td>Soleus</td>
<td>4.15 ± 1.16</td>
<td>4.48 ± 1.03</td>
<td>6.22 ± 0.98</td>
<td>2.32 ± 0.37</td>
</tr>
<tr>
<td>Tibialis anterior</td>
<td>1.80 ± 0.65</td>
<td>1.32 ± 0.67</td>
<td>4.02 ± 0.44</td>
<td>1.5 ± 0.15</td>
</tr>
<tr>
<td>Linear regression parameters</td>
<td>-</td>
<td>-0.62</td>
<td>2.1</td>
<td>0.91</td>
</tr>
<tr>
<td>Intercept</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.62</td>
</tr>
<tr>
<td>Slope</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.92</td>
</tr>
<tr>
<td>R$^2$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.75</td>
</tr>
<tr>
<td>Bland–Altman</td>
<td>-</td>
<td>-0.32 (−2.5, 1.9)</td>
<td>1.9 (0.44, 3.3)</td>
<td>−1.2 (−3.0, 0.68)</td>
</tr>
<tr>
<td>Mean (limits of agreement)</td>
<td>-</td>
<td>-0.32 (−2.5, 1.9)</td>
<td>1.9 (0.44, 3.3)</td>
<td>−1.2 (−3.0, 0.68)</td>
</tr>
<tr>
<td><strong>512 × 512</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fat fraction (%)</td>
<td>3.30 ± 0.58</td>
<td>4.00 ± 2.03</td>
<td>6.75 ± 1.52</td>
<td>2.51 ± 0.60</td>
</tr>
<tr>
<td>Gastrocnemius</td>
<td>(P = 0.63)</td>
<td>(P = 0.52)</td>
<td>(P = 2.5 · 10$^{-4}$)</td>
<td>(P = 9.1 · 10$^{-3}$)</td>
</tr>
<tr>
<td>Soleus</td>
<td>3.38 ± 0.56</td>
<td>4.42 ± 1.20</td>
<td>6.81 ± 0.95</td>
<td>2.51 ± 0.39</td>
</tr>
<tr>
<td>Tibialis anterior</td>
<td>2.54 ± 0.62</td>
<td>1.71 ± 0.68</td>
<td>4.79 ± 0.56</td>
<td>1.68 ± 0.21</td>
</tr>
<tr>
<td>Linear regression parameters</td>
<td>-</td>
<td>-0.19</td>
<td>3.4</td>
<td>1.1</td>
</tr>
<tr>
<td>Intercept</td>
<td>-</td>
<td>-0.19</td>
<td>3.4</td>
<td>1.1</td>
</tr>
<tr>
<td>Slope</td>
<td>0.29</td>
<td>1.1</td>
<td>0.87</td>
<td>0.35</td>
</tr>
<tr>
<td>R$^2$</td>
<td>0.34</td>
<td>0.79</td>
<td>0.76</td>
<td>0.73</td>
</tr>
<tr>
<td>Bland–Altman</td>
<td>0.18 (−2.0, 2.4)</td>
<td>0.24 (−1.3, 1.8)</td>
<td>3.0 [1.7, 4.3]</td>
<td>−0.91 (−2.7, 0.90)</td>
</tr>
</tbody>
</table>

$P$-values are given for the comparison against low resolution CSEI. Confidence intervals (CI) are given at a significance level of 0.05.
Problems with fat–water swaps occurred in some of the estimated FF images of the high-resolution CSEI owing to the long TE needed. All high-resolution CSEI data from 3 subjects had to be excluded because of this. The interecho time of high-resolution CSEI may be reduced by means of bipolar or interleaved data acquisition. However, these approaches are associated with problems with phase errors (28, 34–37). Thus, a single acquisition monopolar readout was chosen to avoid any bias of our reference method.

The estimated T2-relaxation times of muscle tissue and subcutaneous fat in this study correspond well with T2-relaxation values of healthy volunteers from literature [T2, W 32–40 milliseconds and T2, F 133–154 milliseconds (16, 17, 38–40)], although the Bayesian fitting model results in a slightly overestimated T2, F. The measured FFs also agree with previously published data, that is, measured FFs ranging between 1.2% and 3.6% (tibialis anterior), 2.5% and 3.3% (soleus), and 0.91% and 5.0% (gastrocnemius) by using MRI and magnetic resonance spectroscopy fat quantification methods (40–43). Estimation of FF using T2-based techniques has been conducted in previous studies (17, 18). However, these studies have studied higher FFs (>5%) not comparable with the results of this study.

The importance of high SNR in the estimation of T2-relaxation times has been studied previously (44). Here, we simulated the effect of noise on fat quantification using T2-based methods and found that the 2- and 3-parameter fit overestimated the FF at lower SNRs. This could be explained by the fact that the presence of noise might be interpreted by the 3-parameter fit as fat signal, as T2, F > T2, W. Using correct T2-values was also shown to be of importance when using the 2- or 3-parameter fit, particularly for water, to measure low FFs correctly. Owing to a high lowest SNR (~200) of the acquired MESE data, SNR may not be the main issue for the 2- and 3-parameter fit at low and intermediate FFs. Instead, incorrect T2 values are a more probable cause of bias. Like the NLLS-based methods, the Bayesian fit performed better as SNR increased. Simulations also showed that using a correct S0 value is important for obtaining an accurate estimation of the FF, particularly for intermediate FFs. The simulated Bayesian fit resulted in a larger standard deviation compared with the NLLS-based methods, suggesting that the Bayesian fitting method, using the slicesample algorithm as described in this work, might be less robust. Although this contrasts previous results (25), the used Bayesian probability approaches are not identical and might therefore not be com-

![Figure 6](image_url)

**Figure 6.** The acquired and fitted signal decay of 3 voxels with different FFs are shown for the low-resolution (A–C) and high-resolution (D–F) MESE images.

