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
Tomography® is published quarterly. Tomography publishes basic scientific (technical and pre-clinical) and clinical articles advancing the development and applications of imaging technologies. Tomography encompasses studies using 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 image-derived metrics within and across modalities toward the development of novel imaging probes for image-based feedback and intervention. 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. 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. 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.

**Tomography is Indexed and Abstracted in:**
- MEDLINE®
- PubMed Central®
- Europe PMC®
- Scopus®
- EMBASE®
- Index Copernicus®

and these Clarivate Analytics Services:
- Science Citation Index Expanded (also known as SciSearch®)
- Journal Citation Reports/Science Edition®
- Current Contents®/Life Sciences

**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.


**Reprints:** To order reprints for educational, commercial, or promotional use, contact Cadmus Reprints, P.O. Box 822942, Philadelphia, PA 19182-2942; E-mail: cjsreprints@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.
Image Reports
Relationship Between Urolithiasis and Fatty Liver Disease: Findings in Computed Tomography
Federico Guillermo Lubinus Badillo, Oscar Leonel Ortiz Cala, Silvia Nathalia Vera Campos, and Erick Daniel Villarreal Ibañez

Advances in Brief
Sex Differences in Kidney Function and Metabolism Assessed Using Hyperpolarized [1-13C]Pyruvate Interleaved Spectroscopy and Nonspecific Imaging
Yibo Wen, Haiyun Qi, Christian Østergaard Mariager, Per Mose Nielsen, Lotte Bonde Bertelsen, Hans Stædkilde-Jørgensen, and Christoffer Laustsen

Research Articles
Assessment of the Prognostic Value of Radiomic Features in 18F-FMISO PET Imaging of Hypoxia in Postsurgery Brain Cancer Patients: Secondary Analysis of Imaging Data from a Single-Center Study and the Multicenter ACRIN 6684 Trial
Mark Muzi, Eric Wolsztynski, James R. Fink, Janet N. O’Sullivan, Finbarr O’Sullivan, Kenneth A. Krohn, and David A. Mankoff

MRI-Based Deep Learning Segmentation and Radiomics of Sarcoma in Mice
M. D. Holbrook, S. J. Blocker, Y. M. Mowery, A. Badea, Y. Qi, E. S. Xu, D. G. Kirsch, G. A. Johnson, and C. T. Badea

Microstructure Modeling of High b-Value Diffusion-Weighted Images in Glioblastoma
Yuan Li, Michelle Kim, Theodore S. Lawrence, Hemant Parmar, and Yue Cao

Efficient CT Image Reconstruction in a GPU Parallel Environment

COVER ILLUSTRATION
Convolutional neural network segmentation and radiomic analysis of multi-parametric MRI of a soft tissue sarcoma was used to provide insights into therapeutic interventions in the paper by Matthew Holbrook, et al.
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
Relationship Between Urolithiasis and Fatty Liver Disease: Findings in Computed Tomography

Federico Guillermo Lubinus Badillo¹, Oscar Leonel Ortiz Cala², Silvia Nathalia Vera Campos³, and Erick Daniel Villarreal Ibañez⁴

¹Medical Specialist in Radiology and Diagnostic Imaging, Foscal Clinic, Floridablanca, Colombia; ²Specialist in Radiology and Diagnostic Imaging, Autonomous University of Bucaramanga, Bucaramanga, Colombia; ³Medical and Surgeon, Industrial University of Santander, Bucaramanga, Colombia; and ⁴Medicine Student, Autonomous University of Bucaramanga, Bucaramanga, Colombia

Corresponding Author: Federico Guillermo Lubinus Badillo, MD
Universidad Autonoma de Bucaramanga, Bucaramanga, Santander, Colombia, Calle 155A #23-60; E-mail: flubinus@hotmail.com

Key Words: nonalcoholic fatty liver disease, urolithiasis, metabolic syndrome

INTRODUCTION
Nonalcoholic fatty liver disease (NAFLD) is a given term to define a spectrum of liver entities ranging from simple steatosis to non-alcoholic steatohepatitis, with the eventual development of non-cholestatic nonalcoholic cirrhosis and the risk of hepatocellular carcinoma (1). According to the variability of the different types of population studies carried out (2), NAFLD affects a range from 1% to 51% of the population. Moreover, nephrolithiasis is considered a problem with a significant economic and health burden. Its prevalence has increased in recent years, in parallel to the increase in the prevalence of metabolic syndrome. Recent evidence has suggested a close link between NAFLD and the increased risk of urolithiasis; however, the biological mechanism remains unclear (3, 4). Among the risk factors associated with NAFLD, the most important is the syndrome of insulin resistance and the metabolic syndrome. Diabetes mellitus and dyslipidemias are considered primary causes (5, 6, 7). Liver biopsy is the gold standard for the diagnosis of NAFLD, but this test is an invasive procedure and not free of complications. Computed tomography (CT) is a noninvasive method with rapid acquisition of images that allows quantitative evaluation of the diagnosis of NAFLD; therefore, it is useful in clinical practice and offers monitoring and follow-up of patients (8, 9). To the best of our knowledge, there are no studies on joint diagnostic or therapeutic intervention for the treatment of fatty liver and urolithiasis, perhaps because it is not known if there is an association between these two diseases. The purpose of this study was to identify the relationship between renal lithiasis and fatty liver disease by examining for common factors that could be used to reduce their incidence and complications.
Protocol for Computed Tomography of the Abdomen

The protocol for CT of abdomen included the following steps:

- Tomography with and without intravenous contrast with 5-mm cuts from the lung base to the pelvic floor.
- Data were obtained using noncontrast CT examinations. Demographic information of age and gender was collected. Additional or incidental findings such as diverticular disease in the colon, cholelithiasis, or atherosclerosis of the abdominal aorta were also recorded.
- Liver density was determined using a circular region of interest (ROI) parameter of 1 cm in diameter in the right lobe, making an isolated measurement of each segment and after that an average of the value obtained in each hepatic segment was done for a final single value.
- Spleen density was obtained by computing and averaging 3 measurements using a 1-cm-diameter circular ROI at the upper, middle, and lower segments.
- The diagnosis of fatty liver was made by using quantitative techniques, obtaining the difference of liver–spleen densities. If the result is <5 UH, fatty liver is absent. Results between 5 UH and −10 UH are classified as mild-to-moderate steatosis and results lower than −10 UH are considered moderate-to-severe steatosis.

Urolithiasis was classified as macrolithiasis and microlithiasis having as a discriminatory value stones less than 2 mm in greatest diameter. Its location in the urinary system from the renal calyces, the ureteral segments and the bladder was also described. To determine the density of the renal stones, a circular ROI was used that would occupy the largest diameter of the stone surface, and the values were averaged once 3 measurements were obtained. In a window for bone, the intrinsic characteristics of the calculation are described taking into account its homogeneity or heterogeneity. The size of the stone was determined based on 3 dimensions and the volume of these measurements.

Statistical Analysis

For the analysis, the information was entered into a database, typed in duplicate, and validated with the STATA 10 statistical package. An exploratory descriptive analysis of the imaging findings was conducted through measures of central tendency and dispersion for the quantitative variables and through proportions with 95% CIs for qualitative variables. A bivariate analysis was performed, calculating relative prevalence, then a multivariate analysis was performed using a regression method to adjust their combined effects, and finally the adjusted relative prevalence of each detected attitude was calculated. The input and output of variables in the regression method were performed based on the criteria established by Sander Greenland (23).

RESULTS

We analyzed 1010 CT scans of 510 women and 500 men who met the inclusion criteria, out of an overall of 4938 CT scans performed by the radiology department during a 12-month period. Fatty liver disease was found in 458 scans, representing 45.3% of the cases, and urolithiasis was observed in 676 images, corresponding to 66.9% of the cases (Figure 1). The coexistence of both findings is more frequent in men, while in the isolated form, it predominates in women. The percentage of CT without urolithiasis and without Fatty liver Disease was 27% in women and 14% in men.

We found that 337 patients, corresponding to 33% of the cases, presented simultaneously urolithiasis and hepatic steatosis. Half of the patients with urolithiasis also presented hepatic steatosis and approximately two-thirds of the patients without urolithiasis had a nonsteatotic liver, with a PR of 1.37, P < .0001, and when adjusted for age and gender, the PR is 1.29, P < .002—a statistically significant finding (Table 1).

We classified the severity of steatosis in ranges of absence of steatosis, mild/moderate, and moderate/severe, according to the difference in the hepato/splenic densities. We found a progressive simultaneous increase in the percentage of urolithiasis, with 61.4% for healthy patients and 71% for those with steatosis in the mild–to-moderate range and 75.7% for patients classified as having severe steatosis (Table 2).

The density of the kidney stones was divided into higher or lower than 500 UH, to differentiate these by their composition given that those <500 UH contain primarily uric acid. The density of the stone is not associated with hepatic steatosis (PR = 0.88; 95%CI (0.75–1.01) P = .12). The intrinsic homogeneity and heterogeneity characteristics of the calculation related to hepatic steatosis had an associated PR of 0.85, 95%CI (0.71–1.02), P = .68.

DISCUSSION

The association of urolithiasis and hepatic steatosis has been poorly studied. In our study, we found this association in 33% of the total number of patients, a result similar to the result of the study of Einollahi B et al. (10) (with a prevalence of 30%), conducted on 11 245 patients in whom ultrasonography was used to determine the presence of hepatic steatosis and urolithiasis. It is interesting to mention that if we evaluate only patients with...
hepatic steatosis, the frequency of urolithiasis is seen in 73.5% of the subjects. If we observe only patients with urolithiasis, we find that 50% of the cases present hepatic steatosis, which indicates that there is a shared causal relationship between the 2 pathologies.

In observational studies that evaluated the nutritional and metabolic characteristics associated with these pathologies, the causal mechanism is not yet known; however, the production of free radicals of oxygen and the lipid peroxidation or hepatocyte mitochondrial damage are important factors for the genesis of the stones. An increase in the systemic production of free radicals of oxygen at the renal level and mitochondrial lipid peroxidation causes an increase in the concentration of calcium oxalate, favoring its precipitation and formation of stones (11, 12), findings that have been supported in the Carrasco-Valiente (13) study.

It has also been reported that fatty acids can modify the urinary excretion of calcium and oxalate within the phosphorylation of arachidonic acid in association with nephrolithiasis. This is because the phosphorylation of arachidonic acid determines a cascade of metabolic effects on calcium homeostasis, which generates hypercalciumia owing to the action of secondary messengers or the PGE2/vitamin D receptor (14).

The study of Kim S et al. (3) observed that the association between NAFLD and nephrolithiasis was more prominent in participants younger than 50 years of age (adjusted HR, 1.19; 95% CI, 1.14 ± 1.24) than in those older than 50 years (adjusted HR, 1.06; 95% CI, 0.95 ± 1.19) (P < .001). We found similar findings when the distribution by age range was made. We observed that the simultaneous association between hepatic steatosis and urolithiasis decreases significantly and inversely with age.

Different studies have evaluated the simultaneous presentation of dyslipidemia and urolithiasis taking into account that a metabolic alteration is one of the most important causes of fatty liver. Although the causal factor has not been determined yet, a common path must exist given its high association. Other studies also support this idea like Torricelli’s study (15). This study reported on 2442 patients and was based on the multifactorial origin of urolithiasis, in which it was found that patients with high levels of total cholesterol and triglycerides presented urolithiasis more frequently than normal controls (P = .006 and P < .0001, respectively), finding with statistical significance that these patients had higher serum concentrations of calcium oxalate and uric acid. Ding Q et al. (16) performed a case–control study and found that patients with hypertriglyceridemia and low high-density lipoprotein cholesterol levels had an increased risk of nephrolithiasis (OR, 1.31; 95% CI: 1.01 ± 1.71; OR 7.57). The study of Kang (17) calculated the risk of recurrence of urolithiasis to be up to 50% at 5 years in patients with dyslipidemia and metabolic syndrome. They followed up 1561 patients and identified a relationship between dyslipidemia and urolithiasis, with patients with high levels of triglycerides being at a higher risk for recurrence of urolithiasis (95% CI, 1.2–2.8; P = .005)—a statistically significant result and an independent risk factor. Masterson et al. (18) evaluated the diagnosis of dyslipidemia and urolithiasis in 52184 patients and confirmed an association of these pathologies, thus suggesting that low levels of high-density lipoprotein cholesterol predispose to the presence of nephrolithiasis (HR, 1.4; 95% CI, 1.1–1.7; P = .003)—a statistically significant result compared with other reports where other components of the lipid panel favor the formation of urinary stones. Thus, treating dyslipidemia could mitigate the risk of urolithiasis.

Observational studies such as those of Paten (19) (98 patients), Cho (20) (712 patients), and Naya et al. (21) assessed patients diagnosed with metabolic syndrome and insulin resistance, and evaluated its effect on the development of urolithiasis, observing that these pathologies cause changes that favor ammoniagenesis and systemic acidosis, increase calcium bone resorption, and decrease hypercalciumia. The reabsorption of citrates causes hypocitraturia and favors the precipitation of

### Table 1. Association between Fatty Liver and Urolithiasis

<table>
<thead>
<tr>
<th>Fatty Liver Disease</th>
<th>Healthy Liver</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urolithiasis</td>
<td>337 (49.8%)</td>
</tr>
<tr>
<td>Without Urolithias</td>
<td>121 (36.2%)</td>
</tr>
<tr>
<td>Overall</td>
<td>458 (45.4%)</td>
</tr>
<tr>
<td></td>
<td>676 (100%)</td>
</tr>
</tbody>
</table>

PR crude = 1.37 (95% CI, 1.17-1.62; P < .001), PR adjusted* = 1.29 (95% CI, 1.1-1.53; P = .002) (*By gender and age). Abbreviation: PR, prevalence ratio.

### Table 2. Severity of Fatty Liver Disease and Urolithiasis

<table>
<thead>
<tr>
<th>Fatty Liver Degree</th>
<th>Without Urolithiasis</th>
<th>Urolithiasis</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absence</td>
<td>213 (38.5%)</td>
<td>339 (61.4%)</td>
<td>552 (100%)</td>
</tr>
<tr>
<td>Mild-to-Moderate</td>
<td>65 (28.6%)</td>
<td>162 (71%)</td>
<td>227 (100%)</td>
</tr>
<tr>
<td>Moderate-to-Severe</td>
<td>56 (24.2%)</td>
<td>175 (75.7%)</td>
<td>231 (100%)</td>
</tr>
</tbody>
</table>

P < .0001.
calcium in urine. In acid urine, the low pH increases the precipitation of uric acid, which is favored by high serum levels secondarily to the increase in intake of protein that chemically predispose to stone formation (4, 22).

The limitations of our study lie in its retrospective type and in the lack of the clinical, nutritional, and anthropometric information, as well as the associated pathologies, treatments, and lifestyle, of the subjects assessed that may favor the genesis of these pathologies. A study taking these factors in consideration is recommended to determine the cause of the relationship between the 2 conditions more accurately. A follow-up study is necessary to evaluate the natural evolution of the disease and to search why the associations between hepatic steatosis and urolithiasis decrease with age, as they are inferred to be due to changes in lifestyle that favor the decrease or regression of liver and kidney disease.

ACKNOWLEDGMENT
Disclosures: The authors have nothing to disclose.

Conflicts of interest: None reported.

REFERENCES
Sex Differences in Kidney Function and Metabolism Assessed Using Hyperpolarized [1-13C]Pyruvate Interleaved Spectroscopy and Nonspecific Imaging

Yibo Wen1,2, Haiyun Qi1, Christian Østergaard Mariager1, Per Mose Nielsen1, Lotte Bonde Bertelsen1, Hans Stødkilde-Jørgensen1, and Christoffer Laustsen1

1MR Research Centre, Department of Clinical Medicine, Aarhus University, Aarhus, Denmark; and 2The First Affiliated Hospital of Zhengzhou University, Zhengzhou University, Henan, China

Corresponding Author:
Christoffer Laustsen, PhD
The MR Research Centre, Department of Clinical Medicine, Aarhus University, Aarhus University Hospital, Palle Juul Jønsens Boulevard 99, 8200 Aarhus N, Denmark; E-mail: cl@clin.au.dk

Key Words: MRI, sex, kidney, renal metabolism, hyperpolarization, BSSFP, spectroscopy

Abbreviations: Blood oxygenation level-dependent (BOLD), dynamic contrast-enhanced (DCE), magnetic resonance imaging (MRI), fatty acid (FA), magnetic resonance spectroscopy (MRS), magnetic resonance spectroscopic imaging (MRSI), dynamic nuclear polarization (DNP), dissolution dynamic nuclear polarization (dDNP), free induction decay (FID), balanced steady-state free precession (BSSFP), glomerular filtration rate (GFR)

ABSTRACT

Metabolic sex differences have recently been shown to be particularly important in tailoring treatment strategies. Sex has a major effect on fat turnover rates and plasma lipid delivery in the body. Differences in kidney structure and transporters between male and female animals have been found. Here we investigated sex-specific renal pyruvate metabolic flux and whole-kidney functional status in age-matched healthy Wistar rats. Blood oxygenation level–dependent and dynamic contrast-enhanced (DCE) magnetic resonance imaging (MRI) were used to assess functional status. Hyperpolarized [1-13C]pyruvate was used to assess the metabolic differences between male and female rats. Female rats had a 41% ± 6% and 41% ± 5% lower absolute body and kidney weight, respectively, than age-matched male rats. No difference was seen between age-matched male and female rats in the kidney-to-body weight ratio. A 56% ± 11% lower lactate production per mL/100 mL/min was found in female rats than in age-matched male rats measured by hyperpolarized magnetic resonance and DCE MRI. Female rats had a 33% ± 11% higher glomerular filtration rate than age-matched male rats measured by DCE MRI. A similar renal oxygen tension (T2*) was found between age-matched male and female rats as shown by blood oxygenation level–dependent MRI. The results were largely independent of the pyruvate volume and the difference in body weight. This study shows an existing metabolic difference between kidneys in age-matched male and female rats, which indicates that sex differences need to be considered when performing animal experiments.

INTRODUCTION

Metabolic sex differences are gaining increasing attention as an important factor in tailoring treatment strategies. Men are at an increased risk of developing nondiabetic renal disease compared to age-matched women (1). Type-2 diabetic women have a greater risk of renal dysfunction than type-2 diabetic men, potentially resulting from an increased susceptibility to poor metabolic control (2). Furthermore, women typically have considerably more body fat, higher whole-body fat turnover rates, and higher plasma triglycerides than men (3). Recent reports have shown sex-specific myocardial metabolic patterns, suggesting differences in treatment effects among male and female individuals. Divergent responses to metformin in myocardial fatty acid (FA) metabolism between women and men appear to explain the lack of treatment effects when treatment groups were analyzed as a whole (4, 5). It is likely, that similar renal metabolic and functional differences exist between men and women. The existence of sex differences could play an important role in designing and interpreting preclinical and clinical experiments. An increasing number of experimental studies indicate that there are in fact sex differences in both renal rodent models and in human subjects (6–9). Moreover, the differences of renal transporters and renal ammonia metabolism have been reported in rodents with fundamental structural difference between males
Sex differences in Kidney Function and Metabolism

and females (10, 11). Thus, caution in the translation of the results between sexes is imperative. Previous reports have shown morphological differences between male and female rats of >50% for epithelial volume and absolute volume of proximal tubule (12). However, sex difference in renal pyruvate metabolism has, to the best of our knowledge, not been previously reported. The development of hyperpolarized 13C magnetic resonance spectroscopy (MRS) and spectroscopic imaging (MRSI) based on dissolution dynamic nuclear polarization (dDNP) has enabled unprecedented noninvasive visualization of normal and abnormal metabolism in living systems, without the use of ionizing radiation (13–15). This novel method relies on the ability to trace the metabolic conversions of intravenously injected molecules, such as hyperpolarized 13C-pyruvate, as it undergoes metabolic conversion into its metabolic derivatives: 13C-lactate, 13C-alanine, and 13CO2/13C-bicarbonate (16–19). Here we investigated the potential for monitoring the functional and metabolic signatures of age-matched (8 weeks) female and male rats, which received the same volume of hyperpolarized [1-13C]pyruvate. To investigate the dose-related impact of pyruvate, an additional group of dose-matched male rats (5 weeks) was included. Dynamic contrast-enhanced (DCE) MRI in combination with hyperpolarized [1-13C]pyruvate MRI was performed on both age-matched male and female rats. The combination of perfusion and metabolic conversion, that is, fractional perfusion, allowed us to account for the substrate delivery (20), which could be different between males and females. This metric can be simplified by using the hyperpolarized tracer itself, and thus, we investigated if a simple combination of 2 standard sequences could be used as surrogates for the more elaborate imaging sequences typically used. A new interleaved imaging and spectroscopy approach (Figure 1) was used. The sequence consists of a standard slice selective fully balanced steady-state free precession (BSSFP) sequence interleaved with a slice-selective free induction decay (FID) spectroscopic sequence. This would allow high temporal–spatial localization of hemodynamics and the spectroscopic information with high spectral resolution and as such represent a good compromise for many examinations. The rationale for the proposed method is that as the first pass-perfusion is largely dominated by the substrate signal (21, 22), the metabolic conversion is dominated by kidney metabolism (23), thus offering an alternative approach to image the functional and metabolic features of the kidneys.

**METHODS**

**Animal Model**

Male (8 week old; 343.9 ± 6 grams; n = 8) and female (8 week old; 203.7 ± 7 grams; n = 8) Wistar rats were included in this study. The age-matched male and female rats received the same volume of hyperpolarized [1-13C]pyruvate, that is, 1 mL. In addition, 8 male Wistar rats (5 week old; 183.2 ± 3 grams) exposed to similar experimental conditions was compared to the age-matched male and female rats to test if the dose of the pyruvate injection influences the results. The dose-matched male rats received a hyperpolarized [1-13C]pyruvate volume similar to the female rats, that is, 4.9 mL/kg. The details on the rats and the dose of received pyruvate are listed in Table 1. The rats were kept in standard rodent cages with 12/12 h light/dark cycle, temperature of 21 ± 2°C, and humidity of 55% ± 5%. All rats had free access to water and standard rat chow throughout the study. The experiment started after 1 week of acclimatization. All procedures regarding this experiment complied with the guidelines for use and care of laboratory animals. The study was approved by the Danish Inspectorate of Animal Experiments. The rats were anesthetized with 2.5% sevoflurane in 2 L/min of air. A tail vein catheter (0.4 mm) was inserted for injection of both hyperpolarized [1-13C]pyruvate, which was polarized in a 5-T SPINLab (GE Healthcare, Brondby, Denmark) as previously described (24). A MRI contrast agent (Dotarem; Guerbet, Roissy, France) was used for DCE measurements. Respiration, body temperature, and peripheral oxygen saturation were monitored during the experiment as previously described (24).

Magnetic resonance (MR) scanning procedures were performed using a 9.4-T preclinical MR system (Agilent, Yarnton, UK) with VnmrJ 4.0A (Agilent, Santa Clara, CA), using a dual-tuned 13C/1H volume rat coil (Doty Scientific, Columbia, SC) as previously described (25, 26). The MR data were reconstructed and analyzed using MATLAB (MathWorks, Natick, MA) and Osirix (27). The kidney parenchyma regions of interest (ROIs) were manually drawn.

**Hyperpolarized MR**

Anatomical 1H MRI scans were conducted to determine the location of the kidneys. Afterwards, hyperpolarized [1-13C]pyruvate MR imaging and spectroscopy was performed to quantify renal hemodynamics and metabolic conversion of [1-13C]pyruvate, respectively. A balanced steady-state free precession (BSSFP) imaging sequence was interleaved with a standard slice-selective free induction decay (FID) acquisition to acquire spectra, both sequences using the same slice selection parameters (see online supplemental Figure 1 and Figure 1). Parameters for the BSSFP sequence were: 1 echo, repetition time (TR) = 2.2 milliseconds, echo time (TE) = 1.1 milliseconds, matrix = 32 × 32, field of view (FOV) = 60 × 60 mm², flip angle = 30°, and slice thickness = 15 mm, with a total scan time of 70 milliseconds. Parameters for the FID acquisition were: TR = 206 milliseconds, flip angle = 15°, spectral width = 10 000 Hz, and complex points = 2048. A delay was inserted before each acquisition, to fix the TR to 2 seconds. This was chosen to allow renewal of the hyperpolarized [1-13C] pyruvate magnetization pool in the imaging slice. Nineteen images from BSSFP imaging, used for perfusion analysis, and 19 spectra from FID acquisition, used to acquire the metabolites data, were acquired. Age-matched rats received a bolus of 1 mL (125 mM) of hyperpolarized [1-13C]pyruvate. The other 8 male rats received hyperpolarized [1-13C]pyruvate according their body weight, matching the dose per weight that female rats received, that is, 4.9 mL/kg (612.5 μmol/kg). The interleaved acquisition was initiated at the beginning of the injection. The spectra were analyzed in MATLAB for metabolic information, using an in-house general linear model fitting of the known peaks similar to that mentioned in the literature (28). The areas under curve (AUCs) of spectral peaks of [1-13C]pyruvate and its metabolic products were calculated for the total signal. The DCE images were analyzed using the Osirix software. The cortex of kidneys was drawn manually to calculate kidney hemodynamics in the Osirix UMPPerfusion plugin (29). A T1 relaxation time of...
20 seconds was used for pyruvate T1 correction of the perfusion data (21, 30). Fractional perfusion, used to correct the metabolic information for the pyruvate delivery, was calculated by multiplying the metabolic ratios with the DCE-derived perfusion (20).

**BOLD MRI**

Blood oxygen level–dependent (BOLD) MRI was performed (Figure 2) to measure the oxygen tension in the kidneys of the male and female rats. A standard multiecho gradient echo sequence with 10 echoes was used to acquire 5 slices; TR = 118 milliseconds, TE = 2.0 milliseconds, ΔTE = 2.2 milliseconds, FOV = 70 × 70 mm², matrix = 128 × 128, flip angle = 30°, and slice thickness = 1 mm. T2* maps were fitted as previously described and evaluated using the modified 6-layer concentric objects method for ROI analysis (31, 32).

**DCE MRI**

DCE MRI (Figure 1) was performed using a gradient echo sequence with the following parameters: TR = 13.7 milliseconds, TE = 1.9 milliseconds, FOV = 60 × 60 mm², matrix = 128 × 128, flip angle = 15°, and slice thickness = 20 mm. Acquisition was initiated 10 seconds before administration of MRI contrast to acquire a baseline, with a time resolution of 1.75 seconds in 180 time steps. A single 1-mL bolus isotonic saline containing 0.05 mL of Gadoteric acid (Dotarem; 279.3 mg/mL; Guerbet, Villepinte, France) was administered over 10 seconds. Kidney hemodynamics (Perfusion, MTT, and blood volume) were calculated in the Osirix UMMPerfusion plugin as previously described (29). Glomerular filtration rate (GFR) was estimated using the Baumann–Rudin model as previously described (33).

**Tissue Harvesting**

After the MR examination, ~5 mL of arterial blood was collected from the aortic bifurcation from each anesthetized rat before sacrifice. Further, 1 mL of arterial blood was used for ABL blood–gas analysis. The remaining 4 mL of blood was centrifuged for 10 minutes at 1500 g to separate the plasma which was stored in a −80°C freezer for further analysis. The renal cortex and inner medulla were also stored at −80°C.

**Quantitative Polymerase Chain Reaction (QPCR)**

The fresh frozen cortex tissue was used to measure the messenger RNA (mRNA) of monocarboxylate transporter (MCT) 1–4, pyruvate dehydrogenase (PDH), alanine aminotransferase (ALT), and lactate dehydrogenase (LDH) (see online supplemental Table 1). In brief, total RNA was isolated using NucleoSpin RNA II mini kit according to the manufacturer’s instructions (AH Diagnostics, Aarhus, Denmark). RNA was quantified by spectrophotometry and stored at −80°C. cDNA synthesis was performed with RevertAid First Strand cDNA synthesis kit (MBI Fermentas, Burlington, Canada). QPCR was performed using Maxima SYBR Green qPCR Master Mix according to the manufacturer’s instructions (AH Diagnostics). In brief, 100 ng of cDNA was used as template for PCR amplification. Specificity of products was confirmed by melting curve analysis and by gel electrophoresis.

**Statistics**

All data are presented as mean ± SEM. All statistical analysis was performed in GraphPad Prism 6 (GraphPad Software, Inc. La Jolla, CA). Normal distribution was tested by a Shapiro–Wilk normality test and QQ-plots. Data were analyzed using either an unpaired 2-tailed Student t test or a 2-way-ANOVA with repeated measures; multiple comparisons was corrected using a Tukey correction when appropriate. A value of P < .05 was considered statistically significant.
RESULTS
Sex Differences in Anatomy and Blood
A statistically significant difference was observed between age-matched male and female rats, in body weight (BW) and kidney weight (KW), albeit the kidney weight was similar if correction for body weight was performed (Table 2). All data in online supplemental Table 2 was found to be similar between age-matched male and female rats, except for the arterial potassium level (lower in the female rats) and the blood hematocrit and hemoglobin (increased in female rats), which showed an insignificant difference (see online supplemental Table 2).

Sex Differences in Renal Function
A numerical increased renal plasma flow and mean transit time was found in the male rats, but the difference was not statistically significant (Figure 3, A and C). However, a statistically significant lower renal blood volume ($P=.034$) was found in female rats than in age-matched male rats (Figure 3B). Interestingly, an increased GFR was found in female rats than in male rats (Figure

<table>
<thead>
<tr>
<th>Table 2. Body Weight (BW), Kidney Weight (KW, Mean of Both Kidney’s), Blood Glucose Level (Fed State) at the Day of the Scan and Kidney Weight Per Kilogram Rat (KW/BW)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Body Weight (g)</strong></td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>8-week-old male</td>
</tr>
<tr>
<td>8-week-old female</td>
</tr>
</tbody>
</table>

* $P < .05$ versus male group.
A similar renal oxygen tension was found between age-matched male and female rats, demonstrating a significant increased oxygen tension (decreased $^1H T_2^*$ values) from pelvis to cortex (Figure 3E) ($P < .001$).

### Differences in Renal Metabolism Related to Sex

Age-matched male and female rats showed different sex-dependent metabolic responses as measured with $[1-^{13}C]$pyruvate (Figure 4A; $P = .002$). Male rats had a higher lactate/pyruvate ratio than female rats ($P < .0001$), while alanine/pyruvate and bicarbonate/pyruvate did not show a significant difference. When normalized to kidney weight, the 3 metabolites showed no significant difference between the 2 groups (Figure 5A). A significant difference in the perfusion of metabolites was found between the 2 sexes (Figure 4B, $P = .0002$), mainly in the higher fractional perfusion of lactate in the male group than in the female group ($P < .0001$). The fractional perfusion of alanine and bicarbonate showed a similar tendency but without significant difference (Figure 5B). The metabolic balance of the 3 metabolites showed significant differences in the different sex (Figure 4C, $P = .034$). Female rats had a higher alanine/lactate ratio than male rats ($P < .0001$), while the aerobic/anaerobic metabolism (bicarbonate/lactate ratio) did not show significant difference between age-matched male and female groups. When comparing the fraction of different metabolites in the total $^{13}C$ metabolite signal ($\Sigma^{13}C$-metabolites), a higher lactate/$\Sigma^{13}C$-metabolites ratio ($P = .0018$) and a lower alanine/$\Sigma^{13}C$-metabolites ratio ($P = .0003$) was found in male rats than in female rats, and there were no significant differences in the bicarbonate/$\Sigma^{13}C$-metabolites ratio (Figure 4D) between age-matched male and female rats. No correlation was seen between $^1H$ DCE and $^{13}C$-perfusion assessment using a nonselective BSSFP ($P = .20, R^2 = 0.13$). Expression levels of pyruvate metabolic related enzymes of LDH, PDH, ALT, and MCT1-4 did not reveal any sex-specific differences, although a numerical tendency was observed (see online supplemental Figure 1). To test if the constant dose used for the pyruvate injection had an effect on the metabolic profile, another 8 male rats (5 week old) received the matched dose of hyperpolarized $[1-^{13}C]$pyruvate injection that was given to the female rats (Table 1). Interestingly, the fractional lactate production was increased in male dose-matched rats (Figure 5C) than in female rats. Inspecting the pyruvate normalized metabolic distribution (Figure 5D) in the different groups, a similar tendency was observed with an increased lactate production, while the alanine production was less pronounced. To verify the body weight or age dependence on the metabolic difference between male and female rats, a correlation analysis was performed on the lactate production in the male rats with different body weights (spanning the female rat

![Figure 3. Renal hemodynamic features in age-matched male and female rats. A similar plasma flow was observed between male and female rats ($P = .095$) (A). A lower renal blood volume of 35.6% was observed in female rats than in male rats ($P = .034$) (B). A similar, although insignificant, mean transit time was observed in male and female rats ($P = .16$) (C). A 32.8% increase in glomerular filtration rate (GFR) was observed in female rats than in male rats ($P = .01$) (D). A similar decreased $^1H T_2^*$ relaxation (surrogate for oxygen availability) from cortex to pelvis was observed in both male and female rats (E). Data are illustrated as means ± SEM. *$P < .05$ versus male group.](image-url)
body weights). No significant correlation was observed under these conditions (Figure 5E).

**DISCUSSION**

The main finding of this study was the significant differences in renal function and associated metabolic profile between the female and male groups, showing a generally elevated metabolic activity in male kidneys. The mRNA expression levels of pyruvate metabolic-related enzymes of LDH, PDH, ALT, and MCT1-4 did not reveal significant sex-specific differences. One potential reason for this metabolic difference could be the number of nephrons, which previously has been correlated with kidney size (34, 35). This is supported by the higher body weight and kidney weight levels found in the male group than in the female group, while the kidney-weight to body-weight ratio shows no significant difference between the groups. This suggests that kidney growth is well matched with body weight. However, correction for the metabolic conversion by kidney weight reduced the metabolic differences, which indicates that the number of nephrons
could be the origin of this sex difference. In humans, females also generally have smaller kidneys with fewer nephrons than males (36, 37). To examine if the metabolic differences could be affected by different body weights and thus the injected dose between age-matched male and female rats, 1 additional male group with a lower body weight was examined. The injected dose of hyperpolarized $[^{1-^{13}}C]$pyruvate was matched to that of the female rats. The lower-body-weight male rats showed an increase in lactate production and body weight in male rats ($P = .0023$), in the fractional perfusion of lactate, while no differences were seen in alanine and bicarbonate perfusion ($P = .0077$, interaction term sex $\times$ metabolites $P = .17$, metabolites $P < .0001$) (C). A distinct metabolic distribution was seen between the female and male rats in the pyruvate normalized case ($sex \times .99$, interaction term sex $\times$ metabolites $P < .0001$) (D). Linear regression showed no significant correlation of lactate production and body weight in male rats ($P = .32$), while the female rats has a significantly lower lactate production compared to males (E).
production compared to female rats. The differences were less pronounced between the age-matched males and dose-matched males and thus support sex differences as an important parameter in designing and interpreting preclinical results. We speculate that this could be similar in humans, as female patients with diabetes are often at a higher risk of renal dysfunction than males (38), as well as the fact that renal protection has been associated with an increased metabolic activity (39). The acquisition strategy was found to have a less pronounced effect, although it was found to impact the metabolic pools and as such represents an important point when comparing preclinical data even among same-sex individuals. Recent publications using hyperpolarized $^{13}$C magnetic resonance imaging suggest that this technique may enable human examinations, using more selective imaging methods. The impact the metabolic pools and as such represents an important point when comparing preclinical data even among same-sex individuals. The acquisition strategy was found to have a less pronounced effect, although it was found to impact the metabolic pools and as such represents an important point when comparing preclinical data even among same-sex individuals. Recent publications using hyperpolarized $^{13}$C magnetic resonance imaging suggest that this technique may enable human examinations, using more selective imaging methods.

Equal Contribution: Y.W. and H.Q. contributed equally to this work.

The project was supported by Aarhus University Foundation and Karen Elise Jensens Foundation. We thank Duy Dang and Mette Dalgaard for expert technical assistance.

REFERENCES

Sex differences in Kidney Function and Metabolism


Assessment of the Prognostic Value of Radiomic Features in $^{18}$F-FMISO PET Imaging of Hypoxia in Postsurgery Brain Cancer Patients: Secondary Analysis of Imaging Data from a Single-Center Study and the Multicenter ACRIN 6684 Trial

Mark Muzi¹, Eric Wolsztynski²-⁴, James R. Fink¹, Janet N. O’Sullivan², Finbarr O’Sullivan², Kenneth A. Krohn¹, and David A. Mankoff³

¹Department of Radiology, University of Washington, Seattle, WA; ²Department of Statistics, University College, Cork, Ireland; ³Department of Radiology, University of Pennsylvania, Philadelphia, PA and; ⁴Insight Centre for Data Analytics, Cork, Ireland

Corresponding Author:
Mark Muzi, PhD
1959 NE Pacific Street, RR650 Neurology UW356465, Seattle, WA 98195 USA;
E-mail: muzi@uw.edu

Key Words: Fluoromisonidazole, PET imaging, brain cancer, ACRIN 6684, radiomics

Abbreviations: $^{18}$F-fluoromisonidazole ($^{18}$F-FMISO), University of Washington (UW), American College of Radiology Imaging Network (ACRIN), World Health Organization (WHO), Response Assessment in Neuro-Oncology Working Group (RANO), Karnofsky Performance Status (KPS), time-to-progression (TTP), overall survival (OS), overall survival at 1 year from diagnosis (OS-1), 3-dimensional iterative reconstruction method (3D IR), ordered subset expectation maximization reconstruction method (OSEM), point spread function reconstruction method (PSF), standard uptake value (SUV), maximum voxel in a tissue region in SUV units (SUVmax), average SUV from a 1-cm circular ROI centered over the hottest voxel (SUVpeak), maximal activity voxel in a tissue region normalized by the blood activity (T/Bmax), hypoxic volume of pixels in a region that are above T/B = 1.2 hypoxic threshold (HV), gray-level co-occurrence matrix (GLCM), gray-level run-length matrix (GLRLM), gray-level size-zone matrix (GLSZM)

ABSTRACT
Hypoxia is associated with resistance to radiotherapy and chemotherapy in malignant gliomas, and it can be imaged by positron emission tomography with $^{18}$F-fluoromisonidazole ($^{18}$F-FMISO). Previous results for patients with brain cancer imaged with $^{18}$F-FMISO at a single center before conventional chemoradiotherapy showed that tumor uptake via T/Bmax (tissue SUVmax/blood SUV) and hypoxic volume (HV) was associated with poor survival. However, in a multicenter clinical trial (ACRIN 6684), traditional uptake parameters were not found to be prognostically significant, but tumor SUVpeak did predict survival at 1 year. The present analysis considered both study cohorts to reconcile key differences and examine the potential utility of adding radiomic features as prognostic variables for outcome prediction on the combined cohort of 72 patients with brain cancer (30 University of Washington and 42 ACRIN 6684). We used both $^{18}$F-FMISO intensity metrics (T/Bmax, HV, SUV, SUVmax, SUVpeak) and assessed radiomic measures that determined first-order (histogram), second-order, and higher-order radiomic features of $^{18}$F-FMISO uptake distributions. A multivariate model was developed that included age, HV, and the intensity of $^{18}$F-FMISO uptake. HV and SUVpeak were both independent predictors of outcome for the combined data set ($P < .001$) and were also found significant in multivariate prognostic models ($P < .002$ and $P < .001$, respectively). Further model selection that included radiomic features showed the additional prognostic value for overall survival of specific higher order texture features, leading to an increase in relative risk prediction performance by a further 5%, when added to the multivariate clinical model.

INTRODUCTION
Evolving and immature tumor vasculature can lead to chronic tissue hypoxia, which is associated with resistance to radiotherapy and chemotherapy in malignant gliomas, leading to treatment failure and recurrence (1). Chronic hypoxia in patients with glioma has been quantified by using $^{18}$F-fluoromisonidazole ($^{18}$F-FMISO) positron emission tomography (PET) imaging. This tracer is irreversibly reduced and trapped when the level of O$_2$ is below...
3 mmHg—the threshold for the oxygen enhancement effect of ionizing radiation (2). Several studies have shown that metrics of brain tumor hypoxia measured by 18F-FMISO PET were associated with time-to-progression (TTP) and/or overall survival (OS) (3–6). Previous studies at the University of Washington (UW) of 22 patients imaged with 18F-FMISO PET following surgical resection and before radiotherapy with concomitant chemotherapy showed that the maximum 18F-FMISO tissue–blood ratio (T/Bmax) and the hypoxic volume (HV) were associated with shorter OS and TTP (4). However, in a multicenter clinical trial of 42 patients imaged at a similar time point [ACRIN 6684 (5)], T/Bmax and HV were not found to be prognostic, although SUVpeak (defined as the average SUV from a 1-cm-diameter circular region centered on the tumor voxel with greatest activity) predicted OS at 1 year (OS-1) (5). The current analysis considered both the UW and ACRIN 6684 study cohorts with a view to examining differences in the analytical results, and to also explore the utility of agnostic measures of 18F-FMISO uptake using radiomic features to further assess patient risk on the combined cohort of patients.

**METHODOLOGY**

**Patient Characteristics**
The study consisted of 2 cohorts of newly diagnosed patients with high-grade glioma imaged with 18F-FMISO PET after histopathological diagnosis and before initial standard treatment of radiation with concomitant chemotherapy. For the UW cohort, 30 patients (27 WHO grade IV glioblastoma multiforme and 3 WHO grade III anaplastic astrocytoma; median age, 58 years [range, 32 to 77], females, 9; and males, 21) underwent 18F-FMISO imaging studies after surgical resection and before standard treatment (see online supplemental Table S1). Patients were selected and recruited by an experienced neuro-oncologist at the UW-affiliated hospital system on the basis of histopathologic diagnosis, clinical performance status, and presence of residual tumor volume on postsurgical contrast-enhanced T1 magnetic resonance imaging (MRI). Signed informed consent, as approved by the respective Investigational Review Boards and Radiation Safety Committees, was obtained for all patients before 18F-FMISO imaging. All UW patients were followed through periodic calibrations that included Karnofsky Performance Status (KPS) and routine contrast-enhanced MRI surveillance for tumor progression and OS. Tumor progression was determined by RANO criteria for high-grade glioma (7). The majority of these patients (28 out of 30) were included in previous reports examining the relationship of 18F-FMISO imaging for the prediction of TTP and OS (4, 8).