![Figure 7](image_url)

**Figure 7.** The difference between the true and estimated FFs when using incorrect T2, W (A) and incorrect T2, F (B) in the 2-parameter fit and when using incorrect T2, F (C) in the 3-parameter fit. The true T2, W and T2, F are 40 milliseconds and 160 milliseconds, respectively.
parable. In addition, the number of estimated parameters is larger in the Bayesian fit compared with the 2-parameter fit which affects the robustness.

Although the 2-parameter fit resulted in accurate FFs compared with low-resolution CSEI, it depends on whether it is possible to obtain both $T_2,F$ and $T_2,W$ without any contamination, that is, fluid accumulation due to edema or extramyocellular and intramyocellular fat. It has also been reported that $T_2,W$ varies between muscle groups (40). Using a $T_2,W$ calculated from an ROI placed in 1 muscle group could therefore result in incorrect FFs in other muscles. In this study, one volunteer had too little available subcutaneous fat, making it impossible to draw a ROI to obtain an individual $T_2,F$. Alternatively, one could use $T_2$-relaxation times obtained from literature. However, simulations in this study suggest that it is important to use correct $T_2$-relaxation times to avoid biases. A fat quantification method without the need of ROIs, like the Bayesian method, might therefore be preferred.

Owing to varying $T_2$-values between the muscles, one might expect that keeping $T_2,W$ fixed would result in less accurate FF calculation compared with estimating $T_2,W$ together with $W$ and $F$. Both the in vivo results and the simulations suggested that this was not the case, as the 3-parameter fit overestimated the FF, and was more sensitive to incorrect $T_2$-relaxation times. A recent paper instead described the fat signal decay using a biexponential model, that is, a triexponential model for the total (water and fat) signal (18). It is possible that a biexponential description of the fat signal could improve the results of the NLLS methods in this study. For the Bayesian method, a triexponential signal model has also been studied for intravoxel

**Figure 8.** The simulated effect of using an incorrect $S_0$ value in the Bayesian fitting method showing the difference between the estimated and true FF (left), and the standard deviation of 1000 estimations (right).

**Figure 9.** The difference between estimated FF and true FF equal to 2%, 5%, 10%, 30%, and 95% at different signal-to-noise ratios (SNRs) (20, 50, 150, 300, and 600) using the 2-parameter fit (A), the 3-parameter fit (B), and the Bayesian fit (C). The corresponding standard deviations of 1000 estimations using the 2-parameter fit (D), the 3-parameter fit (E), and the Bayesian fit (F) are also shown. The mean SNR of the collected MESE images ($512 \times 512$) in muscle and fat is shown in the first plot (A). At low FF, all 3 methods overestimate the FF in the presence of a high noise level. The difference between estimated and true FF, of the 3-parameter fit outside the shown interval in (B) are: $-40.9\%$ (SNR = 20) and $-27.8\%$ (SNR = 50). The corresponding standard deviation not shown in (E) is $\pm 41.4\%$ and $\pm 39.0\%$, respectively.
incoherent motion applications (45), which could be adapted for fat quantification purposes. Another source of error might be the use of only 1 T2-relaxation time to describe the fat signal decay, although it consists of several composites, each with an individual T2-relaxation time (29).

Although the Bayesian fit slightly underestimated the FF compared with low-resolution CSEI, there are numerous advantages with the method. For example, all volunteers could be evaluated independent of the amount of subcutaneous fat, as the method is not dependent on the ROI definition. Additional advantages are the possibility of obtaining \( T_{2W} \) maps and that no user input is needed. Further investigations to improve the performance of the Bayesian fit using slicesample are needed, including the choice of the number of echoes to include in the calculations, estimation of \( S_0 \), smoothing level of the parameter probability density functions, and number of generated samples and burn-in factor in the slice sample algorithm.

Several drawbacks with the T2-based methods that were used in this study were found. First, the effect of B1-inhomogeneities was not accounted for, assuming perfect T2 decay of the signal over the course of the MESE acquisition. Methods suggested in previous studies include dismissing voxels where large B1-inhomogeneities are present by obtaining a B1-map by an additional data acquisition (18) and using a method based on extended phase graphs (46). Second, owing to the long T1-relaxation time of muscle tissue, an impractically long TR (\( >4 \) seconds) is needed to avoid T1 bias. Alternatively, as was done here, a T1-correction can be performed in the postprocessing steps. In this study, T1-relaxation times obtained from the literature were used to correct for T1 bias. This can introduce errors if the true T1-relaxation times are different from the ones obtained from the literature. Using individual T1-relaxation times could possibly improve the correction, but it will require additional data acquisition. Third, owing to the long acquisition times of MESE images, parallel imaging had to be used to reduce the scan time. This caused varying noise levels and therefore varying SNR over the images.

Several other drawbacks and limitations of the study were identified during this work. The study was conducted in healthy subjects only, and no pathological fat accumulation was seen. Thus, the range of FFs investigated was likely to be lower than that of a patient group. Phantom studies are not included, as it was not possible to construct a phantom which worked for all methods simultaneously. A completely fair comparison of the precision of the various methods was not possible, as they were acquired using different acquisition times and the number of estimated parameters differed between the methods. However, the effect of increasing the acquisition time of high-resolution MGRE images to that of MESE was investigated in 1 subject and it was found to still result in a high noise level and inferior image quality.

In conclusion, all the T2-based methods could produce high-resolution FF images of the calves of healthy volunteers, where the FF was 1%–6%. The 2-parameter fit showed the best quantitative agreement to low-resolution CSEI. The method can thus be an alternative to CSEI when the latter method fails to produce high-resolution FF images owing to low SNR or fat–water swaps. However, the NLLS-based methods are sensitive to incorrect T2-values, particularly \( T_{2W} \) for low FFs. Although the Bayesian fit avoids this particular limitation, further development is needed before it can be used for accurate fat quantification.

ACKNOWLEDGMENTS

This research was supported by Magnus Bergvalls stiftelse and Direktör Albert Pählsson's Stiftelse.

Disclosure: No disclosures to report.