The second cohort (n = 42) was recruited through the American College of Radiology Imaging Network (ACRIN) 6684 trial from 11 participating academic sites including UW, and the patients included in the cohort had histopathologically proven, newly diagnosed glioblastoma multiforme. The trial was approved by the institutional review boards at all participating sites. Before enrollment, all patients signed an informed consent document. Eligible patients had undergone surgical resection before planned standard-of-care treatment of radiation with chemotherapy. In addition, patients could receive investigational agent(s) at any time during the clinical trial. Study candidates were required to have evidence of residual tumor following surgery based on MRI imaging, although the minimal amount of residual tumor was not specified and, unlike the UW cohort, the presence of contrast-enhancement was not required. The ACRIN protocol thus had much less restrictive inclusion criteria. Other key patient characteristics included a median age of 60 years (range, 30–77 years), with 15 females and 27 males. Patients were followed every 3 months for tumor progression and survival for a minimum of at least 1 year or until death (median follow-up period until death, 227 days; range, 74–1034 days), while the patients in the UW group was followed continuously until death and therefore had a longer follow-up. Eleven of the 42 ACRIN patients were alive at the end of the study the follow-up period, which terminated 12 months after the last study patient was imaged. Routine clinical parameters considered for statistical analysis included age, gender, and KPS.

**Scanner Qualification**
The UW tomography (GE Advance, GE Medical Systems, Waukesha, WI) was calibrated on a quarterly basis for conversion of counts to Becquerel per milliliter using the ACR flood phantom containing known activities of 18F imaged separately from a patient and reconstructed using identical parameters as patient emission images. Sample aliquots of the flood phantom were used to cross-calibrate the gamma well counter (Packard COBRA II, Meriden, CT) where patient blood samples were assayed to determine 18F concentration. For the ACRIN study participants (11 institutions), PET/computed tomography (CT) scanners (GE Discovery STE, GE Discovery RX, GE Discovery LS, Siemens ECAT Exact HR+, Siemens Biograph 40, and Siemens Biograph 64) were prequalified by ACRIN using ACR flood phantom scans of known activity and sample patient images and every 2 years following initial accreditation. The phantom scans required reconstruction with identical parameters as patient emission images. Scanners were required to undergo cross-calibration to their local gamma well counter (Packard, Perkin-Elmer, Capintec, Beckman, Baird Atomic) at each participating center for determining blood activity within a week of patient scanning. Cross-calibrations were done at the time of scanner calibration by counting 1-mL aliquots from the calibration phantom following a standardized protocol developed by UW for ACRIN. All qualification data were centrally reviewed. One scanner, in which 3 patients were enrolled, was a prototype Siemens brain MR/PET scanner (no model number) that could not be calibrated with a head coil in place.

**Radiosynthesis**

18F-FMISO was initially prepared at UW using the glycidyl tosylate method (9), and then changed to the method developed by Lim (10) as modified by Adamsen (11). In all cases, the same purification by high-performance liquid chromatography was used. The product-specific activity ranged from 37 to 74 GBq/μmol at the time of injection with >98% radiochemical purity. 18F-FMISO was administered by venous injection of a 10-mL solution of isotonic saline containing <10% (v/v) ethanol USP. The average injected dose for the UW studies was 259 MBq (range, 218–350 MBq). Early studies were done under Radioactive Drug Research Committee, but the majority was performed under
Investigational New Drug approval. The ACRIN patients’ radiotracer doses were supplied by local production of $^{18}$F-FMISO or through Cardinal Health with product characteristics specified by the NCI-held IND (#76042). High-performance liquid chromatography purification was not required. The mean injected dose was 237 MBq (range, 185–285 MBq).

**PET Imaging Acquisition**

For both cohorts, imaging with $^{18}$F-FMISO PET was performed at a clinically relevant time after surgical tissue confirmation of GBM diagnosis and within 2 weeks before the start of conventional therapy. After patient immobilization, venous lines were established in each arm, one for tracer injection and the other for blood sampling. At UW, this was followed by either a 25-minute transmission scan for the GE Advance or for the ACRIN cohort, a low-dose CT scan for PET/CT scanners. For both cohorts, static 20-minute $^{18}$F-FMISO PET emission scans were acquired in 3-dimensional, high-resolution mode starting ~110 min after injection (mean, 118; range 90–155 minutes; n = 72) of $^{18}$F-FMISO. Three venous blood samples were acquired at 5, 10, and 15 minutes after scan initiation for quantitative assessment of hypoxia status by normalizing tissue uptake to the blood concentration as previously described (8). Averaged blood activity was decay-corrected to the injection time and converted to the same units as the scanner (Bq/mL).

**Image Reconstruction**

UW $^{18}$F-FMISO PET emission images were reconstructed by 3D-filtered backprojection (12) with corrections for attenuation, scatter, random events, dead time, and sensitivity, and these resulted in an in-plane spatial resolution of 2.34 mm with a post-reconstruction 6-mm Hanning filter. Most ACRIN studies (39 out of 42) were scanned on PET/CT scanners, which used CT-based attenuation correction. A prototype Siemens brain MR/PET scanner used an estimated attenuation map based on the magnetic resonance (MR) signal (13). Reconstruction of emission images for the ACRIN studies used various iterative 3D techniques (3-dimensional iterative reconstruction method [3D IR], ordered subset expectation maximization reconstruction method [OSEM], point spread function reconstruction method [PSF]), wherein postreconstruction filter sizes ranged from 2 to 6 mm and reconstructed in-plane image resolution varied between 1.25 and 6 mm with corrections for attenuation, scatter, random events, decay, deadtime, and sensitivity (see online supplemental Table 2). Tomography sensitivity was calibrated as described in aforementioned sections.

**Image Analysis**

For the ACRIN cohort, $^{18}$F-FMISO images were analyzed at UW by an experienced team that performed analysis of the original UW cohort and served as the central analysis laboratory for the ACRIN 6684 $^{18}$F-FMISO PET studies. Images acquired with $^{18}$F-FMISO PET were decay-corrected to the time of radiotracer injection and converted to SUV units by normalization of the emission image to the recorded dose and patient weight. Traditionally, volumes of interest (VOIs) were either segmented from the contrast-enhancing volumes on the MR T1+ contrast images and dilated by several voxels or manually constructed on the $^{18}$F-FMISO image based on areas of $^{18}$F-FMISO tracer uptake (4). Both methods proved adequate for routine extraction of $^{18}$F-FMISO uptake information but proved inadequate for radiomic texture feature extraction. Therefore, VOIs over tumor were segmented from either coregistered Gd-enhanced MRI fluid-attenuated inversion recovery images or T2-weighted images obtained within 2 weeks of PET image acquisition, and before conventional therapy. Because $^{18}$F-FMISO is a freely diffusible tracer (14) and tumor hypoxia may exist beyond the margins of Gd-enhancement on MRI, each tumor VOI was segmented to include the entire area of abnormal fluid-attenuated inversion recovery signal hyperintensity taking care not to include postsurgical changes such as the resection cavity and extra-axial fluid collections. An experienced neuroradiologist reviewed each MRI-based tumor volume for all patients. The segmented tumor VOI was then applied to the $^{18}$F-FMISO PET SUV image using PMOD software (PMOD Tech., Version 3.8, Zurich CH) for data extraction. Owing to the relatively large size of the MRI-segmented tumor VOIs (4 cc–750 cc), partial volume correction was not required. In addition, a circular 1-cm-diameter SUVpeak region of interest was centered over the hottest voxel in the tumor. Extracted image data from $^{18}$F-FMISO scans were normalized using the average measured blood activity (Bq/mL) during the scan to produce tissue–blood ratio (T/B) values for all voxels in each image slice. The pixel with maximum T/B value (T/Bmax) and the hypoxic volume (HV) of voxels above the T/B threshold of 1.2, indicating significant hypoxia, were determined to quantify the volume of tissue hypoxia in each tumor region (4). To include the 3 patients scanned on the MRI/PET scanner, their image data were scaled by the average ratio of blood to cerebellar cortex activity, as these values are highly correlated (8) for standard PET imaging. Conventional PET and hypoxia parameters extracted from tumor regions included VOI volume, SUVmax, SUVmean, SUVpeak, T/Bmax, HV, and TLH (total lesion hypoxia = SUVmean × VOI).

**Radiomic Analysis**

We evaluated a set of radiomic features assessing texture characteristics of the hypoxia distribution of the 72 patients with brain cancer from the 2 cohorts who had $^{18}$F-FMISO PET studies with concurrent blood sampling. An adaptive thresholding method (30% of SUVmax) was then applied to the segmented tumor VOIs, yielding volumes that were restricted from void regions, but including voxels with $^{18}$F-FMISO uptake (Figure 1). Preprocessing of $^{18}$F-FMISO emission images for radiomic analysis included spatial normalization of the image data and tumor VOI segmentation mask, based on the segmentations used for standard analyses noted above, and interpolated in 3 dimensions at the lowest resolution of the 2 cohorts (1 mm) via trilinear interpolation. The exploratory investigation on the association to outcome considered the following biomarkers detailed in the Image Biomarker Standardization Initiative (IBSI) reference manual (15) for analysis: 12 first-order statistical variables reflecting intensity (IBSI Section 3.4), 20 histogram features of the discretized uptake values (IBSI Section 3.5), 22 texture features of the gray-level co-occurrence matrix (GLCM) (IBSI Section 3.6), 16 GLRLM-based features emphasizing discretized images with long-run lengths, or sequences of voxels with the
same discretized intensity (IBSI Section 3.7) and 16-GLSZM-based features evaluating entropy (or uniformity) in the distribution of groups of connected voxels with the same discretized intensity (IBSI Section 3.8). Following the guidelines of Vallières et al. (16) on responsible radiomics, the online supplemental online data provides a description of the radiomic extraction process and ontology of the radiomic features used in this report. An open-source software implementation of radiomic features in R (mia 1.0.4, https://github.com/ericwol/mia) was used for radiomic quantitation of the 18F-FMISO uptake distribution. Certain features based on the GLCM (eg, contrast, entropy, dissimilarity, homogeneity) have been shown to have high predictive value for PET imaging (17–19).

In total, there were 97 variables for each patient that were considered for model selection—11 conventional clinical and PET variables and 86 radiomic PET features.

Statistical Analysis
Comparisons between cohorts for parameters such as uptake time and the 97 other variables were made using a 2-sample Wilcoxon (Mann–Whitney) test for equal distributions (20). To examine the relationship of clinical, PET, and textural biomarkers to OS, the following methods were used.

Univariate Model. Univariate Cox proportional model analyses were applied to all variables on each cohort and the combined cohort to assess the statistical significance of each variable separately for OS prognosis. In these analyses, significance levels were set at the 5% level with Bonferroni correction to limit the risk of a type 1 error (false positives).

Multivariate Model. Multivariate prognostic modeling was performed in 2 steps: first to explore the prognostic value of a model that comprises conventional clinical variables, thereby selecting the most parsimonious base reference clinical model, before evaluating the potential for prognostic improvement by adding radiomic features.

In the first step, we performed optimal subset selection by evaluating all possible Cox proportional hazards models comprising combinations of the clinical variables to identify the AIC (Akaiake information criterion)-minimizing (21) model from only conventional clinical and PET variables (see online supplemental Figure 1). The analysis resulted in a base clinical model with a parameter combination of age, HV of the tracer, and the intensity of the tracer (T/Bmax and SUVpeak).

In the second step, preliminary feature elimination was performed to remove radiomic variables that had a strong correlation with other clinical or radiomic variables. We eliminated features with a Pearson correlation >80% in absolute value. Forward stepwise selection (22) was then carried out to yield a second multivariate Cox model. This search allowed selection in both forward and backward directions of any remaining conventional or radiomic variables to the base clinical model from the first step. This step tested the additional prognostic benefit from radiomic features.

The hazard ratio along with its 95% confidence interval and associated P-value based on Wald’s statistic were reported for the Cox models. Prognostic significance was assessed by bootstrapped median P-value (using 1000 bootstrap resamples), and considered both at the 5% significance level and with Bonferroni correction. The likelihood ratio test (LRT) was used to confirm statistical significance of the addition of textural variables to the base clinical model (23, 24). The concordance index of the multivariate prognostic model assessed potential gain in predictive performance over the base clinical model. Kaplan–Meier survival curve estimates from diagnosis to time-to-death were generated. All analyses were done using R (CRAN R-Project; https://cran.r-project.org).

Figure 1. Image segmentation for data extraction. The tumor region is segmented from the magnetic resonance imaging (MRI) fluid-attenuated inversion recovery (FLAIR) (blue line) as a region of hyperintensity (A). The region is applied to the 18F-fluoromisonidazole (18F-FMISO) standard uptake value (SUV) image (B). Adaptive thresholding of the 18F-FMISO SUV image yields the tumor volume for radiomic analysis that avoids regions of low uptake (C).
RESULTS

Considering patient characteristics of the 2 cohorts, the UW one had lower KPS (mean, 80; range, 60–100 KPS) than the ACRIN cohort (mean, 87; range, 70–100 KPS; Wilcoxon test, \( P = .028 \)). The uptake time between injection and scanning for the UW cohort (mean, 128 minutes; range, 100–155 minutes) was longer than the ACRIN cohort on the average (mean, 111 minutes; range, 90–145 minutes; Wilcoxon test, \( P < .001 \)). While the UW group had a combination of grades III and IV patients, removing grade III patients did not yield significantly different results. Methodological differences in attenuation, reconstruction algorithm, and filter size varied significantly between cohorts (see online supplemental Table 1). Conventional PET uptake parameters for \(^{18}\text{F-FMISO} \) (SUVmax, SUVpeak, T/Bmax, and HV) appear in Table 1. A notable difference in distribution between UW and ACRIN cohorts was observed for only KPS, on the basis of Bonferroni-corrected Wilcoxon tests. At the 5\% significance level, without Bonferroni correction, only one other variable (run-length nonuniformity, a GLRLM-based texture feature) yielded a difference in cohort distributions (\( P = .002 \)). Reconstruction methods are known to have a substantial effect on radiomic features (25, 26), and there were clearly reconstruction differences between the 2 cohorts.

Univariate Analysis

PET parameters were assessed on the basis of overall patient survival status at last follow-up, including censoring information. Univariate Cox proportional modeling combining the 2 cohorts showed that 24 variables were significant without Bonferroni correction, but only 9 were significant for prognosis at the 5\% significance level with Bonferroni correction (Table 2). In particular, age, SUVpeak, and HV were all independent predictors of outcome in the combined data set after Bonferroni correction (\( P < .001 \) for all). Comparing the results between the 2 data sets, the UW cohort had all 9 variables from Table 2 remain significant at the 5\% significance level, with HV and SUVpeak both remaining independent predictors after Bonferroni correction. Six of the 9 variables were significant independent predictors at the 5\% significance level in the ACRIN group (Table 2), where HV, SUVpeak, and GLCM energy were not significant.

Multivariate Analysis

Following selection of the best parameter subset for model composition, the most parsimonious base conventional parameter Cox model with optimal AIC value included age, HV, SUVpeak, and T/Bmax, and is summarized in Table 3. All 4 variables in this model were statistically significant prognostic markers at the 5\% significance level, with HV and age remaining significant after Bonferroni correction, using threshold significance level 0.05/\( P_{\text{clinical}} < 0.005 \), where \( P_{\text{clinical}} = 11 \) clinical variables.

The second multivariate prognostic model obtained by extending the base conventional model is presented in Table 4. It illustrates the enhanced prognostic potential obtained by adding texture features to the base model. Although clinical variables were permitted in the model selection procedure, texture features were selected over conventional variables by the more objective stepwise selection algorithm. This supports a potential benefit in introducing texture analyses for postresection risk characterization in glioma. The likelihood ratio test confirmed the statistical significance (\( P < .010 \)) of the additional contribution of this set of texture variables. We observed a reasonably high concordance index of 77.4\% for the model of Table 2. This index measures the pairwise probability of assessing a lower patient risk, based on

### Table 1. \(^{18}\text{F-FMISO} \) Tumor Uptake Values

<table>
<thead>
<tr>
<th></th>
<th>SUVmax (g/mL)</th>
<th>SUVpeak (g/mL)</th>
<th>T/Bmax (unitless)</th>
<th>HV (mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UW</td>
<td>Mean</td>
<td>3.15</td>
<td>2.53</td>
<td>2.13</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.11</td>
<td>1.04</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>2.94</td>
<td>2.15</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>6.53</td>
<td>5.64</td>
<td>4.26</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>1.85</td>
<td>1.34</td>
<td>1.27</td>
</tr>
<tr>
<td>ACRIN</td>
<td>Mean</td>
<td>3.20</td>
<td>2.49</td>
<td>2.10</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.14</td>
<td>0.84</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>2.96</td>
<td>2.26</td>
<td>2.02</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>6.26</td>
<td>5.00</td>
<td>4.14</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>1.80</td>
<td>1.45</td>
<td>0.95</td>
</tr>
<tr>
<td>ALL</td>
<td>Mean</td>
<td>3.18</td>
<td>2.51</td>
<td>2.11</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.12</td>
<td>0.92</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>2.94</td>
<td>2.24</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>6.53</td>
<td>5.64</td>
<td>4.26</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>1.80</td>
<td>1.34</td>
<td>0.95</td>
</tr>
</tbody>
</table>

\(^{18}\text{F-FMISO} \) PET uptake parameters for the single-center UW, multicenter ACRIN 6684 and the combined cohorts of all patients. T/Bmax is the maximum tumor voxel normalized by the blood activity at the time of scanning, and HV refers to the hypoxic volume of pixels above the T/B threshold of 1.2 indicating significant hypoxia.
DISCUSSION

We investigated the association of 18F-FMISO PET hypoxia parameters with survival between a single-center (UW) cohort and a multicenter (ACRIN 6684) cohort. Although both studies and higher curves respectively) was obtained by maximizing the combined cohort into higher- and lower-risk groups (lower blue lines) multivariate Cox models. Dichotomous separation of obtained from the base (dashed red lines) and extended (solid blue lines) multivariate Cox models. Dichotomous separation of the combined cohort into higher- and lower-risk groups (lower and higher curves respectively) was obtained by maximizing the statistical separation in terms of the log-rank test statistic. The separation of high-risk from low-risk patients was statistically significant for both models (log-rank test, \( P < .001 \)).

Figure 2 shows the Kaplan–Meier OS curve estimates obtained from the base (dashed red lines) and extended (solid blue lines) multivariate Cox models. Dichotomous separation of the combined cohort into higher- and lower-risk groups (lower and higher curves respectively) was obtained by maximizing the statistical separation in terms of the log-rank test statistic. The separation of high-risk from low-risk patients was statistically significant for both models (log-rank test, \( P < .001 \)).

The model, given the longer survival time, and is equivalent to the AUC value in an ROC analysis. The addition of the texture features to the base clinical model led to an increase in relative risk prediction performance, from 72.2% to 77.4%.

Table 2. Univariate Analysis Hazard Ratios (Pvalue)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Combined</th>
<th>ACRIN Cohort</th>
<th>UW Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.94 (.001)</td>
<td>2.45 (.002)</td>
<td>1.58 (.017)</td>
</tr>
<tr>
<td>SUVpeak</td>
<td>1.71 (.001)</td>
<td>1.46 (.070)</td>
<td>2.24 (.001)</td>
</tr>
<tr>
<td>HV</td>
<td>1.80 (.001)</td>
<td>0.98 (.957)</td>
<td>1.85 (.001)</td>
</tr>
<tr>
<td>SUVmean</td>
<td>1.87 (.001)</td>
<td>2.03 (.001)</td>
<td>1.70 (.004)</td>
</tr>
<tr>
<td>Median</td>
<td>1.82 (.001)</td>
<td>2.36 (.001)</td>
<td>1.53 (.012)</td>
</tr>
<tr>
<td>p10</td>
<td>1.71 (.001)</td>
<td>2.08 (.006)</td>
<td>1.48 (.026)</td>
</tr>
<tr>
<td>p90</td>
<td>1.71 (.001)</td>
<td>1.64 (.002)</td>
<td>2.02 (.001)</td>
</tr>
<tr>
<td>GLCM Energy</td>
<td>1.73 (.001)</td>
<td>1.27 (.451)</td>
<td>1.73 (.001)</td>
</tr>
<tr>
<td>RMS</td>
<td>1.88 (.001)</td>
<td>1.99 (.001)</td>
<td>1.72 (.003)</td>
</tr>
</tbody>
</table>

Variables that were found statistically significant in univariate survival analyses after Bonferroni correction, with associated \( P \)-values for the combined, ACRIN, and UW cohorts respectively. \( P \)-values in bold indicate significance with respect to the Bonferroni correction threshold of 0.05/97 = 0.0005. Median, p10, p90, and RMS, respectively, correspond to the sample median, 10th and 90th percentiles, and root mean square (quadratic mean) of the tumor SUV values.