REFERENCES


Vascular Deformation Mapping (VDM) of Thoracic Aortic Enlargement in Aneurysmal Disease and Dissection

Nicholas S. Burris¹, Benjamin A. Hoff¹,², Ella A. Kazerooni¹, and Brian D. Ross¹,²,³

¹Department of Radiology; ²Center for Molecular Imaging; and ³Department of Biological Chemistry, University of Michigan, Ann Arbor, MI

Corresponding Author:
Nicholas Burris, MD
Department of Radiology, University of Michigan, 1500 E. Medical Center Drive, TC B1-132, SPC-5030, Ann Arbor, MI 48109-5030;
E-mail: nburris@med.umich.edu

Key Words: aortic aneurysm, spatial Jacobian, aortic dissection, 3D printing, aneurysmal enlargement, vascular deformation mapping

Abbreviations: Vascular deformation mapping (VDM), abdominal aortic aneurysm (AAA), thoracic aortic aneurysm (TAA), electrocardiogram (ECG), computed tomography angiography (CTA), Hounsfield unit (HU), computed tomography (CT), matrix metalloproteinases (MMPs)

Thoracic aortic aneurysm is a common and lethal disease that requires regular imaging surveillance to determine timing of surgical repair and prevent major complications such as rupture. Current cross-sectional imaging surveillance techniques, largely based on computed tomography angiography, are focused on measurement of maximal aortic diameter, although this approach is limited to fixed anatomic positions and is prone to significant measurement error. Here we present preliminary results showing the feasibility of a novel technique for assessing change in aortic dimensions, termed vascular deformation mapping (VDM). This technique allows quantification of 3-dimensional changes in the aortic wall geometry through nonrigid coregistration of computed tomography angiography images and spatial Jacobian analysis of aortic deformation. Through several illustrative cases we demonstrate that this method can be used to measure changes in the aortic wall geometry among patients with stable and enlarging thoracic aortic aneurysm and dissection. Furthermore, VDM results yield observations about the presence, distribution, and rate of aortic wall deformation that are not apparent by routine clinical evaluation. Finally, we show the feasibility of superposing patient-specific VDM results on a 3-dimensional aortic model using color 3D printing and discuss future directions and potential applications for the VDM technique.

INTRODUCTION

The thoracic aorta is the largest blood vessel in the human body and is subject to most extreme hemodynamic forces. A healthy aorta is extremely durable and it is able to absorb forces generated by the heart owing to its thick walls and its elastic nature. Because of multiple factors (eg, hypertension, atherosclerosis, genetic aortic syndromes, infection), the structural integrity and elasticity of the aortic wall can deteriorate, leading to progressive dilation of the aortic lumen and formation of aortic aneurysm (1). Aortic dissection is a related form of aortic disease characterized by tearing of the inner layers of the aortic wall (ie, intima and media), leading to the creation of a false lumen—or channel—within the aortic wall itself, which is structurally compromised and is subjected to high pressures. This results in aneurysm formation in ~60% of patients with chronic aortic dissection of the descending thoracic aorta (Stanford type B) (2). The incidence of aortic aneurysm is increasing in the US population, and mildly dilated aortas are being incidentally detected at higher rates owing to increased use of thoracic cross-sectional imaging for nonaortic indications (eg, lung cancer screening) (3). Recent data suggest that the prevalence of thoracic aortic dilation (~4 cm) is ~3% of individuals older than 55 years of age, which, on the basis of current US population estimates, means that ~2.7 million people in the USA would be recommended to undergo regular imaging of the thoracic aorta on the basis of the current American Heart Association guidelines for imaging surveillance (4–8).

Imaging surveillance has a central role in the management of asymptomatic patients with aortic disease. The vast majority of patients with an aortic aneurysm, ~95%, are asymptomatic until the aneurysm ruptures, and only 40% of patients in whom the aneurysm ruptures reach the hospital alive (6, 9). Although the topic of aortic enlargement in abdominal aortic aneurysm (AAA) before and after endovascular repair has been the focus of significant research effort, the natural history and mechanisms of thoracic aortic aneurysm (TAA) progression remain poorly understood, and only a handful of studies have attempted to measure growth rates of the thoracic aorta (10–16). A major limitation in improving our understanding of TAAs is that the current clinical imaging surveillance techniques rely primarily on measurement of maximal aortic diameter. This parameter has been most widely studied and is shown to correlate with future
VDM of Thoracic Aortic Enlargement

risk of aneurysm rupture (17). Although the simplicity of diameter measurements is appealing, this approach is subject to a high degree of measurement error, in the range of 2–5 mm despite optimal measurement technique (18, 19). Error of this magnitude makes confident determination of aortic enlargement challenging considering that typical aortic growth rates are slow (eg, 1 mm/y in the ascending aorta and 3 mm/y in the descending aorta), and this issue is further compounded when shorter follow-up intervals are analyzed (3 or 6 months) and when the aortic geometry is ovoid (17).

Although several sections of the aorta are vertically oriented and can be viewed in cross-section on axial images, most of the aorta cannot be viewed in cross-section on standard image planes, requiring image-processing software to effectively straighten the aorta and allow true orthogonal diameter measurements to be made. The 2010 ACCF/AHA/AATS/ACR/ASA/SCA/SCAI/SIR/STS/SVM Guidelines for the Diagnosis and Management of Patients With Thoracic Aortic Disease was the first set of guidelines to raise this issue and to recommend standard measurement locations in addition to measurement of maximal aortic diameter (4). Even with orthogonal measurements, the aorta is often not perfectly round in the cross-section, but it is rather ovoid or irregular, particularly in the setting of disease, further compounding the issue of exactly which diameter measurements to record and use for follow-up. Measuring the aortic diameter at predefined anatomic locations fails to capture interval growth at nonmaximal locations, and this measurement does not detect the components of aortic enlargement in circumferential or longitudinal directions. Height-/weight-adjusted aortic area has been proposed as a better predictor of future rupture than maximal diameter, and several studies have investigated the use of volumetric measurements of TAA and AAA to improve the sensitivity for detecting aortic growth (10, 20–24). However, similar to diameter measurements, measurement of aortic area and volumetric should be performed at predetermined anatomic boundaries to ensure that measurements are comparable between studies, and small focal changes in aortic dimension may be camouflaged by a volumetric measurement approach. Although area or volume measures may be more sensitive to detect overall growth of an aortic segment, information about localized change at a specific point along the aortic wall is not captured. Considering that surgical management recommendations are based on thresholds of size and growth rate, a diameter-based measurement technique may lead to treatment recommendations that are either overly aggressive or conservative on the basis of the measurement error alone. In addition, such size criteria used for surgical decision-making are based on historical measurement data and its inaccuracies, further emphasizing the significant ongoing need for accurate and reproducible aortic measurements.

A potential solution to these challenges exists in the field of diffeomorphic image registrations, particularly the application of spatial Jacobian matrices to quantify the deformation of the aortic wall by extracting nonrigid image transformations to match high-resolution thoracic electrocardiogram (ECG)-gated computed tomography angiography (CTA) images acquired at different time points, a technique that we term vascular deformation mapping (VDM). Spatial Jacobian matrices in this context describe the relative local distortion at each point in the image resulting from the automated image-based registration. Although nonrigid image warping coregistration techniques have been broadly used in diseases of the lungs and brain, to the best of our knowledge, no prior studies have quantified spatial Jacobian maps to assess interval aortic enlargement (25–27). Image intensity-based registration techniques have been reported to have submillimeter precision in many applications, which also translate to accuracy in the calculation of the resulting spatial Jacobian with well-optimized workflows (28). Quantitative assessment of registration accuracy is not straightforward, with many potential sources of error and a large number of degrees of freedom; however, the use of cost function penalties such as bending energy helps to constrain and smooth spatial Jacobian results while also maximizing anatomical feature alignment. Jacobian maps may be directly calculated from the optimized nonrigid transform, and it could offer information about the local deformation of the aortic wall, including in the circumferential and longitudinal directions, information that is not currently assessed by other techniques.

A significant need exists for a more sensitive and accurate method of measuring change in thoracic aortic dimensions, considering that accurate detection of small-magnitude changes has important implications for improving understanding of aortic aneurysm progression and better informing treatment decisions. The aim of this study was to demonstrate the feasibility of using the proposed VDM technique to measure interval change in aortic wall dimensions on routine clinical CTA studies of patients with TAA and dissection. In addition, we aim to compare the results of the VDM analysis with routine clinical CTA assessments in seeking to better understand the potential benefits and limitations of this novel technique.

METHODS AND MATERIALS

Study Population

All procedures were approved by the local institutional review board with a waiver of informed consent obtained for this retrospective study and were Health Insurance Portability and Accountability Act–compliant. Patients were identified through a review of local picture archiving and communication system archives to identify adult (>18 years) patients with dilation of the thoracic aorta undergoing imaging surveillance, with at least 2 prior ECG-gated CTA examinations available for review. Patients were excluded if thoracic aortic enhancement was suboptimal (<250 Hounsfield unit [HU]) or there was significant motion/respiratory artifact affecting the clinical evaluation of thoracic aortic segments. After reviewing 15 patients, several were excluded owing to obvious pulsation artifact affecting the diseased aortic segment on CTA images (n = 5) or low-resolution baseline CTA studies (section thickness > 1.5 mm) that were acquired at outside hospitals and uploaded to our picture archiving and communication system (n = 4). One patient with type B aortic dissection was excluded from analysis owing to difficulties with accurate segmentation of the false lumen due to poor enhancement related to slow flow and partial thrombosis. The aortic pathologies of those patients selected for analysis included ascending thoracic aortic...
dilation \( (n = 1) \), descending TAA \( (n = 1) \), and thoracic aortic dissection \( (n = 3) \).

**Computed Tomography Angiography**

CTA examinations were performed on 64-detector computed tomography (CT) scanners using helical acquisition mode (LightSpeed VCT or Discovery CT750HD, GE Healthcare, Waukesha, WI). Images were acquired through the entire thoracic aorta (lung apices to 2 cm below the celiac artery) during intravenous injection of 95-mL iopamidol 370 mg I/mL (Isovue 370, Bracco Diagnostics, Inc., Princeton, NJ) at 4 mL/s, followed by a 100-mL saline chaser at 4 mL/s. Retrospective ECG-gating was used, with axial reconstructions at 0.625-mm section thickness at 75% of the R–R cycle with ECG-modulated tube current technique (20% of maximum milliamperes) and 40% adaptive statistical iterative reconstruction for dose efficiency. Other scan parameters included the following: detector coverage, 40 mm; Display Field of View, 25 cm; gantry rotation time, 0.4 seconds; maximum tube current, 400–700 milliamperere as determined by patient size; and tube voltage, 100–120 kVp.

**Image Segmentation**

Segmentation of the aortic blood volume was accomplished with a user-defined threshold in contiguous regions followed by manual adjustments, all performed using custom in-house algorithms developed in Matlab (The MathWorks, Inc, Natick, MA). In brief, a threshold was chosen on a case-by-case basis to separate contrast-enhanced blood from the surrounding tissues and organs. Manual separation was required at the aortic valve and arch vessel levels. The surface structure was determined on the basis of the segmentation mask and then subject to curvature flow smoothing.

**Image Registration**

Image registration was performed between sequential CTA studies using a custom Matlab interface to the Elastix open source software (Utrecht, Netherlands). Images were processed temporarily for registration with the following after manually cropping around the region of the aorta:

1. A 3D Wiener filter \( (3 \times 3 \times 3) \) was applied to limit the effects of noise.
2. Image values \(<0\) HU were set to 0 HU to avoid including the lungs.
3. The aortic blood segmentation mask was dilated by 6 mm to include the aortic wall.

Automated image registration included an affine optimization followed by a multiresolution nonrigid b-spline warping optimization using mutual information (subsampled within the dilated segmentation mask) with bending energy penalty (set to 50). Three resolutions of b-spline grid spacing were used in the descending order, as follows: 12, 6, and 3 mm. Total time for image registration was \( \sim 10 \) min on a standard high-end personal computer.