DISCUSSION

We investigated the association of 18F-FMISO PET hypoxia parameters with survival between a single-center (UW) cohort and a multicenter (ACRIN 6684) cohort. Although both studies showed prognostic value for 18F-FMISO, the 2 studies found the association with different measures of tracer uptake. Previous results on a subset of 22 patients of the UW cohort showed that age, T/Bmax, and HV were significant in predicting OS (4) in univariate analyses. A similar result for OS was shown here using a larger UW cohort of 30 patients (Table 2), while in the multicenter ACRIN cohort, only SUVpeak was significant in univariate analyses for OS-1. Although all patients in both studies had pathologically proven brain tumors imaged before conventional therapy, the 2 cohorts appear different with respect to traditional 18F-FMISO uptake parameters, particularly HV. A potential explanation of the differences between the 2 individual cohorts most likely lies in patient selection. The UW cohort had larger hypoxic volumes of 18F-FMISO and incidentally had lower KPS, indicating greater disease involvement, although there were no differences in postsurgery contrast-enhancement volume on MRI T1 between the UW cohort and the ACRIN cohort. Unlike the UW group, a few of the ACRIN patients had no 18F-FMISO uptake and therefore no measurable hypoxic volume. For example, hypoxic volumes were <4 cc in nearly 30% (12 out of 42) of patients in the ACRIN cohort compared with in 7% (2 out of 30) of patients in the UW cohort. Low residual tumor volume at the time of 18F-FMISO imaging may not permit adequate assessment of hypoxia by any measure, such as tracer distribution, intensity, or structural characteristics of spatial uptake distribution. A smaller HV limits image count statistics in the ACRIN data set and along with the added noise of variability in cross-calibration where multiple sites differ in scanner, image reconstruction, counters, and blood sampling/counting skill levels may well explain why T/Bmax and HV did not work better for univariate ACRIN data set analysis. In addition, the T/B threshold of 1.2 that defines the volume of hypoxic voxels within the tumor was developed using filtered backprojection image data, and could well be different for images using an estimation reconstruction analysis.

Heterogeneity of treatment in the ACRIN cohort versus relatively uniform treatment in the single-center UW study may also affect study outcome. The ACRIN cohort had many patients with alternative therapies (Lomustine, Paclitaxel, Poliglumex, Selinexor, Vorinostat, Dasatinib, Carmustine, Bevacizumab) during and following conventional therapy; the UW cohort had

Table 3. Base Multivariate Cox Model Analysis

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effect</th>
<th>HR (95% CI)</th>
<th>Median P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.74</td>
<td>2.10 (1.52, 2.91)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>HV</td>
<td>0.54</td>
<td>1.72 (1.23, 2.41)</td>
<td>.002</td>
</tr>
<tr>
<td>SUVpeak</td>
<td>0.66</td>
<td>1.93 (1.18, 3.17)</td>
<td>.006</td>
</tr>
<tr>
<td>TBmax</td>
<td>-0.50</td>
<td>0.61 (0.39, 0.96)</td>
<td>.023</td>
</tr>
</tbody>
</table>

The base multivariate Cox model was comprised of 4 conventional variables, showing predictor effects, associated hazard ratio (HR) with associated 95% confidence interval (CI), and bootstrap median \( P \)-values obtained from 1000 bootstrap resamples. All 4 variables are significant at the 5% level; \( P \)-values in bold indicate significance after bonferroni correction (\( P < .005 \)). This model has a concordance index of 0.722.
fewer agents (Carboplatin, Irinotecan, Bevacizumab) and they were used only after conventional therapy had been completed. Treatments during and after $^{18}$F-FMISO imaging may modulate the time to progression and potentially delay death, altering the survival outcome variables. Notable differences between the early survival characteristics of the 2 cohorts, the degree of tumor burden as observed with $^{18}$F-FMISO HV, and the inclusion of alternative therapies may partially explain these findings. Both

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effect</th>
<th>HR (95% CI)</th>
<th>Median P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.84</td>
<td>2.31 (.165, 3.25)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>HV</td>
<td>0.67</td>
<td>1.95 (.128, 2.97)</td>
<td>.006</td>
</tr>
<tr>
<td>SUVpeak</td>
<td>0.97</td>
<td>2.63 (.152, 4.56)</td>
<td>.001</td>
</tr>
<tr>
<td>TBmax</td>
<td>-0.75</td>
<td>0.47 (.028, 0.80)</td>
<td>.010</td>
</tr>
<tr>
<td>HIST CoV$^a$</td>
<td>1.29</td>
<td>3.65 (.172, 7.73)</td>
<td>.001</td>
</tr>
<tr>
<td>HIST QCOD</td>
<td>-0.59</td>
<td>0.55 (.036, 0.86)</td>
<td>.015</td>
</tr>
<tr>
<td>HIST Kurtosis</td>
<td>-0.55</td>
<td>0.58 (.034, 0.98)</td>
<td>.048</td>
</tr>
<tr>
<td>GLRLM Long runs emphasis</td>
<td>-1.26</td>
<td>0.28 (.014, 0.58)</td>
<td>.001</td>
</tr>
<tr>
<td>GLSZM Zone size entropy</td>
<td>0.74</td>
<td>2.10 (.118, 3.71)</td>
<td>0.020</td>
</tr>
<tr>
<td>GLSZM Low gray-level zone emphasis</td>
<td>0.41</td>
<td>1.51 (.094, 2.45)</td>
<td>0.110</td>
</tr>
</tbody>
</table>

Extended multivariate Cox model obtained by selecting additional variables (conventional or radiomic) to the base model, showing predictor effects, associated hazard ratio (HR) with associated 95% confidence interval (CI), and bootstrap median P-values obtained from 1,000 bootstrap resamples. Nine of the 10 variables are significant at the 5% level; P-values in bold indicate significance after Bonferroni correction. This model has a concordance index of 0.774. The likelihood ratio test confirmed the significant contribution ($P = .010$) of the additional set of radiomic variables.

$^a$ HIST CoV, HIST QCOD, and HIST Kurtosis, respectively refer to the coefficient of variation (CoV), quartile coefficient of dispersion (QCOD), and kurtosis of the histogram of discretized uptake values; these 3 features evaluate various aspects of discretized uptake distribution variability. GLRLM long runs emphasize discretized images with long-run lengths (ie, sequences of voxels with the same discretized intensity). GLSZM zone size entropy evaluates entropy (or uniformity) in the distribution of groups of connected voxels with the same discretized intensity. GLSZM low gray-level zone emphasizes discretized images containing more low-intensity regions.

**Figure 2.** Survival analysis for the base and extended multivariate models on the combined cohorts. Kaplan–Meier analysis of overall survival (OS) with stratification into lower- and higher-risk groups (the higher and lower curves, respectively) based on maximized statistical separation in terms of the log-rank test statistic ($P < .001$), for the base 4-variable clinical model comprised age, hypoxic volume (HV), T/Bmax (maximal activity voxel in a tissue region normalized by the blood activity) and SUVpeak (average SUV from a 1-cm circular ROI centered over the hottest voxel; dashed red lines), and its extension using additional radiomic features (solid blue lines). The figure shows how the latter model led to different patient classification.
independent studies were carried out on limited sample sizes, and as such, differences in (or lack of) significance for some of the variables may also be a direct consequence of low statistical power.

The present analysis on the aggregated cohort highlighted the preserved significance of both sets of markers from each cohort. Each of the 4 variables in the multivariate base model of Table 3 was a statistically significant jointly prognostic marker at the 5% significance level. The multivariate model applied to the combined cohort thus assesses complementary aspects of $^{18F}$-FMISO uptake by including the tracers’ overall distribution (HV) and intensity of uptake (SUVpeak and T/Bmax).

Texture features, such as GLCM entropy and other radiomic measures could improve the prediction of patient outcome. The addition of 6 radiomic variables, identified via a formal statistical stepwise selection process, to the initial 4-variable clinical model led to an added 5% in concordance index, and their inclusion contributed to a multivariate Cox proportional hazards model that was deemed statistically significant ($P < .001$).

**LIMITATIONS**

The modest sample size and constraints on follow-up in this study are important to highlight. These issues arose owing to the limited access to FMISO radiotracer imaging for patients with glioma during the period following pathological confirmation and before conventional therapy. ACRIN funding limited patient follow-up to be a minimum of 1 year. At the end of the study, 9 patients were alive and thereafter lost to follow-up, and thus were censored.

Selection of the prognostic multivariate models presented in Tables 3 and 4 could be performed in a number of different ways. The base multivariate clinical model of Table 3 was selected using conventional best-subset selection, an exhaustive comparison of all possible models on the basis of some goodness-of-fit criterion adjusted for model dimension. We use AIC for this test, as it is one of the most commonly used metrics and a reasonable choice considering the relatively small number of candidate variables. The extended model of Table 4 is subject to selection bias, as it was constrained to include the 4 conventional variables of the base model of Table 3. As such it should be validated on an independent $^{18F}$-FMISO PET-imaged glioma cohort. The modest sample size prevented the ability to carry out convincing training and test validation. A much larger cohort might have facilitated such an approach and thereby reduced the lingering risk of optimistic bias associated with retrospective analysis of small data sets. However, our models can be considered as demonstrations of prognostic potential, as an illustration of the opportunity for potential prognostic enhancement provided by radiomic analysis of $^{18F}$-FMISO imaging data when combined with standard clinical models.

**CONCLUSIONS**

Differing significance of quantitative metrics for $^{18F}$-FMISO uptake resulted between a single-center cohort and a multicenter cohort of glioma patients for the assessment of hypoxia using $^{18F}$-FMISO PET before combined radiotherapy and chemotherapy. These may be explained by single-center patient selection bias for patients in the UW cohort with larger tumors that allow adequate assessment of hypoxia and by a substantial subset of multicenter patients with low or no HV in the other cohort. The present analysis on the aggregated cohort highlighted the preserved significance of $^{18F}$-FMISO PET markers from both sets. Multiparametric modeling on the combined cohort showed that all of SUVpeak, HV and T/Bmax together with age have a significant predictive value when combined in a multivariate prognostic model. Our results also showed that adding selected radiomic features to this prognostic model further increased predictive performance for the combined group of patients and might be considered in future studies and potential future clinical applications.

**Supplemental Materials**

Supplemental Data: [https://doi.org/10.18383/j.tom.2019.00023. sup.01](https://doi.org/10.18383/j.tom.2019.00023. sup.01)

This study was supported by NIH/NCI R50-CA211270, P01-CA042045, U01-CA079778, U01-CA080098, and Science Foundation Ireland Grants PI 11/1027 and 12/RC/2289-F2.

Disclosure: The authors have nothing to disclose.


MRI-Based Deep Learning Segmentation and Radiomics of Sarcoma in Mice

M. D. Holbrook1, S. J. Blocker1, Y. M. Mowery2, A. Badea1, Y. Qi1, E. S. Xu2, D. G. Kirsch2, G. A. Johnson1, and C. T. Badea1

Departments of 1Radiology, Center for In Vivo Microscopy; and 2Radiation Oncology, Duke University Medical Center, Durham, NC

Corresponding Author:
C. T. Badea, PhD
Center for In Vivo Microscopy, Duke University Bryan Research Building, Room 161F, 311 Research Drive, DUMC Box 3302, Durham, NC 27710;
E-mail: cristian.badea@duke.edu

Key Words: Radiomics, MRI, preclinical imaging, deep learning, segmentation

Abbreviations: Magnetic resonance imaging (MRI), convolutional neural network (CNN), receiver operating curve (ROC), area under the curve (AUC), volume overlap error (VOE), radiation therapy (RT), neural network (NN), support vector machine (SVM)

ABSTRACT

Small-animal imaging is an essential tool that provides noninvasive, longitudinal insight into novel cancer therapies. However, considerable variability in image analysis techniques can lead to inconsistent results. We have developed quantitative imaging for application in the preclinical arm of a coclinical trial by using a genetically engineered mouse model of soft tissue sarcoma. Magnetic resonance imaging (MRI) images were acquired 1 day before and 1 week after radiation therapy. After the second MRI, the primary tumor was surgically removed by amputating the tumor-bearing hind limb, and mice were followed for up to 6 months. An automatic analysis pipeline was used for multicontrast MRI data using a convolutional neural network for tumor segmentation followed by radiomics analysis. We then calculated radiomics features for the tumor, the peritumoral area, and the 2 combined. The first radiomics analysis focused on features most indicative of radiation therapy effects; the second radiomics analysis looked for features that might predict primary tumor recurrence. The segmentation results indicated that Dice scores were similar when using multicontrast versus single T2-weighted data (0.863 vs 0.861). One week post RT, larger tumor volumes were measured, and radiomics analysis showed greater heterogeneity. In the tumor and peritumoral area, radiomics features were predictive of primary tumor recurrence (AUC: 0.79). We have created an image processing pipeline for high-throughput, reduced-bias segmentation of multiparametric tumor MRI data and radiomics analysis, to better our understanding of preclinical imaging and the insights it provides when studying new cancer therapies.

INTRODUCTION

Because imaging is a standard means for assessing disease state and therapeutic response in clinical oncology, small-animal imaging for coclinical cancer trials enhances the simulation of clinical practice in animals. High-resolution images can noninvasively describe tumor morphology and composition, as well as how tumors change over time or with treatment.

Magnetic resonance imaging (MRI) is the clinically preferred method for imaging soft tissue sarcomas owing to its excellent soft tissue contrast (1). Assessing treatment response requires tumor measurements that are both accurate and precise. Manual segmentations suffer from variability that is in part due to individual human rater biases. In the clinic, tumor regions are often identified with input provided by radiologists or radiation oncologists. However, advances in computer vision have made automated segmentation processes possible. Specifically, segmentation algorithms based on convolutional neural networks (CNNs) have shown comparable efficacy in identifying tumors as other automated methods (2). Several CNN-based methods have been proposed for tumor segmentation from multicontrast MRI, based on both 2D slices (3) or 3D volumes (4). Many current architectures for tumor segmentation use a patch-based approach, in which a 2D or 3D patch is processed by convolutional and fully connected layers to classify the center pixel of the patch (3, 5). Other networks operate semantic-wise by classifying each pixel in an input image or a patch using fully convolutional networks or U-nets (2, 6). Deep learning solutions are particularly attractive for processing multichannel, volumetric image data, where conventional processing methods are often computationally expensive (7).

The extraction of high-dimensional biomarkers using radiomics can identify tumor signatures that may be able to monitor disease progression or response to therapy or predict treatment outcomes (8, 9). Radiomics analysis generates complex high-dimensional data, and trends are often difficult to extract. The
utility of radiomics benefits greatly from the use of machine learning algorithms (10). In this way, radiomics provides an exciting approach for identifying and developing imaging biomarkers in the context of precision medicine.

Our group has established quantitative imaging techniques for the preclinical arm of a coclinical sarcoma trial studying the treatment synergy between immune checkpoint blockade with an antibody against programmed cell death protein 1 (PD-1) and radiation therapy (RT) in a genetically engineered mouse model of soft tissue sarcoma (11). Our first objective was to develop and evaluate a deep learning method based on CNNs to perform automatic tumor segmentation of preclinical MRI data acquired in these sarcomas. The high-throughput capacity of a fully automated segmentation pipeline offers a significant time advantage in a large-scale study, as is often the case when studying cancer therapeutics. Even more importantly, a CNN-based segmentation protocol has the advantage of removing observer bias, which can have a significant effect on defining tumor tissue in magnetic resonance (MR) images (12).

The second objective was to perform radiomics analyses on the acquired MRI data sets. Recently, radiomics analysis has been successfully applied to clinical sarcoma data (13). There is evidence that radiomics features extracted from MRI may serve as biomarkers for predicting overall survival in patients with soft tissue sarcomas (14). To the best of our knowledge, no radiomics studies exist on mouse models of sarcomas. This study describes and evaluates our small-animal MRI-based image analysis pipeline on sarcomas treated with RT, including automated tumor segmentation and radiomics analysis.

METHODS

**Sarcoma Model and Experimental Protocol**

The preclinical arm of our coclinical trial uses a genetically engineered model of soft tissue sarcoma developed in p53R172H mice. Primary sarcoma lesions were generated in the hind limb by intramuscular delivery of Adeno-Cre followed by injection of the carcinogen 3-methylcholanthrene (p53/MCA model) (15). Tumors resembling human undifferentiated pleomorphic sarcoma developed 8–12 weeks after injection. Imaging studies were initiated when tumors were palpable (>100 mg), continuing through subsequent stages of disease progression. Mice were killed at the end of the study either once the tumor burden became excessive or 6 months after surgery. Excessive burden was defined as recurrent tumors >1.5 cm in length or presence of large lung metastasis.

Three MRI images, 1 T2-weighted, and 2 T1-weighted (before and after contrast injection), were obtained 1 day before delivering RT (20 Gy) on a small-animal irradiator (Precision X-ray X-RAD 320) with 100 kV. One week later, the mice were reimaged with MRI using the same imaging protocol. After the second MRI, the primary tumor was surgically removed by amputating the tumor-bearing hind limb, and mice were followed for up to 6 months. Some mice developed local recurrence of the primary tumors near the site of amputation or developed distant metastases.

**Multicontrast MRI**

All MR studies were performed on a 7.0 T Bruker Biospec small-animal MRI scanner (Bruker Inc., Billerica, MA), with a 20-cm bore, equipped with an AVANCE III console, using Paravision 6.0.1, and a 12-cm-inner-diameter gradient set capable of delivering 440 mT/m. Tumor images were acquired using the 4-element surface coil array (receive), coupled with the 72-mm linear volume (transmit) coil. All animal handling and imaging procedures were performed according to protocols approved by the Duke Institutional Animal Care and Use Committee (IACUC). Acquisition parameters have previously been described in detail (11) and are as follows:

1. **T1-weighted:** A 2D T1 FLASH sequence was performed with echo time of 4.5 milliseconds and repetition time of ~0.9 seconds. A fixed field of view of 28 x 28 mm (read and phase) was selected to ensure full coverage of the tumor volume with reasonable margins. In-plane resolution was 100 µm over 60, 300-µm thick axial slices (slice direction typically along the tibia). Three images were acquired and averaged to reduce the effects of motion (which are limited in the hind limb). The flip angle was 30°.

2. **T2-weighted:** A 2D T2 TurboRARE sequence was performed with an effective echo time of 45 milliseconds and repetition time of ~8.6 seconds. The echo spacing for the T2TurboRARE sequence was 15 milliseconds. Scans were performed with identical slice geometry and field of view to the preceding T1 sequence. As with the T1 images, 3 averages were acquired, with a RARE factor of 8.

Following the acquisition of nonenhanced scans, T1 contrast enhancement was achieved via injection of Gd-DTPA at 0.5 mmol/kg via the tail vein catheter, which was placed under anesthesia before imaging. Contrast was injected at 2.5 mL/min and allowed to circulate for 3 minutes before a second T1-weighted scan to allow peak enhancement. The MR images were bias-corrected in a 3D Slicer using the N4 algorithm (16). The sarcoma tumors were next segmented using the T2-weighted images in 2 steps: first, semiautomatic segmentation was performed via the 3D Slicer with the GrowCut tool (17), and second, an observer refined the segmentations by hand. Two researchers acted as observers to create binary segmentation labels with initial guidance from a radiation oncologist experienced with mouse models. In total, 70 manual segmentations were created to serve as the ground truth to train the CNN, including both pre- and post-RT images. Examples of each multicontrast MR image together with the tumor segmentation used as labels for training are shown in Figure 1A. An overview of the data used for this study is given in Table 1. Although we have used 79 mice with tumors, only 62 mice had 2 MRI scans and 42 mice qualified for radiomics analysis.

**CNN-Based Segmentation**

We have implemented a 3D fully convolutional U-net network similar to Ronneberger et al. (18) using Tensorflow (19) to segment soft tissue sarcomas in mice1. To assess the utility of collecting multiple scans with different contrasts, we have compared the segmentation performance of networks trained on

---
1Our open source code for segmentation and radiomics analysis can be found at: https://github.com/mdholbrook/MRI_Segmentation_Radiomics
2 sets of images. The network is presented with inputs of either only T2-weighted images or the MRI images with 3 different contrasts (see Figure 1A). In the case of multicontrast segmentation, the 3 images are concatenated together as channels of a single image, i.e., the network takes a single, 4-dimensional (read, phase, slice, channel) input. Image intensities are normalized on each image volume to map voxel values to within a standard, zero-centered reference scale. This mapping serves to address differences in bias between images of similar contrast and optimize image values for consistent CNN processing. Our network operates on 3D image patches (dimensions, in voxels: 142, 142, 18) selected out of larger image volumes (280, 280, 60). The output of the network is a single 3D segmentation map, showing the probability that a given pixel belongs to the background or foreground (tumor). Thresholding this map yields a binary segmentation from which the tumor volume and radiomics features are determined.

The network comprises 2 halves: an encoder with convolutional layers and max pooling and a decoder layer with deconvolutional layers and up-sampling operations. The structure of the network is shown in Figure 2. The use of multiple max pooling operations (size: $2 \times 2 \times 2$) allows for detection of multiscale features using small, $3 \times 3 \times 3$, convolution kernels, greatly increasing the computational efficiency of the network. The number of convolution filters and, by extension, feature maps, are increased after pooling to preserve information found at finer resolutions. Activation layers after each convolution operation were set as rectilinear activation units. The final activation is a sigmoid function, setting the network’s output to be within the range of 0 to 1.

To increase the amount of information available to the decoder portion of the network, skip connections are present. Skip connections take a set of feature maps from the encoder and concatenate them with feature maps in the decoder. These connections reintroduce higher frequency data directly into the decoder.