**Vascular Deformation Mapping**

Using the deformation fields generated from the final optimized nonrigid transformation, the spatial Jacobian tensor can be defined as all first-order derivatives at each voxel location:

\[
J = \begin{bmatrix}
\frac{\partial \Delta x}{\partial x} & \frac{\partial \Delta x}{\partial y} & \frac{\partial \Delta x}{\partial z} \\
\frac{\partial \Delta y}{\partial x} & \frac{\partial \Delta y}{\partial y} & \frac{\partial \Delta y}{\partial z} \\
\frac{\partial \Delta z}{\partial x} & \frac{\partial \Delta z}{\partial y} & \frac{\partial \Delta z}{\partial z}
\end{bmatrix}
\]

The determinant of the 3D spatial Jacobian, \( |J| \), or simply referred to as the Jacobian map, was calculated from the final optimized image transform and normalized by the time difference between imaging sessions \( (|J|/\text{years}) \) to indicate a deformation rate. Values of \(|J|\), henceforth referred to as VDM values, were linearly interpolated to the vertex points of the aortic segmentation surface for display. Image expansion is visualized by greater values \((\text{red}, |J| > 1)\), compression by \(|J| < 1 \text{ (blue)}\), and no general deformation by \(|J| = 1 \text{ (green)}\). Areas of expansion or compression were considered artificial if one of the following criteria was present:

1. visible motion artifact was present on source CTA images;
2. visible error was noted in image alignment after the image warping coregistration step; or
3. regions of expansion/compression were adjacent to the cut-planes of the 3D aortic segmentation \((\text{eg. at the level of the aortic valve or proximal arch vessels})\), as these areas are susceptible to minor differences in geometry resulting from manual segmentation.

A simplified workflow of the VDM technique is displayed in Figure 1.

**RESULTS AND DISCUSSION**

**Evaluation of Aortic Aneurysm**

The VDM analysis clearly depicted interval enlargement of the descending aortic dimensions in our first representative case of a 76-year-old female patient with a prior history of surgical repair of an ascending aortic aneurysm. The aortic arch and descending aorta were not included in the initial surgical repair given the mild degree of preoperative dilation; however, the distal arch and descending aorta were noted to progressively enlarge over 3 subsequent CTA examinations spanning a period of 3.8 years (Figure 2). It is interesting to note that although the VDM shows enlargement of the proximal descending aorta at each interval, the extent and rate of enlargement progress from the first interval to the last, consistent with the gradually accelerating and outwardly expanding nature of aortic enlargement described with aortic aneurysm \( (6) \). In addition, although the clinical radiologist’s assessment using maximal aortic diameters identified enlargement at each interval, the growth rate appeared to be decelerating by diameter measurements, and the growth was reported to be limited to the distal arch, whereas the VDM clearly highlighted more extensive enlargement along the length of the aorta, involving the proximal and mid-descending aorta at the second and third intervals. In an attempt to quantify and validate the VDM results, aortic area measurements were performed at a single level in the distal descending aorta with close attention paid to placing the measurement plane at precisely the same level and orientation in each study (Figure 2). The luminal area measurements revealed a small...
increase in area at the first interval (3.7 mm$^2$), a larger increase in luminal area at the second interval (22.1 mm$^2$), and the greatest increase in luminal area at the third interval (100.2 mm$^2$) consistent with the accelerating growth visualized on the VDM map. For reference, an overall luminal increase of 100 mm$^2$ is equal to a $\sim$1.1 mm increase in the diameter assuming that the lumen is circular.

In contrast to the above case, the VDM analysis was performed on a 66-year-old female patient undergoing imaging surveillance of a mildly dilated ascending aorta (maximally 4.1 cm at baseline), which revealed little deformation (Figure 3). This case was selected for analysis to serve as a negative control, as no enlargement was detected by clinical diameter assessment, and enlargement of the ascending aorta is both significantly slower and less common than enlargement of the descending aorta, particularly when the degree of dilation is mild [15]. VDM did not reveal any areas of rapid growth in the ascending or descending aorta, and the majority of the 3D surface area of the thoracic aorta showed $|J|$ values close to 1 (green), compatible with stable aortic dimensions. However, several small regions of moderate deformation were detected, one at the level of the sinotubular junction, another in the proximal arch in the region of the origin of the innominate artery, and the last at the mid-descending level. No motion artifact or image registration error was visually apparent, and, although the significance of these findings remains unclear, it is possible that they represent areas of slow growth that are beneath the threshold of detection by diameter measurements.

**Evaluation of Aortic Dissection**

Aortic dissection and aortic aneurysm are unique in their pathophysiology; however, the ultimate consequence of both pathologies is the same—dilation of the aortic wall due to weakened structural integrity. In both aneurysm and dissection, clinical surveillance guidelines and surgical decision-making are based on observation of the absolute aortic dimensions and the rate of aortic enlargement. As such, the VDM technique can be used to monitor progression (ie, enlargement) of patients with aortic dissection. In the first representative case, we present the results of a 56-year-old patient with a prior history of surgically repaired dissection of the ascending aorta, with a residual dissection flap involving the native aortic arch and descending aorta (Figure 4). VDM showed values close to 1 (green) throughout the majority of the aorta, compatible with stable dimensions of the true and false lumen during the 2-year time interval, in agreement with the clinical diameter assessment. There were several small areas of apparent mild enlargement in the ascending aorta and the distal descending aorta, which are thought to be due to imprecisions in coregistration caused by slight differences in cardiac and respiratory phases between studies, resulting in minor differences in aortic angulation (ie, “bending”), although no definite misregistration was visually apparent.