---

### Table 1. Overview of Study Mice and Scan Segmentations Used for Training and Radiomics

<table>
<thead>
<tr>
<th>Total Mice</th>
<th>Mice with 2 scans</th>
<th>Mice for radiomics</th>
<th>MR Sets</th>
<th>Manual Segmentations</th>
<th>Unique Manually Segmented Scans</th>
<th>Overlap with radiomics scans</th>
<th>K-Folds Validation Size</th>
<th>K-Folds Test Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>79</td>
<td>62</td>
<td>42</td>
<td>141</td>
<td>70</td>
<td>49</td>
<td>39</td>
<td>6</td>
<td>12</td>
</tr>
</tbody>
</table>

a These mice were not excluded for surgical complications or other health reasons.

b 21 of these scans were segmented twice, once by each reader for a total of 70 segmentations.
and have been shown to increase accuracy for segmentation tasks (18). We have tested the performance of this network both with and without skip connections.

Care was taken to prevent the network from overfitting. To this end, spatial dropout was applied before max pooling and upsampling layers. Spatial dropout is a variation of classical neural network dropout that is more applicable to CNNs. Spatial dropout will randomly zero out feature maps, forcing the network to learn redundancy in identifying important image features rather than overfitting to the peculiarities native to the training data (20, 21).

The selection of an appropriate loss function is critical for designing an effective CNN model. We chose to train and compare models using 2 loss metrics: Dice loss (22), which is a measure of similarity between the prediction and label, and cross entropy (23) which is based on the distributions of the prediction and label images. Both are popular cost functions for segmentation tasks (3, 4, 6, 18). The Dice score for predictions \( P \in \{0, 1\} \) and the expert’s consensus ground truth \( T \in \{0, 1\} \) are defined as:

\[
D = \frac{2|P \cap T|}{|P| + |T|}
\]

where \(|P|\) and \(|T|\) represent the cardinality of the prediction and ground truth sets, respectively. Predictions that perfectly match the ground truth will have a Dice coefficient of 1, whereas predictions with little intersection with the ground truth will have a score near 0. Dice loss requires the use of binary inputs, so network outputs were thresholded by 0.5 before loss is computed.

For use as a loss function, we used the Dice score minus one. Cross entropy loss is computed as the measure of similarity between estimated probabilities and ground truth.

We have trained and tested our network using a single MR image contrast (T2-weighted) and all 3 contrasts (T1-weighted, T1-weighted with contrast, and T2-weighted). Visually, the T2-weighted images show the highest contrast for tumor detection (Figure 1A); however, we aimed to determine the utility of including the T1-weighted images for tumor segmentation, as they are typically included in clinical MRI acquisition protocols.

In total, 8 networks were trained and evaluated. These networks iterated combinations of 3 parameters: Dice versus cross entropy loss functions, networks with skip connections versus those without, and multicontrast versus T2-only MR images as inputs. Training data comprised a random selection of 70% training data, 20% validation data, and 10% test data. All networks were trained for 600 epochs, requiring between 6 to 9 hours each. Networks were trained with a batch size of 20 and a learning rate of 1e-4. Early stopping was used by saving network weights at validation loss minimums. Training was performed on a standalone workstation equipped with an NVidia Titan RTX GPU (NVidia, Santa Clara, CA).

The output of each segmentation CNN is a set of probabilities, one for each input voxel. Once the CNN was trained, a decision threshold was found to convert floating point probability maps to binary segmentations. The decision threshold is selected to maximize the fit of the predicted segmentation, and in this

![Diagram of Convolutional Neural Network (CNN) architecture used for segmentation. Volumetric patches are taken from either single or multicontrast MR images. The patches are normalized and passed through a successive 3D convolutional encoder (gold) and deconvolutional decoder (blue). Skip connections (green) take feature maps from the encoder and concatenate them with decoder feature maps before deconvolution operations. All convolution/deconvolution operations are performed with 3 x 3 x 3 kernels with a stride of 1 x 1 x 1.](image)
case, to maximize the Dice coefficient between the predicted results and the label. The threshold is calculated by running the training data through the trained network and generating precision/recall curves. The threshold at the intersection of precision and recall gives the highest agreement between the predicted segmentation and label images. This threshold gives the highest true-positive and lowest false-positive performance, maximizing the quality of the segmentation.

The images in our data set contain a single primary soft tissue sarcoma located in the hind leg, and the segmentation ground truth of these tumors is continuous. Because the CNN processes the image volume in patches without spatial references, the resulting segmentations are not guaranteed to have these properties, and the small structures outside the tumor are occasionally misclassified. To address this, after recomposing the image volume from processed patches, a postprocessing step was implemented, which rejected all but the largest continuous region in the predicted segmentation.

To provide uniform treatment of mice for subsequent radiomics analysis, the top-performing segmentation network was selected and trained again, this time using 5-fold cross validation with 10% validation and 20% testing splits (Table 1). This was necessary owing to the overlap in training data and valid radiomics sets. Each trained network would be responsible for processing only its test set, allowing these networks to cover the entirety of the radiomics data. Data that did not contribute to radiomics analysis were used for only training. Scans that contributed to radiomics analysis but did not have an associated hand segmentation were processed via an ensemble of the 5 networks via majority voting.

Segmentation Quality Metrics. We have evaluated the performance of our CNN segmentations using several metrics, the simplest of which is binary accuracy. Binary accuracy is given as the percent of voxels correctly classified by the network. The predicted probabilities are thresholded by 0.5 before calculating this metric.

We also included the Dice score, which is a standard evaluation metric in the medical imaging and computer vision communities. In addition to being used as an image quality metric, Dice was used as a loss function for half of our networks. We have also calculated the volume overlap error (VOE), also called the Jaccard similarity score, given by the intersection over the union of the ground truth ($T$) and prediction ($P$):

$$VOE = \frac{|P \cap T|}{|P \cup T|}$$

Radiomics

Using the CNN-produced segmentation maps, we performed radiomics analysis for multiple regions (Figure 1B): (A) tumor; (B) peritumoral area obtained by morphological dilation that spans 3 mm outside of the tumor; and (C) tumor combined with the peritumoral area. Each of these regions serve as masks in which to compute radiomics features. Dilation was performed using the scikit-image Python package (24). Before computing radiomics features, the MR images were normalized to account for variations in intensities, which can substantially impact feature extraction and classification. Normalization was performed by zero-centering image data and scaling values to have unit standard deviation $\sigma$. Intensity outliers (values outside $\pm 3\sigma$) were excluded from calculations as described in a study (25). Radiomics features were calculated using the PyRadiomics package (26). From each MR contrast image, 107 radiomics features were calculated, creating a high-dimensional feature space. The multicontrast data were appended to create a single feature vector for each data set (321 features). All analysis was performed using multicontrast data. The extracted features were computed from normalized images only and did not include those calculated from derivative or filtered images (e.g., Laplacian of Gaussian, wavelet, etc.).

Our first analysis focused on determining which radiomics features are the most affected by RT. Radiomics features calculated from pre- and post-RT sets were compared to find the features that changed in a statistically significant manner, as determined by paired t tests with multiple t test corrections.

Our second analysis aimed to determine if differences in radiomics features showed promise for predicting which individuals on study would develop a recurrence of the primary tumor following surgical resection. Features selection using minimum redundancy–maximum relevance (mRMR) (27) was performed on the radiomics features to rank features in an effort to reduce redundancy and increase relevance based on

| Table 2. Comparison of the Networks Trained for Sarcoma Segmentation |

<table>
<thead>
<tr>
<th>Cost</th>
<th>Data</th>
<th>Network</th>
<th>Threshold</th>
<th>Precision</th>
<th>Recall</th>
<th>AUC</th>
<th>Dice</th>
<th>VOE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dice</td>
<td>Multi-contrast</td>
<td>No skip</td>
<td>0.900</td>
<td>0.833</td>
<td>0.820</td>
<td>0.957</td>
<td>0.827</td>
<td>0.994</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Skip</td>
<td>0.995</td>
<td>0.891</td>
<td>0.826</td>
<td>0.979</td>
<td>0.857</td>
<td>0.995</td>
</tr>
<tr>
<td></td>
<td>T2 only</td>
<td>No skip</td>
<td>0.900</td>
<td>0.849</td>
<td>0.787</td>
<td>0.950</td>
<td>0.817</td>
<td>0.994</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Skip</td>
<td>0.998</td>
<td>0.906</td>
<td>0.776</td>
<td>0.977</td>
<td>0.836</td>
<td>0.994</td>
</tr>
<tr>
<td>Cross Entropy</td>
<td>Multi-contrast</td>
<td>No skip</td>
<td>0.656</td>
<td>0.814</td>
<td>0.858</td>
<td>0.996</td>
<td>0.835</td>
<td>0.994</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Skip</td>
<td>0.540</td>
<td>0.869</td>
<td>0.856</td>
<td>0.998</td>
<td>0.863</td>
<td>0.995</td>
</tr>
<tr>
<td></td>
<td>T2 only</td>
<td>No skip</td>
<td>0.636</td>
<td>0.833</td>
<td>0.803</td>
<td>0.992</td>
<td>0.818</td>
<td>0.993</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Skip</td>
<td>0.516</td>
<td>0.873</td>
<td>0.849</td>
<td>0.997</td>
<td>0.861</td>
<td>0.995</td>
</tr>
</tbody>
</table>

Values have been calculated from the test set which contains 10% of sets (ie, 7) image sets. In this data set, the network trained with cross entropy loss, skip connections, and on multicontrast images performed best according to 4 of the 5 metrics used in the current study.
recurrence. To visualize differences between groups of mice with and without local recurrence, correlation maps were computed using the 200 most relevant radiomics features selected via mRMR. For prediction of primary tumor recurrence, the top 10 features were used. The optimal number of features for prediction was found via a search that used from 2 to 200 features to maximize the area under the curve (AUC) of the receiver operating characteristic (ROC) curve. Two types of classifiers, based on simple neural networks (NNs) and support vector machines (SVMs), were used to create models predicting primary tumor recurrence. The models were trained using stratified K-fold validation. Because the quantity of data was not large by machine learning standards (data from 42 mice), each model was trained 5 times de novo. The data were split into training (75%) and validation (25%) sets, which were cycled for each training and selected to contain similar numbers of recurrences and nonrecurrences. The performance of the prediction models based on NNs or SVMs was assessed using the AUC of the ROC curves. Training and validation were performed using radiomics vectors from each of the 3 regions, that is, tumor, peritumoral area, and both combined, as illustrated by Figure 1B.

### RESULTS

#### CNN Segmentation

The measure of agreement between the label and predicted segmentation for each network trained is given with metrics of precision, recall, Dice, and VOE. These values were calculated on the test set using all 8 networks trained (Table 2). Ideal metrics would have precision and recall be close to 1 and about equal. The closeness of these 2 values represents the quality of the decision threshold that was used for these data sets. Dice and VOE scores are shown to be significantly higher for the network with skip connections. In addition, the networks trained using multicontrast data performed better than those trained using only T2-weighted images. Differences in loss function between networks was shown to be small, with cross entropy being slightly favored. The results in the table were calculated after postprocessing image volumes to remove all but the largest continuous segmentation region. This step improved segmentation performance, increasing average AUC for ROC by 0.7%, Dice by 2.2%, and VOE by 0.8%.

The best performing network, multicontrast with skip connections trained with cross-entropy loss, was retrained using 5-fold cross validation. The results are given in Table 3. CNN

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Precision</th>
<th>Recall</th>
<th>AUC</th>
<th>Dice</th>
<th>VOE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4220 ± 0.0680</td>
<td>0.8365 ± 0.0414</td>
<td>0.8497 ± 0.0260</td>
<td>0.9972 ± 0.0014</td>
<td>0.8422 ± 0.0187</td>
<td>0.9933 ± 0.0009</td>
</tr>
</tbody>
</table>

**Figure 3.** Results of CNN segmentation comparing the ground truth (label, red) with the model predictions (green) for the k-fold networks trained on cross entropy loss with skip connections and multicontrast data. Each row shows a single slice taken from separate tumor in the test set. The T2-weighted image is given for reference. The difference column shows errors in the CNN segmentation relative to the label: red for false negatives and green for false positives.
segmentations from 4 test volumes as computed with the k-fold networks are shown in Figure 3. The ground truth and predicted segmentations largely agree, with the greatest disagreement occurring on tumor edges and small extrusions. Visually, the CNN output and the original segmentation labels are generally well-matching, with only minor discrepancies. The average time required to process a single scan was 0.53 seconds.

Radiomics
Our first radiomics analysis sought to identify tumor features that change significantly when comparing images acquired before and 1 week post RT. Figure 4A displays the most statistically significant radiomics features (both shape and texture related) between pre- and post-RT sets. The gray-level features show that after RT, tumor images often acquire a more heterogeneous texture (e.g., T2 gray-level run

Figure 4. Change in radiomic features after radiation therapy (RT). Statistically significant differences in radiomics features before and after RT as calculated from the segmented tumors (A). Spatial maps of one of the intensity-based radiomics features, the gray-level dependence non-uniformity (GLDM) (B). Differences in feature intensity are clearly visible between the 2 time points. This image also shows changes in tumor size and shape 1 week after RT.

Figure 5. Comparison of radiomic features from groups which did and did not experience primary tumor recurrence. Correlation maps of the 200 most significant of radiomics features for tumor non-recurrence (A) and recurrence (B). These data were calculated from the peritumoral area in pre-RT images. The 2 correlation maps show clear differences between features based on tumors which will and will not recur.
length matrix (GLRLM)) and that the total tumor volume (i.e., shape voxel volume) typically increases 1 week post RT. A spatial map of the gray-level dependence matrix (GLDM) for dependence nonuniformity is shown in Figure 4B and illustrates both the changes in gray values and shape of a tumor. In total, 76 radiomics features were found to be significantly different with RT, including 11 shape features, 23 T1-weighted texture features, 19 T1-weighted postcontrast texture features, and 23 T2-weighted texture features. This suggests that changes induced by radiation may not be limited to tumor size and shape, but also tissue properties provided in images with multiple MRI contrasts.

**Figure 6.** A comparison of some of the most relevant radiomic features for predicting tumor recurrence as determined by minimum-redundancy-maximum-relevance (mRMR). Features are shown for data collected pre-RT and one-week post RT. These features were calculated from the peritumoral area.

**Figure 7.** Recurrence classification from features calculated from pre-RT images receiver operating curve (ROC) curves for prediction accuracy using neural networks and support vector machines (SVMs) for 3 regions examined. The best predictive power is found from features in the peritumoral area (neural network [NN] AUC: 0.78), followed by the tumor and tumor combined with the peritumoral area (NN AUC: 0.68 vs 0.96). The NN outperforms the SVM for all regions.
The second radiomics analysis we performed examined if radiomics features can be used to predict primary tumor recurrence in animals on study. It is important to note that all imaging was performed on the primary lesion before resection, and local recurrence would not be detectable until weeks after the second MRI date. The correlation maps of the 200 most significant radiomics features (peritumoral area, before RT) for tumors that would and would not recur are shown in Figure 5. Several regions within these maps demonstrated clear differences between features in the MRI data corresponding to animals that would eventually experience local recurrence after primary lesion excision, compared to animals in whom no recurrence was observed. A sample of radiomics features used for comparing animals in whom recurrence is observed and those who achieved successful local control is given in the plots of Figure 6. These features were calculated from the peritumoral area for both before and after RT and suggest that there may be measurable differences in radiomics features before and after RT that correlate with the potential for local recurrence (eg, T2 GLDM dependence entropy). A similar analysis was performed for the other 2 regions (tumor and tumor plus peritumoral area); however, the most visually stunning results came from the peritumoral area.

The features found most relevant for differentiating tumors that would eventually recur and tumors that were locally controlled came from all MR contrasts. Of the 10 radiomics features used for this purpose, 7 came gray-level features. The remaining 3 features were derived from the shape of the tumor segmentation.

The prediction model performance is illustrated in Figures 7 (pre-RT data) and 8 (post-RT data) using plots of ROC curves for NN and SVM classifiers. The best predictive power is found from features in the peritumoral area identified using an NN and based on post-RT data (AUC: 0.79, Figure 8). These were followed closely by the performance of an NN classifier trained on the same area in the pre-RT data (AUC: 0.78). In all cases, the NN classifier outperformed the SVM.

**DISCUSSION AND CONCLUSIONS**

Our results show that CNN-based segmentations with supervised learning using either T2-weighted or multicontrast MRI images are viable methods for automatic tumor volumetric measurements. Our segmentation performance (Table 2) was better in the configuration using skip connections (ie, a U-net configuration) similar to what has been reported in other studies (28). The best overall segmentation performance was achieved using a cross entropy loss, with Dice scores of 0.861 for T2-weighted images versus 0.863 for multicontrast data. When trained in K-fold cross validation, the performance drops slightly from the initial training (Table 3, mean Dice: 0.8422). This may be attributed to variation in test sets between trainings.

Automated tumor segmentation of the images before and after RT showed that tumor volumes increase in the week between the 2 imaging time points. Because a single dose of RT alone is unlikely to inhibit growth of a palpable primary lesion, these trends were expected. Gray-level intensity radiomics features indicated that tumor images acquire a more heterogeneous texture 1 week post RT. Although the administered RT was not expected to inhibit tumor growth, high-dose exposure is likely to damage tumors, causing tissue-level changes such as...
inflammation, edema, and necrosis. These changes alter tumor signal patterns in each of the MR contrasts, contributing to heterogeneity of the tumor radiomics features that differentiate the pre- and post-RT MRI data. This suggests that radiomics with multicontrast MRI may be useful in detecting and monitoring the effects of high-dose radiation in solid tumors.

In addition, our data suggest that radiomics features could aid in determining the likelihood of primary tumor recurrence. In our study, pre-RT data that include tumor and surrounding tissues were the most effective at identifying individuals likely to recur locally (Figure 7). Recently, peritumoral radiomics was also used to predict distant metastases in locally advanced non–small cell lung cancer (14). Although T2-weighted data alone are sufficient for tumor volume segmentation, only lagging slightly behind multimodal segmentation in performance, there are advantages in using multicontrast MRI data in radiomics analysis. This is particularly true when considering that both T1-weighted and T2-weighted scans are frequently included as standard protocols in clinical cancer imaging. In our study, the most relevant features for assessing changes over time with RT, as well as the prediction of local recurrence, were identified when including multiple MRI contrasts (Figures 4–7).

One limitation of this study is its dependence on successful and complete surgical resection of the primary tumor. A major advantage of the described mouse model is that the tumors grow and spread much like human soft tissue sarcomas, with little control by the investigators aside from site of initiation. Surgical resection of the primary lesion, just as is the case in human patients, is not always complete. It is possible that some recurrences were secondary to microscopic positive margins or regional nodal disease. Thus, surgical margins remain an important consideration when discussing local recurrence of the primary lesion. Although encouraged by our radiomics findings, we acknowledge that there are additional factors potentially confounding that cannot be evaluated by MRI data alone.

In future work, we aim to improve our prediction models by adding other biomarkers related to immune response, and we will also attempt to predict distant metastases to the lungs. More importantly, our imaging analysis pipeline will also be applied in studies adding immunotherapy to RT as part of our coclinical trial of sarcoma (11).

In conclusion, we have created and tested an image processing pipeline for high–throughput, reduced-bias segmentation of multiparametric tumor MRI data that serves the preclinical arm of our coclinical trial. Furthermore, we have implemented the architecture for radiomics analysis of tumor images, to better our understanding of preclinical imaging and the insights it provides when studying new therapeutic strategies for cancer.

ACKNOWLEDGMENTS

All work was performed at the Duke Center for In Vivo Microscopy supported by the NIH National Cancer Institute (R01 CA196667, U24 CA220245). Additional support was also provided by an NIH training grant from the National Institute of Biomedical Imaging and Bioengineering (T32 EBO01040) and NIH National Cancer Institute (R35 CA197616).

REFERENCES


28. Kayalibay B, Jensen G, van der Smagt P. CNN-based segmentation of medical imaging data. arXiv:170103056 [cs] [Internet]. 2017 [cited 2018 Jul 27]; Available from: http://arxiv.org/abs/1701.03056. Here the best predictive power is found in the peritumoral area (NN AUC: 0.79), which is the strongest performing predictor overall. The peritumoral area shows the next strongest performance (NN AUC: 0.70 and SVM AUC: 0.67).
Microstructure Modeling of High b-Value Diffusion-Weighted Images in Glioblastoma

Yuan Li, Michelle Kim, Theodore S. Lawrence, Hemant Parmar, and Yue Cao

 Departments of Radiation Oncology, Radiology, and Biomedical Engineering, University of Michigan, Ann Arbor, MI

Corresponding Author:
Yue Cao, PhD
519 W. William Street, Ann Arbor, MI 48103;
E-mail: yuecao@umich.edu

Key Words: Diffusion-weighted imaging, high-order diffusion MR methods, quantitation, glioblastoma

Abbreviations: apparent diffusion coefficient [ADC], apparent restriction size of intracellular water [ARS], cell radius [R], cerebrospinal fluid [CSF], diffusion sensitive signals of intracellular water [Sin], diffusion sensitive signals of extracellular water [Sout], diffusion-weighted (DW), extracellular diffusion coefficient [Dex], fractional volume of intracellular water [Vin], fractional volume of water with the small diffusion coefficient [Vs], Gaussian Phase Distribution [GPD], glioblastoma (GBM), intracellular diffusion coefficient [Dim], microstructure model (MSM), mean squared errors (MSE), pulse gradient spin echo [PGSE], signal-to-noise ratio (SNR), tumor volume (TV), volumes of interest (VOIs), white matter (WM)

INTRODUCTION
Glioblastoma (GBM) is the most common and aggressive primary brain tumor in adults and has poor prognosis with a median survival of ~14 months despite multimodality therapy with surgery, concurrent chemoradiation therapy, and adjuvant chemotherapy (1, 2). Postcontrast T1-weighted and fluid-attenuated inversion recovery (FLAIR) T2-weighted magnetic resonance imaging (MRI) images are the primary images to guide treatment and assess tumor progression or therapy response (3). However, the postcontrast T1-weighted MRI identifies blood–brain barrier disruption, which is affected by tumor growth but also by radiation, antiangiogenesis drugs, and chemotherapy. Abnormality on T2 FLAIR images is influenced by T2 changes of tumor cells as well as edema that always coexists within GBM. Limitations of conventional MRI in clinical management of GBM have been recognized and motivate investigations of physiological and metabolic MRI.