In contrast, Figure 5 illustrates a 52-year-old man with a history of type B aortic dissection who had 2 ECG-gated CTA studies available for analysis, the first $\sim$1 year after the onset of his dissection and a follow-up study performed 6 months after the first. On the basis of the clinical report, there was suspicion for $\sim$1–2 mm of interval enlargement of the distal aortic arch,
but the conclusion of the clinical assessment was that there had been no definite enlargement, as the observed change in diameter was within the range of measurement error. VDM of this patient showed nearly diffuse enlargement of the false lumen throughout the distal aortic arch and descending aorta, with a corresponding decrease in size of the true lumen, changes that are frequently observed in chronic aortic dissection (29). It is important to note that although the absolute change in maximal aortic dimension was thought to be small (1–2 mm), the rate of growth is noted to be significant owing to the short interval (6 months) between the 2 studies. This ability to detect growth over short intervals is particularly useful in the setting of patients with recent aortic dissection, as there is a proven clinical benefit to endovascular (TEVAR) repair in the subacute period (2 weeks to 3 months after dissection) (30). Of note, there was a visually apparent motion artifact in the ascending aorta on CTA images, leading to difficulty with image coregistration, which is manifested on the VDM as a wavy aortic wall contour and areas of high and low Jacobian determinant (red and blue, respectively) on adjacent areas of the aortic wall.

Finally, Figure 6 illustrates a 76-year-old man with a history of ascending aorta replacement, who developed a type B aortic dissection on the baseline study and had 4 surveillance examinations...
CTAs performed over a 3-year period available for analysis. Comparing the clinical reports of the first and most recent CTA studies, the patient experienced up to 6 mm of enlargement overall at various points along the descending aorta during the 4-year follow-up period; however, using clinical diameter measurements interval enlargement was only confidently detected at the second interval. During the first interval, the VDM analysis revealed several areas of enlargement along the descending aorta, with the most intense areas in the proximal and distal descending aorta characterized by enlargement of the false lumen and compression of the true lumen. Despite the distal descending aorta being susceptible to image coregistration error related to respiratory variation, we did not visually detect any issues with image coregistration. On re-examination of the CTA studies, there was suggestion of 1–2 mm of aortic enlargement at these levels by diameter assessment. Furthermore, the entry tears that allow blood to flow from the true lumen into the false lumen were located at the proximal descending and distal descending levels, the locations of the most rapid growth, supporting the VDM results that enlargement had occurred in these regions. At the second interval, the VDM map again showed regions of false lumen enlargement; however, the rate of aortic enlargement decreased. Finally, during the third interval, the descending aorta that appeared showed only a small area of continued false lumen enlargement at the mid-descending level, and otherwise no interval change. This observed gradual deceleration of aortic growth over time has previously been described in patients with chronic aortic dissection (29). Although the mechanisms underlying the evolution of chronic type B dissection remain poorly understood, the wall of an acutely dissected aorta contains minimal fibrosis, and structural integrity of the wall is low. With increasing chronicity, the aortic wall undergoes a process of adaptive remodeling, mainly through increased collagen deposition, leading to increased wall rigidity and a decreased rate of enlargement (29). Unfortunately, such remodeling processes are insufficient for preventing aneurysm formation in some patients who remain at a risk for rupture. The ability to accurately measure the rate of aortic enlargement at each follow-up interval may better inform clinical management through improved depiction of the overall growth trend (ie, accelerating vs decelerating), and may contribute to a better understanding of the natural history of aneurysm formation among patients with chronic aortic dissection.
Advantages and Clinical Opportunities

We believe that the early results presented here clearly demonstrate several unique advantages of the VDM technique over maximal diameter measurements for assessment of aortic enlargement in the setting of aneurysm and dissection. First, and most importantly, we believe that the application of the VDM technique can result in reduced measurement error, as it uses the full 3D image data along the entire length of the aorta, rather than diameters placed at fixed locations along the aortic length. Furthermore, our approach relies on modern semiautomated nonrigid image registration techniques that can align CTA images with a precision in the range of 0.5–1 mm (28). The reported 2–5 mm range of error associated with aortic diameter measurement on CT arises from several potential sources, including variability in placement of the diameter calipers along the near and far walls, as well as variability in the rotation of the diameter plane when the aorta is not uniformly circular as shown in Figure 7 (31). However, the VDM technique relies on modern nonrigid image coregistration techniques that are able to match each location along the 3D aortic wall surface with submillimeter accuracy.

Second, the VDM technique offers the distinct advantage of being able to map a continuous range of growth rates in a 3D fashion, both along the entire length of the aorta and around its circumference, whereas aortic diameter measurements are limited to a single radial position at a fixed anatomic location. In addition, wall deformation can be assigned vectors, allowing for measurement of directional deformation in addition to overall magnitude, a characteristic of aortic aneurysm growth that has not been previously quantified in situ. Separating the full Jacobian tensor into components of normal, circumferential tangent, and longitudinal tangent magnitudes may provide an even more nuanced understanding of changes in aortic wall geometry.

Volumetric and cross-sectional area measurements, although reported to be more sensitive for aortic enlargement, rely on discreet predefined anatomic boundaries of the aorta (ie, start and stop points along the length of the aorta), and are therefore limited in determining the spatial location and gradation of aortic enlargement. The 3-dimensional nature of VDM lends itself to robust and easily interpretable data visualization modalities that are customizable and approachable for surgeons and other nonimager aortic specialists. Furthermore, physical models displaying VDM may now be easily and economically 3D-printed and provided to surgeons preoperatively to aid in surgical planning. Although surgical decision-making remains a complex and patient-specific task, mapping the distribution of growth along the entire thoracic aorta, particularly areas of growth at nonmaximal locations, allows for the possibility of tailoring the surgical repair technique to include areas of slow growth that may not otherwise be detected by diameter measurements, and it could potentially necessitate future reoperation.

Lastly, because of the spatially continuous and quantitative nature of the VDM and its reduced measurement error, increased sensitivity for detection of eccentric and small-magnitude aortic enlargement is possible. The rate of aortic enlargement, rather than absolute increase in maximal diameter, can be easily calculated and visualized. The rate of aortic enlargement is more closely related to the underlying structural and cellular mechanisms that drive aneurysm progression, and is likely a better indicator of risk among patients with aortic aneurysm. Unfortunately, the rate of enlargement often cannot be accurately calculated from aortic diameter measurements owing to significant measurement error, particularly when time intervals—the denominator in a rate measurement—are short. The decreased measurement error attainable with the VDM technique may allow confident determination of slow aortic enlargement over short time intervals (eg, 3–6 months), rather than the several-year time frame often required for diameter measurements.

Earlier detection and more accurate quantification of aortic growth may allow for more targeted and aggressive treatment of aortic disease, may provide for better-informed decisions to undergo major aortic surgical or endovascular procedures, and may be useful in the research setting where aortic enlargement is an outcome of interest and follow-up periods are limited by cost or other logistical considerations. Furthermore, the frequency of surveillance imaging can be better tailored to an individual if the stability of their aorta can be more accurately assessed; patients with slow-growing or stable aneurysms may have imaging spaced to 2- to 3-year intervals, allowing for more efficient health-care utilization, whereas patients with rapid enlargement can undergo imaging more frequently in the hopes...
of minimizing the incidence of potentially predictable and preventable complications.

**Technical Challenges**

The results presented here constitute the first steps in the application of a VDM technique for the evaluation of aortic disease, and while preliminary results show great promise, several challenges remain. First, because this analysis has a high degree of sensitivity to aortic wall deformation, errors can be introduced by factors resulting in differing spatial alignment of the 2 compared aortic geometries. The 2 areas most susceptible to such error are at the aortic root (sinuses of Valsalva) and at the distal descending aorta at the level of the diaphragm, with the 2 main contributing factors being cardiac and respiratory motion. The effects of these factors on variation in aortic geometry have been previously described (32). The aortic root has the highest degree of pulsatory motion of any thoracic aortic segment owing to its close proximity to the heart, with the degree of pulsation amplified during expiration. In addition, the entire thoracic aorta has a relatively uniform lateral and posterior displacement with expiration. However, although uniform displacement could be easily corrected for during image coregistration, the distal descending aorta remains relatively fixed in position by the diaphragm, and this nonuniform motion introduces potential misalignment during image coregistration.

These observations stress the importance of acquiring the CTA images during the same phase of respiration (preferably inspiration) and with ECG-gating (preferably in late diastole) to minimize errors attributable to the small phasic variations in aortic geometry. In addition to respiratory and pulsation artifacts, “stair-step” artifact is occasionally encountered in ECG-gated CTAs, particularly when studies are performed on scanners with detector rows numbering 64 or less. The stair-step artifact is problematic for the VDM technique, as it creates an abrupt shelf-like defect in the 3D aortic segmentation that limits image coregistration. Fortunately, modern CT scanners that have been optimized for cardiovascular imaging can greatly minimize the frequency and severity of stair-step artifacts owing to the increased number of detector rows and decreased gantry rotation time.

Second, the VDM method has significant technical demands, both in terms of user expertise and computing requirements. Many software platforms are currently available to perform accurate semiautomated aortic segmentation, which, if combined with our technique, may significantly decrease the time required for analysis. In addition, although the required computing power makes the VDM technique less practical for smaller centers that lack dedicated high-performance computing capabilities, the recent rise of cloud-based medical image analysis software may allow for analysis to be centralized, lessening technical demands at the level of the end user. Development of this VDM technique is still in the early stages of development; however, routine clinical applicability of this approach will require streamlining of the analysis with improved semiautomated image segmentation techniques and code optimization to minimize technical demands and user input.

**Future Technical Directions**

Future technical efforts will be focused on the following 5 key topics:

1. Further validating the accuracy and consistency of the VDM analysis is an important next step, as there is no clear noninvasive—or even invasive—tool to measure changes in aortic dimension in situ. As such, VDM results will need to be compared with other more sensitive measurements of aortic enlargement such as CTA luminal area and volumetric assessments. In addition, phantoms with precisely controlled dimensions could be used to further validate VDM results in a controlled experimental setting.

2. Further developments in computer algorithms will be needed to further automate and improve image segmentation, coregistration, and analysis steps. In particular, additional efforts need to be focused on accurate and reproducible false lumen segmentation in patients with aortic dissection, as the false lumen is prone to heterogeneous enhancement related to either partial thrombosis and/or slow blood flow.

3. Development of a user-friendly console for easy visualization and interaction with the VDM results will be an important step in moving this technology into the realm of clinical practice.

4. Evaluating the feasibility of quantifying changes in the aortic dimensions between different points in the cardiac cycle, rather than between different studies acquired at the same point in the cardiac cycle. The aortic dimensions have been shown to change significantly with pulsation, and measurement of these changes using VDM could provide important insights into the elasticity/rigidity of the aortic wall, a characteristic that has been associated with a large variety of cardiovascular diseases (33).

5. Investigating the feasibility of the VDM technique to quantify enlargement of other pathologies that manifest as progressive vascular enlargement such as AAA, cerebral aneurysm, pulmonary artery enlargement related to pulmonary hypertension, and endoleak after endovascular aortic repair.

**Future Clinical Directions**

In addition to ongoing technical developments, this technology has several unique potential clinical and research applications that we plan to investigate. First, given the rapidly increasing availability and decreasing costs of color 3D printing, we plan to study the clinical utility of superimposing VDM results on full-scale, color 3D-printed aortic models. To demonstrate feasibility of this approach, we submitted the VDM results from Figure 2 (Interval 2) for color 3D printing with the results shown in Figure 8. To the best of our knowledge, there are no prior published reports of superimposing imaging-based measurements of physiologic, such as the VDM results of aneurysm enlargement, on 3D-printed anatomic models. A large proportion of the literature in the field of 3D printing has focused on the production of high-fidelity anatomic models from medical imaging data, with some proposed applications including rapid prototyping for medical device development, creation of individualized medical implants/prostheses, and operative planning and pa-
basis now identified to involve 3 distinct pathomechanisms that include perturbation of the transforming growth factor β signaling pathway, disruption of the vascular smooth muscle cell contractile apparatus, and impairment of extracellular matrix synthesis (36, 37, 45). Advances in our understanding of the underlying pathogenetic alterations involved in the pathogenesis of thoracic aortic disease are providing significant new opportunities for therapeutic interventions using novel pharmaceutical approaches. The development of a validated imaging biomarker would allow for longitudinal quantification of the effects of drug interventions on modulation of disease progression. This capability would provide unique opportunities to use this imaging biomarker to facilitate development of therapeutic strategies in both preclinical aneurysm models and for use in clinical translational trials undertaking novel therapeutic strategies.