Diffusion-weighted (DW) imaging is a technique to measure water molecule mobility in the tissue microscopic environment, and is sensitive to cell density and size, cell membrane permeability, and extracellular space tortuosity. Apparent diffusion coefficient (ADC) quantified from conventional DW images with b-values between 0 and 1000 s/mm² using a mono-exponential decay is a commonly reported parameter in literature. The correlation between high cellularity and low ADC in tumor animal models and human cancers motivates investigations of roles of ADC in clinical GBM (4-7). One limitation of DW imaging is that coexistence of edema in clinical GBM results in elevated ADC compared with normal white matter (WM) and gray matter (GM). To overcome this challenge, high b-value DW images and high-order diffusion MR methods were applied to 30 patients with newly diagnosed GBM. Diffusion-weighted (DW) images acquired on a 3 T scanner with b-values from 0 to 2500 s/mm² were fitted in volumes of interest (VOIs) of solid tumor to obtain the apparent restriction size of intracellular water (ARS), the fractional volume of intracellular water (Vin), and extracellular (Dex) water diffusivity. The parameters in solid tumor were compared with those of other tissue types by Students’ t-test. For comparison, DW images were fitted by conventional mono-exponential and bi-exponential models. ARS, Dex, and Vin from the MSM in tumor VOIs were significantly greater than those in WM, GM, and edema (P values of .01-.001). ARS values in solid tumors (from 21.6 to 34.5 um) had absolutely no overlap with those in all other tissue types (from 0.9 to 3.5 um). Vin values showed a descending order from solid tumor (from 0.32 to 0.52) to WM, GM, and edema (from 0.05 to 0.25), consisting with the descending cellularity in these tissue types. The parameters from mono-exponential and bi-exponential models could not significantly differentiate solid tumor from all other tissue types, particularly from edema. Further development and histopathological validation of the MSM will warrant its role in clinical management of GBM.

Apparent diffusion coefficient has limits to differentiate solid tumor from normal tissue or edema in glioblastoma (GBM). This study investigated a microstructure model (MSM) in GBM using a clinically available diffusion imaging technique. The MSM was modified to integrate with bi-polar diffusion gradient waveforms, and applied to 30 patients with newly diagnosed GBM. Diffusion-weighted (DW) images acquired on a 3 T scanner with b-values from 0 to 2500 s/mm² were fitted in volumes of interest (VOIs) of solid tumor to obtain the apparent restriction size of intracellular water (ARS), the fractional volume of intracellular water (Vin), and extracellular (Dex) water diffusivity. The parameters in solid tumor were compared with those of other tissue types by Students’ t-test. For comparison, DW images were fitted by conventional mono-exponential and bi-exponential models. ARS, Dex, and Vin from the MSM in tumor VOIs were significantly greater than those in WM, GM, and edema (P values of .01-.001). ARS values in solid tumors (from 21.6 to 34.5 um) had absolutely no overlap with those in all other tissue types (from 0.9 to 3.5 um). Vin values showed a descending order from solid tumor (from 0.32 to 0.52) to WM, GM, and edema (from 0.05 to 0.25), consisting with the descending cellularity in these tissue types. The parameters from mono-exponential and bi-exponential models could not significantly differentiate solid tumor from all other tissue types, particularly from edema. Further development and histopathological validation of the MSM will warrant its role in clinical management of GBM.

© 2020 The Authors. Published by Grapho Publications, LLC This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
ISSN 2379-1381 https://doi.org/10.18383/j.tom.2020.00018
order diffusion models have been explored in clinical gliomas to differentiate tumor grade and assess therapy response (8–16).

Among the high-order diffusion models, a few models attempt to quantify microstructures in tumors (17, 18). For example, a model, called VERDICT, has been proposed to quantify microstructural properties of colorectal cancer cell lines (19). This model considers cell size, vascular volume fraction and associated pseudodiffusivity, and intracellular and extracellular fractional volumes and diffusivities. Owing to complicity of the VERDICT model, prior knowledge of intracellular and extracellular diffusion coefficients is used to fit in vivo DW images in 2 xenograft animal models. Another model, called temporal diffusion spectroscopy, uses oscillating diffusion gradients to probe cellular structures that restrict intracellular water diffusion by assuming that intracellular water is restricted in impermeable cells (20, 21). Compared with VERDICT, this model simplifies the extracellular water diffusion to a single free diffusion term, and it fits 4 free parameters, including intracellular and extracellular water diffusivities.

Recent studies show that hypercellular tumor volumes (TVs) that can be detected on the DW images with $b = 3000\ mm^2/s$ in GBM have a prognostic value (22, 23). This technique used a widely available imaging technique and could be easily translated into a clinical trial. A phase II clinical trial targets this hypercellular TV with intensified radiation doses (24).

In this study, we modified the model described in Jiang et al.’s study (25, 26) and applied it to clinical DW images in the patients with GBM. In our clinic, bipolar pulse diffusion gradient waveforms and a high parallel imaging factor were used to reduce eddy current and geometric distortion in the clinical DW images, respectively. We applied the modified model to the DW images with high $b$-values to investigate whether there were any significant differences in the quantified microstructure parameters between the hypercellular tumors and normal tissue and edema in the patients with GBM. Similar comparisons were made for the conventional ADC and the parameters quantified from the bi-exponential model. This study was the first step to test the possibility of the application of the model quantifying the microstructure parameters using a widely available diffusion technique on the clinical scanner for GBM.

**MATERIALS AND METHODS**

**Microstructure Model with Bipolar Diffusion Gradients**

We assume a DW signal in tissue that can be considered as a sum of water signals from intracellular and extracellular compartments (21):

$$S = S_0[V_{in}S_{ia} + (1 - V_{in})S_{ex}]$$  \[1\]

where $S_0$ is the total magnetization from both water compartments, $V_{in}$ is the fraction volume of intracellular water, and $S_{in}$ and $S_{ex}$ are respective diffusion sensitive signals of intracellular and extracellular water. In the previous works of restricted intracellular diffusion (19, 27, 28), a cell has been modeled as an impermeable sphere, which completely restricted diffusion of intracellular water molecules within the spherical space. The analytical formulae of the restricted diffusion signals have been derived for the conventional monopolar diffusion pulse gradient spin echo (PGSE) and oscillating gradient spin echo sequences using the Gaussian phase distribution (GPD) approximation (20, 29–31). It has shown that the GPD approximation of restricted diffusion for the conventional PGSE sequence has sufficient accuracy for most experimental conditions and for sphere and parallel-plane geometry assumptions (32). Therefore, we adopt this formulation to express the restricted diffusion signal of the intracellular water as:

$$S_{in} = \frac{γ^2}{2} \sum B_n \int_0^{2\pi} dt_1 \int_0^{2\pi} g(t_1) g(t_2) \exp\left(-D_n λ_1 |t_1 - t_2|\right) dt_2$$  \[2\]

where $γ$ is the gyromagnetic ratio of proton spin, $g(t)$ is the gradient waveform, $D_n$ is the intracellular diffusion coefficient, and $λ_n$ and $B_n$ are structure-dependent parameters. The analytical formulae of $λ_n$ and $B_n$ for the spherical geometry [provided by a previous work (30)] depend upon the radius ($R$) of the sphere or cylinder and the $n$th root of a Bessel function of the first kind. The integral in equation [2] depends upon the specific gradient waveform $g(t)$ used in the diffusion pulse sequence. On the clinical scanner, the most commonly used gradient waveform is the conventional monopolar PGSE, and the oscillating diffusion gradient waveforms are not available. However, large eddy currents generated in the monopolar diffusion PGSE can produce artifacts on DW images.

To minimize eddy current–caused artifacts and improve quality of DW images, bipolar diffusion gradient pulse sequences have been introduced on clinical scanners (33). There are a few variations in bipolar diffusion gradient waveforms that have been implemented on the clinical scanners by different vendors. Three common bipolar gradient waveforms, $g(t)$, are illustrated in Figure 1. The first one is introduced by Fordham et al. (34), in which bipolar gradient pulses simply replace the monopolar gradients before and after the 180° radiofrequency (RF) pulse. The second one contains 4 diffusion gradient pulses that are placed before, between, and after 2 180° RF pulses. The 4 diffusion gradient pulse durations and time intervals can be tuned to minimize eddy current effects on DW images, resulting in asymmetric waveforms (Figure 1B). We derived $S_{in}$ in equation [2] for 3 bipolar diffusion gradient waveforms shown in Figure 1 and provided them in the online supplemental Appendix.

Finally, the extracellular diffusion signal is formulated as follows (20):

$$S_{ex} = \exp(-bD_{ex})$$  \[3\]

where $D_{ex}$ is the extracellular water diffusion, and $b$ is the $b$-value and depends upon the gradient waveform. For the 3 bipolar diffusion gradient waveforms, $b$-values are given in equations [A3], [A6], and [A9] in the online supplemental Appendix. Note that 4 free parameters ($R, D_{in}, D_{ex}$, and $V_{in}$) can be estimated by fitting the microstructure model (MSM) to DW images. Here, $R$ can be considered to be an apparent restriction size (ARS) of intracellular water. Also, whether $D_{in}$ was sensitive to the cost function in fitting was investigated.

**Bi-Exponential Model**

The bi-exponential diffusion model, considered 2 free diffusion components, has been investigated to differentiate tumor from normal tissue and assess tumor response to therapy (35). To
The small diffusion coefficient $b = 1000 \text{s/mm}^2$ by a mono-exponential diffusion function as the bi-exponential model.

Conventional ADC is usually

$$S = S_1 \exp(-bD_1) + S_2 \exp(-bD_2)$$

where $S_1$ and $S_2$ are respective amplitudes of apparent diffusion coefficients of $D_1$ and $D_2$. The fractional volume of water with the small diffusion coefficient is given by:

$$V_s = S_1/(S_1 + S_2).$$

Again, 4 free parameters $(D_1, D_2, S_1, \text{and} \ S_2)$ are fitted from the bi-exponential model.

Conventional Mono-Exponential Model

Conventional ADC is usually fitted to DW images with $b = 0$ and $b = 1000 \text{s/mm}^2$ by a mono-exponential diffusion function as

$$S = S_0 \exp(-bADC)$$

Patients

Thirty patients (median age, 62 years; males, 19; females, 11) with histologically diagnosed new glioblastoma were included in this study that has been approved by an institutional review board. All patients had MRI scans post surgery but before chemoradiation therapy.

In Vivo MR Imaging

All scans were performed on a 3.0 T scanner (Skyra, Siemens Healthineers, Erlangen, Germany) using a 20-channel head coil. Conventional MR images, 2D T2-FLAIR images, and 3D pre- and postcontrast T1-weighted images using a MPRAGE sequence were acquired. DW images were acquired by a spin-echo echo-planar pulse sequence with diffusion weighting in 3 orthogonal directions and 11 b-values from 0 to 2500 s/mm$^2$ with an incremental step of 250 s/mm$^2$. A bipolar diffusion gradient waveform (shown in Figure 1B) was used to reduce eddy currents and improve quality of DW images. In this sequence, there were 2 180° RF pulses and 4 diffusion gradient pulses. The four diffusion gradient pulses had durations of $\delta_1 = 9.94$ milliseconds, $\delta_2 = 15.14$ milliseconds, $\delta_3 = 19.8$ milliseconds, and $\delta_4 = 5.28$ milliseconds. The times intervals between the first and second, the first and third, and the first and fourth gradient pulses were $\Delta_1 = 20.84$ milliseconds, $\Delta_2 = 36.64$ milliseconds, and $\Delta_3 = 67.34$ milliseconds, respectively. Other acquisition parameters included the parallel imaging factor of 4 (GRAPPA) (to reduce echo spacing and hence geometric distortion), TE/TR = 93 / 9300 milliseconds, bandwidth of 1040 Hz/pixel, voxel size of $\sim 1.3 \times 1.3 \times 5.2$ mm, 30 slices to cover the whole brain, 1 average, and total scan time of 4.50 minutes. It has been shown that diffusion anisotropy is lost or reduced dramatically in T2 abnormalities of GBM owing to tumor infiltration and edema (36). To test the loss or reduction of anisotropic diffusion in GBM, DW images were also acquired in 30 directions with $b = 1000 \text{s/mm}^2$ and 3 averages. To determine the hypercellular tumor in GBM, a diffusion image volume was acquired at $b = 3000 \text{s/mm}^2$ and 4 averages.

Definition of Volumes of Interest

First, we investigated whether microstructure and diffusion parameters in solid components of GBM, estimated by this model, were significantly different from ones in edema regions, normal gray matter (GM), and normal WM. Owing to anticipated low SNRs in DW images, we performed this test in several volumes of interest (VOIs). Previous studies of GBM using advanced imaging have shown that solid tumor components can be beyond the contrast-enhanced gross TVs (22, 37). Also, the T2 abnormality volume can consist of tumor, edema, and a mixture of the 2. However, at high b-values, water signals of edema are attenuated much faster than the hypercellular tumor. Based upon this hypothesis, in previous studies, a TV was created by combining automated thresholding on the DW images with $b = 3000 \text{s/mm}^2$, and then these were viewed and edited by a neuroradiologist with more than 10 years of clinical experience (22). The initial TV was created using a threshold of the mean intensity plus 2 standard deviations calculated from a volume of interest in the normal-appearing tissue most contralateral to the T2-abnormality in vivo (22). Therefore, we used this TV to characterize microstructure and diffusion parameters by the MSM in this study. Note that this TV is different from the contrast-enhanced TV (Figure 2).
To compare the behavior of this model in edema to solid tumor in GBM, an edema volume was created within the T2-abnormality but had at least 1 cm away from both the TV used in this study and the contrast-enhanced gross TV. In the cases with tumor recurrence, the edema volume was checked and ensured to have no spatial overlap with the recurrent TV. Also, the VOIs of 2 large WM fiber bundles were drawn: one in the frontal lobe and another in the Genu of corpus callosum. To compare to GM, cortex regions in the frontal and parietal lobes were segmented using the fuzzy c-means algorithm on DW images with b = 0 (T2-weighted images) and ADC maps. To avoid influence of cerebrospinal fluid (CSF), a deep GM structure, the head of caudate nucleus, was carefully chosen. As a total, 6 VOIs were created (Figure 2).

Before fitting the MSM, we investigated fractional anisotropy (FA) of diffusion in the defined hypercellular TV to determine whether we could fit the MSM using mean diffusivities in the TV. The averaged FA was 0.15 ± 0.05 in the TVs, which is consistent with previous reports (36), and 0.41 ± 0.07 in the frontal WM. Therefore, it is reasonable to fit the MSM to the mean diffusion signals in the TV using a sphere assumption and omitting anisotropic diffusion. In the WM, the cylinder-shape assumption was used, while the sphere shape was used in other tissue types.

Before the VOI creation, postcontrast T1-weighted images were reformatted into the axial plane with spatial resolution of 1 × 1 × 3 mm. All other images acquired within the same session were reformatted to the postcontrast T1-weighted images using coordinates in the DICOM header.

**Computation of Diffusion Models**

Both the MSM and bi-exponential model were implemented using Matlab. The MSM was fitted to the DW images with 11 b-values using a *Simplex* algorithm in Matlab. We investigated the sensitivity of the objective function to the parameters of $R$, intracellular diffusion coefficient $D_{in}$, intracellular volume fraction $V_{in}$, and extracellular $D_{ex}$. If any parameter was not sensitive to the objective function, we would use a fixed value, which would reduce the number of the free parameters and improve the stability of fitting. *Simplex* was initiated randomly in the ranges of the parameters based on prior knowledge of the physiological parameters given in Table 1. Fitting was run multiple times, and the results were accepted with the minimum mean squared errors (MSE). Similarly, the bi-exponential model was fitted to the same DW images. ADC maps were calculated from DW images with $b = 0$ and 1000 s/mm$^2$ using in-house Functional Image Analysis Tools (*imFIAT*).

**Statistical Analysis**

To evaluate whether the parameters fitted from the MSM can differentiate tumor from other tissue types, Students’ *t* test was used and a *P*-value of 0.05 was considered significant. Similar analysis was applied to the parameters obtained from the bi-exponential model and ADC.

**RESULTS**

**Parameter Characteristics from the MSM**

When investigating sensitivity of the objective function to the parameter variation, we found that $D_{in}$ had little sensitivity. To test the influence of $D_{in}$ variations on other parameters, we varied $D_{in}$ from 0.1 to 1.0 $\mu$m$^2$/ms. We found that the fitted $R$ in the TVs had no more than 1.5% differences, and $D_{ex}$ and $V_{in}$ did not show significant differences (Figure 3). Therefore, we fixed $D_{in}$ at 0.1 $\mu$m$^2$/ms and varied other 3 parameters in fitting the MSM.

The MSM was fitted to the DW data well in all VOIs, as example curves from TV, WM, GM and edema are shown in.

**Table 1. Ranges of Initial Values of Three Parameters for Microstructure Model Fitting**

<table>
<thead>
<tr>
<th>Tissue Type</th>
<th>$R$ ($\mu$m)</th>
<th>$D_{ex}$ ($\mu$m$^2$/ms$^{-1}$)</th>
<th>$V_{in}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor</td>
<td>5–25</td>
<td>1–3</td>
<td>0–1</td>
</tr>
<tr>
<td>Normal Tissue</td>
<td>0.5–1.5</td>
<td>1–3</td>
<td>0–1</td>
</tr>
<tr>
<td>Edema</td>
<td>0.5–1.5</td>
<td>1–3</td>
<td>0–1</td>
</tr>
</tbody>
</table>
than normal GM and WM (the existence of a small amount of tumor cells in the edema VOI. Figure 4. Note that the DW signal in the TV was attenuated slower than normal GM and WM at b-values >1000 s/mm$^2$, while the DW signal in edema was attenuated rapidly at low b-values, indicating that a large portion of extracellular water has relative free diffusion.

Characteristics of 3 fitted parameters in the TV, normal GM, normal WM, and edema are summarized in Table 2 and Figure 5. Three parameters, $R$, $D_{ex}$, and $V_{in}$, in the TV were significantly different from normal GM, normal WM, and edema. Specifically, the mean $R$ in the tumor was $28.1 \pm 0.48$ $\mu$m and significantly greater than those in other tissue types (range, 1.1–2.3 $\mu$m, $P < .001$). In the latter group, edema had a significant greater $R$ than normal GM and WM ($P < .001$), which could be because of the existence of a small amount of tumor cells in the edema VOI.

The mean fractional volume of intracellular water, $V_{in}$ in the tumor was 0.42 and significantly increased compared with all other tissue types ($P < .001$). Among normal tissue and edema, as anticipated, $V_{in}$ had the highest value in the large WM fiber bundles (0.21), intermediate value in the GM regions (0.16–0.13), and the lowest value in edema (0.10) that is consistent with the large amount of extracellular water. Most interestingly, the values of $R$ and $V_{in}$ in the tumor had absolutely no overlap with the ones in other tissue types, suggesting potential high sensitivity and specificity of $R$ and $V_{in}$ for differentiating tumor from edema and normal tissue.

The mean $D_{ex}$ in the tumor was $2.03 \mu m^2/\text{ms}$ and significantly greater than that in the normal tissue and in edema ($P < .01$), which is contributed possibly from edema, micronecrosis, and perfusion mixed in the voxels of GBM. Normal WM and deep GM had $D_{ex}$ values of $1.15–1.31 \mu m^2/\text{ms}$, which is consistent with other reports (38). Edema had $D_{ex}$ of $1.52 \mu m^2/\text{ms}$, largely attributing to the great fractional volume of extracellular water. Cortex had $D_{ex}$ of $1.62 \mu m^2/\text{ms}$, possibly resulting from partial volume average effects with CSF.

Parameter Characteristics from the Bi-Exponential Model

Characteristics of the 3 fitted parameters by the bi-exponential model are summarized in Table 3 and Figure 6. None of the 3 parameters in the tumor significantly and completely differed from the values in all other tissue types. $D_1$ (the large diffusion coefficient) in the tumor was $2.02 \pm 0.07 \mu m^2/\text{ms}$, was not significantly different from edema (1.89 $\pm 0.06 \mu m^2/\text{ms}$), but was significantly greater than normal WM and deep GM ($P < .05$). The cortex had the significantly elevated $D_1$ compared to tumor and other normal tissue ($P < .01$). $D_2$ in the tumor was $0.34 \pm 0.01 \mu m^2/\text{ms}$, was not significantly different from edema (0.33 $\pm 0.04 \mu m^2/\text{ms}$), deep GM, and cortex (0.36 $\pm 0.02 \mu m^2/\text{ms}$), but was significantly greater than normal WM (0.16–0.21 $\mu m^2/\text{ms}$, $P < .01$). $V_s$ in the tumor was 0.42, significantly greater than edema (0.29, $P < .05$) and genu (0.33, $P < .05$), and not significantly different from frontal WM and deep GM. Cortex had the highest values in $D_1$, $D_2$, and $V_s$ than tumor and normal WM and GM, possibly owing to the partial volume average with CSF and suggesting that the bi-exponential model is strongly influenced by fluid components.

To test whether combining all 3 parameters ($D_1$, $D_2$, and $V_s$) could differentiate tumor from all other tissue types, binary multivariate logistic regression was applied to the data. Backward rejection with $P > .1$ was used to eliminate the parameters in the logistic regression models. In the first model including the 3 parameters (Table 4), $D_1$ was significant ($P < .05$), but $V_s$ was not significant, and $D_2$ was marginally significant. However, $D_2$ had a large negative coefficient that could offset the $D_1$ contribution in the model. In the second model where $V_s$ was rejected, $D_1$ was...
marginally significant, and $D_2$ was not significant. In the final model, $D_1$ was not significant after rejecting $D_2$.

**Conventional ADC Model**
There were no significant differences of ADC between tumor and any other nontumor tissue types ($P > .05$) (Figure 7). In general, ADC in the tumor was greater than WM and deep GM but lower than edema and cortex, consistent with other reports (7).

**DISCUSSION**
In this study, we modified the model of (20, 25, 26, 29-31) to fit the DW images acquired with a widely available bipolar pulse diffusion gradient imaging and characterize microstructure and diffusion properties of the hypercellular tumor in patients with GBM. We found that 3 parameters ($V_{in}$, $R$, and $D_{ex}$) in the tumor were substantially and significantly different from edema and normal tissue. The bi-exponential diffusion model that does not explicitly model the restricted diffusion of intracellular water could not robustly differentiate GBM from edema and normal brain tissue. ADC that ignores intravoxel heterogeneous diffusion in brain tissue and tumor failed to differentiate GBM from edema and normal tissue. The microstructure model has a great promise to aid in to conventional MRI for GBM diagnosis, image-guide therapy, and response assessment. Further validation with histopathology will warrant the role of the microstructure mode in the clinical management of GBM.

<table>
<thead>
<tr>
<th></th>
<th>$R$ (um) (Mean ± SEM)</th>
<th>$D_{ex}$ (um$^2$/ms) (Mean ± SEM)</th>
<th>$V_{in}$ (Mean ± SEM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor</td>
<td>28.1 ± 0.48</td>
<td>2.03 ± 0.07</td>
<td>0.42 ± 0.011</td>
</tr>
<tr>
<td>Frontal White Matter$^a$</td>
<td>1.20 ± 0.01</td>
<td>1.15 ± 0.01</td>
<td>0.21 ± 0.004</td>
</tr>
<tr>
<td>Genu$^a$</td>
<td>1.13 ± 0.01</td>
<td>1.31 ± 0.02</td>
<td>0.21 ± 0.005</td>
</tr>
<tr>
<td>Deep Gray Matter</td>
<td>1.19 ± 0.05</td>
<td>1.19 ± 0.04</td>
<td>0.16 ± 0.004</td>
</tr>
<tr>
<td>Cortex</td>
<td>1.16 ± 0.02</td>
<td>1.62 ± 0.05</td>
<td>0.13 ± 0.002</td>
</tr>
<tr>
<td>Edema</td>
<td>2.32 ± 0.07</td>
<td>1.52 ± 0.05</td>
<td>0.10 ± 0.007</td>
</tr>
</tbody>
</table>

$^a$R was fitted by using a cylinder shape assumption.
In the current study, the fractional volume of intracellular water, \( V_{in} \), in GBM estimated by the microstructure model was found substantially different from edema and normal tissue. The \( V_{in} \) in the TV had the largest value, which makes possible to differentiate the GBM from surrounding tissue. Also, GBM has enlarged cell size and nucleus, and increased cell density, which can increase \( V_{in} \) measured in image voxels, and has microcroeosis and edema, which can reduce \( V_{in} \) (39, 40). In the current study, the estimated \( V_{in} \) in GBM by the microstructure model was 1.75-fold greater than that reported in a previous study that included primary and metastatic cancers and used the bi-exponential diffusion model (27). Historically, the bi-exponential diffusion model often results in an underestimated \( V_{in} \), for which several possible causes have been discussed (38). The underestimation in \( V_{in} \) can be caused by the transcytoplemmal water exchange that is omitted in the model as well as low SNR in DW images (41). Also, T2 differences between intracellular and extracellular water could affect the estimated \( V_{in} \) values, which will be discussed further in the last paragraph. Diffusion gradient, b-value range, diffusion model, and subject age all have an influence on the estimated \( V_{in} \) (4). As expected, the lowest \( V_{in} \) value was found in edema, consistent with the notion of a large amount of extracellular water in the region. The \( V_{in} \) values in 2 large WM fiber bundles and GM are between the solid GBM and edema, suggesting that the MSM has the potential to differentiate GBM from surrounding tissue. In this study, we evaluated anisotropic diffusion in the hypercellular TV and found low FA (0.15 as a mean value). Therefore, we did not consider anisotropic diffusion in the MSM. In the future study, anisotropic diffusion could be considered in the microstructure model.

Our microstructure model yielded the substantial greater \( R \) in GBM than in normal GM, normal WM, and edema. \( R \) could be considered as the apparent restriction size of intracellular water and a possible biomarker to differentiate GBM from normal tissue. The \( R \) value should be considered as an average value over a distribution. Previous studies have shown the increased radius of tumor cells compared with that of normal tissue using the microstructure model or VERDICT model, and the reasonable correlation between the DW image–estimated cellularity and histopathologically determined cellularity (19, 21). A pathological

**Table 3.** Characteristics of Three Parameters Fitted by the Bi-Exponential Model

<table>
<thead>
<tr>
<th></th>
<th>( D_1/\mu m^2/ms ) (Mean ± SEM)</th>
<th>( D_2/\mu m^2/ms ) (Mean ± SEM)</th>
<th>( V_{in} ) (Mean ± SEM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor</td>
<td>2.02 ± 0.07</td>
<td>0.34 ± 0.01</td>
<td>0.42 ± 0.01</td>
</tr>
<tr>
<td>Frontal White Matter</td>
<td>1.44 ± 0.03</td>
<td>0.21 ± 0.02</td>
<td>0.38 ± 0.01</td>
</tr>
<tr>
<td>Genu</td>
<td>1.55 ± 0.04</td>
<td>0.16 ± 0.02</td>
<td>0.33 ± 0.02</td>
</tr>
<tr>
<td>Deep Gray Matter</td>
<td>1.73 ± 0.07</td>
<td>0.36 ± 0.02</td>
<td>0.46 ± 0.02</td>
</tr>
<tr>
<td>Cortex</td>
<td>2.78 ± 0.06</td>
<td>0.55 ± 0.01</td>
<td>0.51 ± 0.02</td>
</tr>
<tr>
<td>Edema</td>
<td>1.89 ± 0.06</td>
<td>0.33 ± 0.04</td>
<td>0.29 ± 0.04</td>
</tr>
</tbody>
</table>
A study in GBM shows that the radius of GBM cells can be as large as 20 μm with a mean of 10 μm and a standard deviation of 11 μm (42). The size of the apparent restriction space estimated in our model is larger than the reported cell size. Many factors can cause an overestimation in the restriction size of intracellular water in GBM. It is most likely that the highly permeable cell membrane cannot restrict intracellular water diffusion like a hard boundary. Therefore, the soft boundary increases the estimated radius so that we could call the estimated radius from our MSM as an effective radius. The low SNR in DW images could further cause an overestimate of R (43). Owing to the diffuse nature of GBM, diffused tumor cells can be found in the edema region (44, 45), which could be the cause of the slight but significant increase in the apparent restriction size in the edema region. Further studies could be carried out to investigate whether combining R and \( V_{in} \) in the image voxels with edema can provide information on the GBM cell infiltration and distribution.

In the current study, the value of \( D_{ex} \) in GBM by the microstructure model was found significantly greater than those in normal GM, normal WM and edema, which is similar to one by the bi-exponential diffusion model (Table 3), and also consistent with the values reported by Mulkern et al. (38). \( D_{ex} \) in GBM is affected by the large quantity of extracellular water owing to coexistence of edema, and to a small extent by micronecrosis and perfusion (40). To reduce the number of free parameters in the model, we did not account for the perfusion-caused pseudo diffusion in \( D_{ex} \). However, we have tested the perfusion effect on \( D_{ex} \) in the TV VOIs in 30 patients using the approach in (46). We estimated that the pseudo diffusion coefficient was \( \sim 0.15 \text{ um}^2/\text{ms} \), which was 13.5 times smaller than \( D_{ex} \) (2.03 \text{ um}^2/\text{ms}) in the tumor VOIs. Thus, in the TV VOIs defined in our study, omitting the perfusion effect did not cause a substantial overestimate in \( D_{ex} \). Note that the TV VOI used in our study is not the contrast-enhanced TV. We believe that the high \( D_{ex} \) in GBM could mainly be due to edema. The relative high values of \( D_{ex} \) in cortex and edema are likely because of the partial volume average of CSF in the cortex VOI and the large amount of extracellular water in the edema region, respectively.

As discussed in Introduction, modeling microstructure and diffusion properties in tumors has been attempted by different models (17, 18, 20, 21, 28). In general, these models fit more free

<table>
<thead>
<tr>
<th>Table 4. Multivariate Logistic Regression Using the Parameters from the Bi-Exponential Model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First Model</strong></td>
</tr>
<tr>
<td>Coefficient</td>
</tr>
<tr>
<td>P Value</td>
</tr>
<tr>
<td><strong>Second Model</strong></td>
</tr>
<tr>
<td>Coefficient</td>
</tr>
<tr>
<td>P Value</td>
</tr>
<tr>
<td><strong>Third Model</strong></td>
</tr>
<tr>
<td>Coefficient</td>
</tr>
<tr>
<td>P Value</td>
</tr>
</tbody>
</table>

Figure 6. Bar graphs of estimated parameters of \( D_1 \) (top left panel), \( D_2 \) (top right panel), and \( V_s \) (bottom left panel) in all tissue types using the bi-exponential diffusion model. * .01 < P < .05; ** .001 < P < .01; *** P < .001. Error bar is SEM.

Figure 7. Bar graph of apparent diffusion coefficients in all tissue types using the mono-exponential diffusion model. Error bar is SEM.
parameters and require long acquisition times. The signal-to-noise ratio (SNR) of DW images is also a limiting factor to produce robust fitted parameters. Furthermore, there are great time restriction, technology availability, and other practical considerations when acquiring DW images in the patients with GBM on a clinical scanner. For example, we used a bipolar pulse diffusion gradient waveform to reduce eddy current–induced artifacts and a high parallel imaging factor to improve geometries accuracy.

With these acquisition limitations, we modified and tested the model in the literature (18) and (19) using the widely available technology on a clinical scanner in the patients with GBM. For other bipolar diffusion gradient waveforms, we provided equations of the model in the online supplemental Appendix.

The MSM showed the prognostic potential and also the opportunity to better distinguish GBM from normal tissue and edema. However, the current study has a few limitations. First, our sample size was not large, with only 30 patients. This could result in overfitting. The MSM needs to be further validated in an independent large cohort of patients in future. Second, the restriction space of intracellular water is modeled as an impermeable sphere, which has been used in both VERDICT and temporal diffusion spectroscopy model (19, 21, 28). A previous simulation work suggests that omitting water exchange at the boundary of the restriction space of water diffusion leads to an underestimation of $V_{in}$ but has little influence on the estimation of the space size (43). Whether omitting the permeability of cell membrane or restriction space could lead to an overestimation or underestimation of the size of the restriction space in GBM needs further investigation. Anisotropic diffusion was not considered in this study, which could be considered in the future work, particularly for the regions with tumor cell infiltration. Nevertheless, the microstructure model using a clinically available diffusion pulse sequence leads to better differentiation between the hypercellular tumor and normal tissue or edema than the bi-exponential and mono-exponential models.

CONCLUSION

The MSM considers the restricted diffusion of intracellular water, and differentiates hypercellular tumor of GBM from normal tissue and edema better than free diffusion–based mono-exponential and bi-exponential models. Some parameters from the MSM also reveal the similarity in normal tissue. Further development and histopathological validation of the MSM will warrant its role in clinical management of GBM.

Supplemental Materials
Supplemental Appendix: https://doi.org/10.18383/j.tom.2020.00018. sup.01

ACKNOWLEDGMENT

This work was in part supported by NIH/NCI grant UO1 CA183848. We would like to thank Dr. Himanshu Bhat from Siemens Healthineers for providing information on diffusion gradient pulses.

REFERENCE


16. White NS, McDonald CR, Farid N, Kupemsky JM, Kesari S, Dale AM. Improved conspicuity and delineation of high-grade primary and metastatic brain tumors using


Efficient CT Image Reconstruction in a GPU Parallel Environment

Tomás A. Valencia Pérez1, Javier M. Hernández López1, Eduardo Moreno-Barbosa1, Benito de Celis Alonso1, Martín R. Palomino Merino1, and Victor M. Castaño Meneses2

1Faculty of Mathematical and Physical Sciences, Benemérita Universidad Autónoma de Puebla, Puebla, México; and 2Molecular and Materials, Engineering Department, Universidad Nacional Autónoma de México, Queretaro, México

Corresponding Author:
Victor M. Castaño, Meneses, PhD
Molecular and Materials, Engineering Department, Universidad Nacional Autónoma de México, Queretaro, Mexico 76230;
E-mail: vmcastano@unam.mx

Key Words: Computed tomography, iterative algorithms, GPU, parallelization, reconstruction, image quality
Abbreviations: computed tomography (CT), maximum likelihood expectation maximization (MLEM), central processing unit (CPU), graphics processing unit (GPU), Shepp–Logan phantom (SLP), mean squared error (MSE), peak signal-to-noise ratio (PSNR), structural similarity index (SSIM)

ABSTRACT

Computed tomography is nowadays an indispensable tool in medicine used to diagnose multiple diseases. In clinical and emergency room environments, the speed of acquisition and information processing are crucial. CUDA is a software architecture used to work with NVIDIA graphics processing units. In this paper a methodology to accelerate tomographic image reconstruction based on maximum likelihood expectation maximization iterative algorithm and combined with the use of graphics processing units programmed in CUDA framework is presented. Implementations developed here are used to reconstruct images with clinical use. Timewise, parallel versions showed improvement with respect to serial implementations. These differences reached, in some cases, 2 orders of magnitude in time while preserving image quality. The image quality and reconstruction times were not affected significantly by the addition of Poisson noise to projections. Furthermore, our implementations showed good performance when compared with reconstruction methods provided by commercial software. One of the goals of this work was to provide a fast, portable, simple, and cheap image reconstruction system, and our results support the statement that the goal was achieved.

INTRODUCTION

X-ray computed tomography (CT) is a nondestructive technique in which a source of X-rays (ionizing radiation) revolves around an object of interest, generating axial images of its structure. CT is today an indispensable tool in medicine for the diagnosis of multiple diseases (1). Since its introduction in 1970, there are about 30,000 CT scanners in the world and its number continues to increase exponentially (1). Even if CT techniques represent one of the greatest developments in the field of X-rays in the past 50 years (1), there is still room for improvement. Some of the proposed lines of work to achieve this objective include reduction of patient exposure time, reduction of acquisition and processing times, development of new techniques with new functionalities, and cost reduction (2). Some of the solutions to these points have so far focused on the development of higher performance hardware that will lower costs for and dose delivery to patients without compromising the effectiveness of the diagnosis. Work has also been done on the use of iterative methods for image reconstruction, reducing processing times and reducing the dose necessary for the acquisition of images with diagnostic value.

Current CT scanners use filtered back projection (FBP) for regular image reconstruction (3, 4). This technique is mathematically based on the radon transform (RT) (5, 6). In this technique, the image of an object is reconstructed from a set of X-ray projections of the aforementioned object (7). There are many computational algorithms that can be used to solve the RT. The theoretical framework of reconstruction methods can be found in the literature (2, 8, 9). Nevertheless, these are usually classified into 2 methodologies: analytical reconstruction methods and/or iterative reconstruction methods. In the first method, the use of FBP algorithms has been the standard to date (10), because it is a simple and fast reconstruction algorithm with low computational cost. However, it is prone to the appearance of artifacts and sometimes has low spatial resolution. In addition, it requires a large number of projections without noise for the reconstruction solution to be of quality. This is far from real-life situations, in which only a finite set of projections will be available. Moreover, “approximate” reconstruction methods are also known as iterative algorithms. These are less sensitive to incomplete sets of data, noise, and artifacts (10, 11). This translates into better quality of the reconstructed images. In addition, by requiring less information for reconstruction, they can use fewer projections/radiation by reducing the doses delivered in a study to a patient. For a review of the mathematics behind these methods, see the...
literature (12). The most common between them is maximum likelihood expectation maximization (MLEM) (13). Nevertheless, the main limitation of these techniques with respect to analytical reconstruction methods is their high computational cost.

One way to deal with the limitation of long computational times is to parallelize the reconstruction of an image. To this end, the graphics processing unit (GPU) can be used in combination with the parallel programming environment CUDA (compute unified device architecture) (14-16), same that allows to handle and process data in GPUs. It is made up of execution units, called streaming multiprocessors, which in turn are formed by computing cores called streaming processors, or CUDA cores. It is these nuclei that execute the instructions. That is, the mathematical operations or the routing and transfer of data in memory. The programming model of graphic cards involves both the central processing unit (CPU) and the GPU. The sequential part of the applications is executed in the CPU (Host), and the intensive computing part is accelerated by using the GPU (Device) in parallel. This technology is basic for the use and application of GPUs in fields as diverse as: bioinformatics, computational chemistry, computer imaging and vision, climatologic, numerical analysis, etc. (17).

Medical image processing is one of the first fields in which GPUs were used (18, 19). Tatarchuk et al. (20) presented a set of methods that allowed interactive exploration of medical data sets in real time. Flores et al. (21, 22) presented the development of a fast algorithm implemented in GPUs to reconstruct high-quality images from projected, sampled, and noisy data. Belzunce et al. obtained a parallel implementation under CUDA for nuclear medicine data (SPECT and PET). An acceleration factor of up to 85 times was achieved with respect to a single-wire CPU implementation (23). Xie et al. proposed a way to synchronize the acquisition of CT data with the GPU analysis. This reduced the computational analysis time between 10 and 16 times (24). Finally, Xing et al. (25) showed more applications on the use of GPUs in this area.

In this study, medical CT images will be reconstructed using an iterative method that will solve the MLEM reconstruction algorithm. (13) This will be done using GPUs and CUDA to parallelize the analysis and reduce its computational costs. This analysis will be performed under nonideal noise conditions and their results will be compared with those of commercial medical imaging programs. Another objective of this study is that the work scheme developed can be implemented in a commercial laptop with a graphics card compatible with CUDA. This way and using a system with minimal computing requirements; a fast, simple and low-cost image reconstruction platform can be provided. The article is structured as follows: the next section presents the methodology used in the processing of the images comprising the structure of the algorithm modules, the quality metrics applied to the reconstructions obtained, and the software and hardware used. This is followed by the results section in which the benefits of the CUDA implementation of the MLEM algorithm is presented and compared.

**METHODOLOGY**

First, the MLEM algorithm was implemented in series, both in MATLAB® (MathWorks, Natick, MA) and in C. Second, the same algorithm was implemented in parallel using CUDA. Third, different image quality parameters were calculated for a series of clinical and phantom images based on the reconstruction time and the number of iterations. In a couple of cases Poisson noise was added to blur the original images and perform an evaluation of the effect of noise on the different reconstruction parameters. Finally, results of the reconstructed images were compared with those of commercial software.

**MLEM Algorithm**

Lange et al. (26) presented the MLEM reconstruction algorithm for emission and transmission tomography. The MLEM algorithm was derived from the following mathematical expression:

\[
x_{1}^{t+1} = \frac{x_{1}^{t} \cdot \sum_{i=1}^{M} \sum_{j=1}^{N} a_{ij} y_{j}}{\sum_{i=1}^{M} a_{ij} x_{1}^{t}}
\]

Here \(x_{1}^{t}\) is a pixel, \(j\) is the image \(x\) in the \(r\)-th iteration, \(y_{i}\) is the interval \(i\) of the projection measured given by the CT scanner, \(a_{ij}\) is the system matrix \(M \times N\) whose coefficients connect the \(x_{1}\) values of the image with projections \(y_{i}\) and describe the probability of detecting a photon in pixel \(j\) for the \(i\) bin.