Finally, the potential impact of VDM results on clinical patient care will need to be thoroughly studied. Primary topics of investigation will include investigation of the associations between VDM measures and patient cardiovascular risk factors, the potential of VDM assessment to reclassify patient risk assessments, and the ability of VDM to predict patient outcomes. Although the VDM technique may indeed be more sensitive in detecting change in the aortic dimensions, an analysis of how data from this novel method could change patient management will be needed to show value over the more easily obtained aortic diameter measurements. Considering that the VDM analysis can be performed respectively on routine clinical CTA scans, the VDM results can be compared with clinical reports and a wide variety of patient demographic parameters and outcomes such as surgical repair strategy, surgical complication rate, reoperation rates, and occurrence of aorta-specific adverse events during imaging surveillance. In addition, as the VDM technique allows for assessment of aortic enlargement at specific spatial locations along the aortic wall, growth can be colocalized with pathological features of the aortic wall that are believed to promote aneurysm development such as atherosclerotic plaque (both calcified and lipid-rich), mural thrombus, intimal hyperplasia, or wall thickness. Identifying direct correlations between localized aortic wall pathology and regional wall expansion by VDM analysis could greatly advance our understanding of the underlying pathophysiology that leads to aortic aneurysm, and offer new strategies to predict aortic events, risk-stratify patients, and monitor the effectiveness of pharmacological therapy.

CONCLUSION
We have demonstrated the feasibility of a spatial Jacobian-based technique to measure changes in the size of the aortic lumen between baseline and follow-up ECG-gated thoracic CTA examinations in patients with mild aortic dilatation, aortic aneurysm, and aortic dissection, and that this technique is capable of quantifying and visually displaying the degree of aortic enlargement in a 3-dimensional fashion. Furthermore, we have shown that there are clear discrepancies between the VDM results and clinical diameter assessments, with the VDM technique appearing more sensitive for detection of changes in aortic dimensions owing to reduced measurement error, al-

Figure 8. Color 3D-printed model of the thoracic aorta produced from the VDM results of a patient with progressively enlarging descending aortic aneurysm presented in Figure 2, interval 2. The superimposition of pathophysiologic VDM data on a 3D-anatomic model represents a novel application of medical 3D printing and be valuable in operative planning.

tient-specific device sizing in the setting of complex anatomy (34, 35). However, despite such promising applications, the role of 3D printing in clinical practice remains a topic of debate, largely owing to questions of cost versus benefit if similar observations and measurements can be made by 3D medical imaging software. As demonstrated here, the addition of VDM results superimposed on the 3D model surface significantly increases the amount of information that the model contains and may further justify the effort involved in creating patient-specific 3D models. Although a computerized VDM provides a detailed overview of the aneurysm enlargement, a 3D model that can be handled and closely studied by surgeons may allow for subtle anatomic and functional observations that are not as easily appreciated on a digital image, and could facilitate patient education during clinic visits.

Second, the VDM technique may be extremely valuable as a research tool for studies of aneurysm pathophysiology owing to its high degree of sensitivity to changes in aortic dimensions. There has been increased interest in elucidating the cellular pathways involved in remodeling of the vascular wall leading to the formation of aortic aneurysm (36, 37). Various cellular signaling pathways along with host-immune interactions have been implicated in the pathogenesis of AAAs (38-41). These factors are in part related to the complex underlying cellular processes responsible for the loss of extracellular matrix and wall remodeling. For example, chronic inflammation of the aorta wall has been implicated in the activation of matrix metalloproteinases (MMPs) such as MMP-2 and MMP-9, which have been reported to play a role in aortic wall weakening and subsequent AAA formation (42, 43). Furthermore, transforming growth factor β signaling alterations (44) have been widely associated with vascular smooth muscle disease with the genetic
though formal quantification of the degree of error reduction and the potential clinical impacts of a more sensitive analysis of aortic dimension changes requires further investigation. The VDM technique for measurement of change in aortic wall dimensions holds the promise of considerably improving the accuracy of aortic imaging surveillance, informing clinical decision-making, furthering aortic research questions, and shedding light on the natural history of aortic disease.

ACKNOWLEDGMENTS

The authors would like to thank Drs. Himanshu Patel and Bo Yang for their guidance pertaining to the potential applications of vascular deformation mapping in the field of aortic surgery. This study received support from BDR: National Institutes of Health Research Fellow Grant (RF1502).

REFERENCES


SIGNA™ Returns.

For more than 30 years, SIGNA has stood for quality, trust, and innovation. This year, SIGNA Returns with the introduction of a family of new products that bring you the trusted performance you’ve relied on for over three decades, plus powerful forward-thinking technologies to help enhance your clinical confidence.

Visit gehealthcare.com for more information.
The Future of Preclinical Imaging

The next generation of superconducting 7T, 4.7T and 3T cryogen free MRI.

- Up to 12 hour power protection
- No quenching due to power loss
- No cooling water required
- 1500 mT/m gradients

Multi-modality Imaging

- **PET/MR**
  - Simultaneous & sequential acquisition of PET & MR
  - MRS-PET

- **SPECT/MR**
  - Sequential & simultaneous acquisition of SPECT & MR
  - MRS-SPECT

- **Upgrades**
  - Bring your MRI back to life
  - MRS Upgrade

- **Clinical Kit**
  - Clinical to preclinical conversion kit
  - C2P

- **Confocal**
  - In vivo optical imaging
  - MRS CellLIVE

For more information contact us at:
information@mrsolutions.com
www.mrsolutions.com
Designed to meet the need for simultaneous PET/MRI, the PET Insert from Bruker is compatible with MRI systems up to 15.2 Tesla and allows simultaneous measurements for the perfect correlation of PET and MRI data in space and time. The insert can also be used for standalone PET imaging, providing a flexible option for a range of applications.

For more information please visit: www.bruker.com/pet-insert