All algorithm implementations developed in this study required of an \(a_{ij}\) system matrix that was used in the resolution of the MLEM algorithm given in the literature (1). Creating a system matrix \(a_{ij}\) depended on the number of projection lines \(y_{i}\), the number of projection angles and the size of the image to be reconstructed \(n \times n\). This matrix could be calculated in different ways, and in this project, the methodology given in the literature (27-29) was used to generate and model the matrix in MATLAB.

**Implementation of the Algorithm for CT Image Reconstruction**

As already mentioned, to compare processing times and performance, the MLEM algorithm was implemented, first serially, in MATLAB and C languages, and then in parallel using CUDA on the graphics card.

**Serial Implementations.** Two versions of the MLEM algorithm were developed, one in MATLAB and one in C (Serial-MATLAB and Serial-C, respectively). The different processes that are part of the programmed algorithms are described as follows:

1. **Input Data:** In a first step, the files provided by the CT scanner were loaded. This included the system matrix \(a_{ij}\) and the measured projections \(y_{i}\). These structures were called matrix and vector, respectively. The \(a_{ij}\) matrix and the vector \(y_{i}\) provided information such as the number of rows and columns of the system matrix or the number of elements of the projection vector \(y_{i}\).

2. **Initial Image Estimation:** To reduce the number of iterations, an initial estimate of the image was used—\(x_0^{t}\). It was created either with all its values equal to the average of the projection data or by setting its values at an intensity between 1 and 255 (grayscale).

3. **Forward-Projection:** The first iteration was performed on the image with an initial estimate that simulated the projections of a given \(x_0^{t}\). This was required for the calculation
of \( y_{i}^{\text{simulated}} \) to determine the corresponding projection values.

4. Comparison with the measured data: The comparison step was performed by dividing the original projection data \( y_{i} \) with the data from simulated projections \( y_{i}^{\text{simulated}} \), obtaining a relation \( Y_{i} = \frac{y_{i}^{\text{simulated}}}{y_{i}} \). Here \( Y_{i} \) formed a multiplicative correction factor for each projection.

5. Back-Projection: The calculated \( Y_{i} \) was back-projected to obtain a correction factor \( X_{j} \) for the initial image. This was done with the following system calculation \( X_{j} = \sum_{i=1}^{M} a_{ij} Y_{i} \).

6. Normalization: The correction factor \( X_{j} \) was normalized before being multiplied by \( x_{i}^{0} \). It was divided by a weighting term based on the system model to apply the desired intensity of each image correction factor. This produced \( e_{j} = X_{j} / \sum_{i=1}^{M} a_{ij} \), where \( e_{j} \) represented the normalized correction factor.

7. Image Update: The initial estimation of image \( x_{i}^{0} \) was multiplied by the normalized correction factor \( e_{j} \), obtaining the new image estimation \( x_{i}^{j} \), which was then fed into the algorithm in the initial image estimation step.

These processes were executed iteratively until the relative error between the estimated and the actual image reached a value <5%. To verify that this behavior was maintained, up to 1000 iterations were performed.

From the serial implementation of the algorithm, the performance was measured for each of the processes. It was found that the processes that consumed the most resources were the steps of Forward-Projection and Back-Projection. Between them, they used ~75% of computational time.

Parallel Implementations. Two versions were developed: parallel using GPU (Parallel-GPU) and parallel using GPU but optimizing data transfer (Parallel-Opt). Both implementations resolved equation (1) using both C and CUDA. CUDA was used to exclusively develop the Forward-Projection and Back-Projection processes, known for being the ones that consume the most computing resources.

The calculations performed on the data structures were handled in C by four kernels (Forward-Projection, Back-Projection, Normalization, Image Update), which distributed the data to the blocks of threads. This was done in similarity to the serial case. The different processes of the parallel versions are described as follows:

1. Forward-Projection module: To this end, a Forward-Projection module, using resources from the cuBLAS library from CUDA, was programmed. No Forward-Projection or Back-Projection modules exist per se in cuBLAS, so they were developed specifically by authors for this project. Matrix \( a_{ij} \) was loaded to the Host and transferred to the Device; vector \( x_{i}^{0} \) was generated in the Host and also transferred to the Device. Matrix \( a_{ij} \) was partitioned in grids of threads blocks. Each thread read a row of the \( a_{ij} \) matrix and multiplied it by vector \( x_{i}^{0} \) to obtain a vector in the Device. The resulting \( y_{i}^{\text{simulated}} \) vector was then transferred to the Host.

2. Back-Projection module: As in the previous process, the Back-Projection module was launched on the Device. This module required most of the computational resources for the entire reconstruction process, as it consisted of 2 operations: calculation of the transposition of \( a_{ij} \) and that of its product with vector \( Y_{i} \).

3. Optimization of GPU computing time. The main data transfers between Host and Device were carried out when the Forward-Projection and Back-Projection processes were executed. To improve the performance of the programs, the transfer of data between the RAM of the CPU and the global memory of the GPU was handled by joining the Forward-Projection, comparing the measured data and Back-Projection modules in one function, and maintaining data calculations in the GPU. This created the optimized parallel version (Parallel-Opt).

As in the serial case, these processes were executed iteratively until the relative error between the estimated and the actual image reached a value <5%.

Image Quality Parameters

In this study, 2 clinical CT images were initially reconstructed along with a 512 × 512 image of the Shepp–Logan phantom (SLP) (32). The first clinical image was an axial cut of the brain that covered the cerebellum, as well as the nasal and auditory cavities (33). The second clinical image represented an abdominal axial cut through the lungs, heart, pancreas, and liver (33). The matrices of both CT images were 512 × 512 with a dynamic range of 8 bits per pixel, with a system matrix of 23944 × 262144 and a vector projection of 23944 × 1 for a 5° sweeping angle.

To evaluate the quality of the reconstructed images, different image quality parameters were used. The main parameters studied were: contrast-to-noise ratio and signal-to-noise ratio (34). The first indicated how well 2 neighboring tissues could be distinguished. The latter how strong a signal was with respect to background noise. Other image quality parameters evaluated were: mean squared error (MSE) (35), peak signal-to-noise ratio (PSNR) (36) or structural similarity index (SSIM) (37). MSE provided a measurement of the differences between a reconstructed image and a reference image. The lower the MSE value, the better the result. The PSNR described the relationship of the maximum possible power of a signal with the power of noise corruption. It was represented on a decibel scale, and a high PSNR value meant that the reconstruction was of high quality. Finally, the SSIM was used to measure the similarity between 2 images, namely, original CT image and reconstructed images. SSIM values were normalized between 0 and 1; being 1 the situation in which both images were equal.

All image quality parameters were measured for all the CT images and for the 4 implementations of the algorithm developed and the two versions of TIGRE. In addition, 2 scenarios were considered, in which 5% Poisson noise was added to the images under study to assess the impact of noise on the reconstruction time and quality.

Comparison with Commercial Software

To compare the implementations developed in this work for the MLEM algorithm with those of an external (commercial) code, the TIGRE MatLab library was used (38). From it, 2 iterative
tomographic reconstruction toolboxes based on GPUs were used, namely, the OS-SART algorithm that is used to solve reconstructions of the conical beam CT images using simultaneous algebraic reconstruction techniques with ordered subsets, and the ASD-POCS algorithm that is based on the most pronounced adaptive descent projection in convex subsets (39).

**Hardware and Software**

Serial and parallel implementations were developed and executed on a laptop with an Intel Core i7-5500U processor running at 2.40 GHz and equipped with an NVIDIA GeForce 840 M graphics card, with a global memory of 2048 MB and 384 CUDA cores on an Ubuntu operating system 15.04 (40). MATLAB R2017a was used (27) as well as GCC 4.9.2 (41) for the serial part. CUDA C 7.5 (31) was used for the parallel calculations. The Open Source Computer Vision library (OpenCV 3.2.0) was used to calculate image quality measurements (42, 43).

**RESULTS**

Computer algorithms and code developed here will be available after petition to authors.

Figures 1 to 3 present the results obtained for the 4 implementations of the algorithm developed here (Serial-MatLab, Serial-C, and the 2 parallel versions) and the external test software (2 versions).

Images in Figure 1 correspond to a reconstruction with Parallel-OpT software with a maximum of 1000 iterations. In the first row, SLP and two 512 × 512 CT images are presented. In the following row 4 128 × 128 regions of interest of the previous images are presented. These regions were selected randomly by authors. The original image of the SLP (first column) is shown in the first row of Figure 1. In the second column, the reconstructed image “iteration 64” is shown, which will be referred from now on as “the best” and called iteration 64. Number 64 just quantifies the number of iterations necessary to reach it. In the third column, the reconstructed image called “functional” is shown. In this case it is called “iteration 35,” and it is characterized by the fact that the differences in its quality parameters with those of the “the best” are <50%. The last column shows the “functional” image with the Poisson noise. The CT images of the central row correspond to the axial section of the brain, while the abdominal axial section is presented in the last row of Figure 1. The data of the last 2 rows are presented in the same scheme as in the SLP image.
The values of the different quality parameters of each image are presented in Table 1. The minimum number of necessary iterations to obtain the given image quality value and the time taken to achieve it are also presented.

Figure 2 shows the mean values for the different quality parameters of the reconstructed images of Figure 1: MSE, PSNR, and SSIM. Table 2 vs. number of iterations presents the quantification of these values. Images obtained from reconstruction had 512 x 512 matrices formed from CT projections that varied between 0° and 180° in 5° steps (views). Initially, the quality values of the reconstructed images were calculated for all serial and parallel implementations; however, the differences in the absolute values of these quality parameters were <0.001%.

In addition, the dependence between the optimal number of views/projections for image reconstruction in the implementations was studied. These results are presented as complementary material (see online supplemental Figure 1). In this analysis, the image quality parameters, as well as the time calculations, were compared based on the number of iterations and the number of views. It was found that the use of jumps of 5° per view to cover the entire field of vision, as well as 50 to 100 iterations, produced reliable results without extending the calculation time while obtaining high-quality image parameter values.

It is worth mentioning that when the field of view was covered either with projections that varied between 0°–180° and 0°–360°, the reconstructions of the different programs produced very similar results (regarding image quality parameters). Because of this and from this point, only reconstructions with a system matrix that varied from 0° to 180° were taken into account.

Figure 3 presents the average computation time for the 4 implementations/programs developed by the authors, as well as the 2 versions based on TIGRE. It also shows the results with added Poisson noise. The data are presented versus the number of iterations.

It can be observed that the differences in computation time between the Parallel-OpT and Serial-C implementations were of the order of 12 times faster for the parallel solution. The Parallel-OpT was 170 times faster than Serial-MatLab. Parallel-OpT and Parallel-GPU implementations differed by a factor of 2 when computer processing time was considered. It was observed that the calculation times of the ASD_POCs version of TIGRE were comparable with those of the Serial-C and ~10 times slower than Parallel-OpT. In contrast, TIGRE OS-SART was the fastest of all, being 2 times faster than the Parallel-OpT implementation in CUDA. When the data were analyzed with aggregate noise, no differences were found in the computation time with respect to Parallel-OpT. This was for the reconstruction of the 3 images studied. Finally, it is worth noting that the relationships between the calculation times of implementations of the MLEM algorithm were constant and remained independent of the number of iterations or added noise.

As mentioned earlier and as seen in Figures 2 and 3, Parallel-OpT was up to 13 times faster than ASD_POCs and 7% better in image quality. The OS-SART was, in contrast, 50% faster, but with lower image result quality. The MSE was 720% worse; the PSNR was 7.5% worse.

These differences are highlighted in Figure 4. In Figure 4, a general comparison between Parallel-OpT and the 2 chosen from the TIGRE library are shown. The best quality/time ratio can be appreciated from the first implementation.

Finally, to test the stability and effectiveness of these programs, data from a clinical study were reconstructed using the Parallel-OpT implementation. Further 101 CT images, corresponding to a single axial cut of the pelvis were obtained from each volunteer. Data were downloaded from The Cancer Genome Atlas Sarcoma database (TCGA-SARC) (33); therefore, no ethical permission was needed to perform this study on our side. In each image the SSIM was calculated as in previous sections. The SSIM

<table>
<thead>
<tr>
<th>Table 1. Quality Parameters from Figure 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Original</td>
</tr>
<tr>
<td>The best</td>
</tr>
<tr>
<td>Functional</td>
</tr>
<tr>
<td>Functional + Noise</td>
</tr>
<tr>
<td>Head</td>
</tr>
<tr>
<td>Original</td>
</tr>
<tr>
<td>The best</td>
</tr>
<tr>
<td>Functional</td>
</tr>
<tr>
<td>Functional + Noise</td>
</tr>
<tr>
<td>Abdomen</td>
</tr>
<tr>
<td>Original</td>
</tr>
<tr>
<td>The best</td>
</tr>
<tr>
<td>Functional</td>
</tr>
<tr>
<td>Functional + Noise</td>
</tr>
</tbody>
</table>
Figure 2. Image quality parameters: MSE, PSNR, and SSIM versus number of iterations. Image parameters were obtained after reconstruction with Parallel-OpT. (A) mean squared error (MSE), (B) peak signal-to-noise ratio (PSNR), and (C) structural similarity index (SSIM).
Data are presented in Figure 5 (using MatLab Distribution Fitter app software). SSIM values reached a maximum of 0.9. Data presented a normal distribution form of 0.902 ± 0.019 (mean ± SD). This was confirmed by conducting a Shapiro–Wilk test calculated with SPSS software ($P < .05$). All data, except one, exceeded the threshold value of 0.85 SSIM, a standard limit for good image quality.

**DISCUSSION AND CONCLUSIONS**

This work shows the parallel implementation of the MLEM reconstruction algorithm using graphic card capabilities. Software solutions developed here were compared with reconstructions of clinical images using several image quality parameters (CNR, SSIM, MSE, etc.). Computer time used for image reconstruction in CUDA-based programs was shorter than that associated with serial solutions. Acceleration factors that varied between 2 and 170 times were obtained. In comparison with the TIGRE software, images were obtained of similar quality, but in some cases and for some specific image quality factors, images were 7 times better for our CUDA programs. The addition of 5% noise to signals did not increase the reconstruction time proportionally; in fact, it left it almost unaltered. Moreover, image quality decreased in general by a 5% factor (except for the PSNR parameter that lost 30% of its maximum value in some scenarios). The clinical study results showed that calculations were consistent and reproducible. Furthermore, the quality of image reconstruction was conserved even at short reconstruction times.

**Table 2. Quantification of Data Presented in Figure 2**

<table>
<thead>
<tr>
<th></th>
<th>MSE</th>
<th>PSNR</th>
<th>SSIM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min. Iteration Time (s)</td>
<td>Min. Iteration Time (s)</td>
<td>Max. Iteration Time (s)</td>
</tr>
<tr>
<td>Shepp–Logan</td>
<td>32.2</td>
<td>1000</td>
<td>1548</td>
</tr>
<tr>
<td>Shepp–Logan + Noise</td>
<td>58.95</td>
<td>1000</td>
<td>1595</td>
</tr>
<tr>
<td>Abdominal CT</td>
<td>0.842</td>
<td>146</td>
<td>227.4</td>
</tr>
<tr>
<td>Abdominal CT + Noise</td>
<td>7.46</td>
<td>22</td>
<td>35.98</td>
</tr>
<tr>
<td>Head CT</td>
<td>100</td>
<td>1000</td>
<td>1586</td>
</tr>
<tr>
<td>Head CT + Noise</td>
<td>139.1</td>
<td>1000</td>
<td>1612</td>
</tr>
<tr>
<td>Shepp–Logan with OS_SART</td>
<td>262.2</td>
<td>30</td>
<td>26.8</td>
</tr>
<tr>
<td>Shepp–Logan with ASD_POCS</td>
<td>272.6</td>
<td>40</td>
<td>795.8</td>
</tr>
</tbody>
</table>

**Figure 3.** Calculation of computing times for different implementations developed in this study as well as TIGRE algorithms. An enlarged image of the first 200 iterations is presented in the lower right corner.
All these results imply that the MLEM implementations in CUDA using GPUs’ capabilities presented here were reliable and fast.

With respect to image acquisition, it was shown that to generate the system matrix, only projections in the range of 0° to 180° were needed. This contrasted with the acquisition of full-field projections (from 0° to 360°). This is a well-established fact in the field, so no data or images were presented to support this fact in this work. In general, acquisitions over small angles (1°, 2°, or 3°) at 1000 iterations took up to 5000 seconds. This amount of time was obviously greater than reconstructions in steps of 5° or larger. Measurements with larger angles were described in the online supplemental Figure 1. There it could be appreciated that the 5° measurements presented a perfect compromise between reconstruction quality and reconstruction time. Therefore, this value was the one selected for reconstruction in this study.

Figure 2 showed how the different image quality metrics (MSE, PSNR, SSIM) decreased first and then increased rapidly as the number of iterations increased. This allowed to deduce an “ideal” range of iterations for image reconstruction for this software implementation that was set between 50 and 100. Therefore, as it can be seen in Figure 2B, the PSNR parameter for the SLP and the abdominal image reached a local maximum value within the range of 50 to 100 iterations. For the head image, this effect was similar but less prominent. Obviously when larger number of iterations were performed, values of this last parameter were larger and therefore better. Nevertheless, the time used to improve 5% of this parameter could be easily 5–10 times larger. That made this approach of larger iterations less attractive for practical purposes. The maximum values of the SSIM metric for the SLP, the head, and the abdominal images were 0.992, 0.987, and 0.986, respectively. These maximum values were reached within the proposed range of iterations (see Figure 2C). After the maximum, the values remained stable and were not affected by a greater number of iterations. As indicated before in the text, values of SSIM over 0.85 represented images of good clinical quality.

Once the iteration interval and the degrees between projections used to acquire the images were established, the processing times of the different implementations developed were compared (Figure 3). It was found that the difference in time between Serial-C and Parallel-OpT to achieve the same result was ~1000 seconds at 50 iterations (12 times more) and 9000 seconds in 1000 iterations (170 times more). These results showed that implementations of the MLEM algorithm based on CUDA could perform good image reconstructions in shorter times than traditional CPU-based methods. Surprisingly, the addition, the Poisson noise did not affect the reconstruction times of the Parallel-OpT implementation. When compared with the results from TIGRE implementations, equal or better time performances were achieved, always with better image quality reconstruction.

If the reader considers that one of the basic ideas behind this work was to implement the MLEM algorithm in a novel way by using the GPUs’ capabilities in commercial/low capability computing systems, he/she can observe that this was accomplished. Therefore, the solution for image analysis presented here could be used in institutions or health systems with low financial

![Figure 4. SSIM versus number of iterations versus time for: (A) Parallel-OpT, (B) OS_SART, and (C) ASD_POCS. Embedded in the image, the color code indicates the effect of time on SSIM values.](image-url)
Figure 5. Histogram of SSIM values of reconstructed images from a clinical study. The data present an equally normalized distribution of values. (A) Parallel-OpT, (B) TIGRE 1, and (C) TIGRE 2.
resources in a reliable way, conserving image quality and maintaining reconstruction times as low as possible.

ACKNOWLEDGMENTS
We would like to thank CONACyT México for the support of the first author with a doctoral grant.

REFERENCES

Disclosures: The authors have nothing to disclose.

Conflict of Interest: None reported.
Scope
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. Studies involving hardware and software advances along with chemical and molecular probe developments are also of high interest for publication in Tomography.

Submission
Submission of manuscripts can be accomplished through the web site www.Tomography.org using our online submission system.

Open Access Compliance
Funding Agencies and Article Repositories
Certain funders, including the NIH, members of the Research Councils UK (RCUK), and Wellcome Trust require deposit of the Accepted Version in a repository after an embargo period. Tomography allows all of its authors to deposit their articles on such repositories immediately upon publication.

Instructions for Authors
Research Articles
Original studies considered for publication will provide the field of imaging with significant advances in basic, preclinical, clinical, hardware and software using single or multi-modal imaging technologies. In addition, studies that use imaging technologies to advance basic and clinical medicine through interrogation of biological and pathological processes are also of interest.

Advances in Brief
Advances in Brief are timely high impact contributions on a topic important to imaging researchers. These Advances will undergo an accelerated review process. Papers submitted as Advances in Brief may combine the Results and Discussion sections.

Reviews
Tomography will publish Reviews which are timely and important to imaging researchers. Reviews should be written as concisely as possible. All Review articles will be subject to peer review. The parameters for this category are provide below but serve only as guidelines as longer reviews will be considered if the topic warrants additional space in the journal.

Perspectives
Tomography will publish Perspectives which are used to present new insights on an active or emerging area of imaging research or application of imaging technologies in the preclinical or clinical setting along with the author’s personal viewpoint on the area including unresolved areas for future research along with key issues facing development and/or deployment.

Image Reports
Image Reports are discrete Reports (approximately 1,500 words) that provide significant advances and offer unique insights into the application of imaging technologies for the further advancement of preclinical or clinical diagnosis or treatment response. No specific limits are place on overall numbers of figures or tables but up to 4 would be reasonable.

Consensus Papers
Papers which provide a broad consensus on a timely topic related to imaging which is sponsored from investigators participating at a national meeting or meeting sponsored by a government funding agency. Authors are encouraged to contact the editorial office prior to submission.
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.
Welcome to the World of Preclinical Imaging

- Introducing the PET/CT Si78 high performance PET and CT
- Total body PET for mice and rats
- Powered by ParaVision 360
- Simplified workflow
- Supports a wide range of applications

Discover more at: www.bruker.com/pci

Innovation with Integrity