Quantification of nonenhancing tumor burden in gliomas using effective $T_2$ maps derived from dual echo turbo spin echo MRI

Benjamin M. Ellingson, Ph.D. $^{1-6}$, Albert Lai, M.D., Ph.D. $^{1,6,7}$, Huytram N. Nguyen $^{1,7}$, Phioanh L. Nghiemphu, M.D. $^{1,6,7}$, Whitney B. Pope, M.D., Ph.D. $^2$, and Timothy F. Cloughesy, M.D. $^{1,6,7}$

$^1$ UCLA Neuro-Oncology Program, David Geffen School of Medicine at UCLA
$^2$ Department of Radiological Sciences, David Geffen School of Medicine at UCLA
$^3$ Biomedical Physics Program, David Geffen School of Medicine at UCLA
$^4$ Department of Bioengineering, Henry Samueli School of Engineering and Applied Science at UCLA
$^5$ UCLA Brain Research Institute, David Geffen School of Medicine at UCLA
$^6$ Jonsson Comprehensive Cancer Center at UCLA
$^7$ Department of Neurology, David Geffen School of Medicine at UCLA

Research Support: National Brain Tumor Society Research Grant (BME, TFC); NIH/NCI 1 R21 CA167354-01 (BME); UCLA Institute for Molecular Medicine Seed Grant (BME); UCLA Radiology Exploratory Research Grant (BME); University of California Cancer Research Coordinating Committee Grant (BME); ACRIN Young Investigator Initiative Grant (BME); Art of the Brain (TFC); Ziering Family Foundation in memory of Sigi Ziering (TFC); Singleton Family Foundation (TFC); and Clarance Klein Fund for Neuro-Oncology (TFC). National Institute of Health National Institute of General Med-
ical Sciences training grant, GM08042 (KL) and the University of California Los Angeles Medical Scientist Training Program (KL)

Conflicts of Interest: Drs. Timothy F. Cloughesy, Albert Lai, Whitney Pope, and Benjamin Ellingson are paid consultants for Genentech, Inc., and Hoffman-La Roche, Ltd. Drs. Ellingson and Pope are also a paid consultant for MedQIA, LLC.

Address Correspondence To:
Benjamin M. Ellingson, Ph.D.
Assistant Professor of Radiology, Biomedical Physics, Psychiatry and Bioengineering Departments of Radiological Sciences and Psychiatry
David Geffen School of Medicine
University of California – Los Angeles
924 Westwood Blvd, Suite 615
Los Angeles, CA 90024

Word Count: 5465
STATEMENT OF TRANSLATIONAL RELEVANCE

The current study provides a comprehensive evaluation of the use of T2 relaxation rates estimated from dual echo turbo spin echo images to objectively define nonenhancing tumor burden in patients with gliomas. Results suggest T2 measurements are relatively stable across scanners and may be useful for delineating nonenhancing tumor from other tissues. Results also suggest the use of T2-defined nonenhancing tumor burden may be useful for prognostic and response assessment purposes.
ABSTRACT (250/250 WORDS)

**Purpose:** Evaluation of nonenhancing tumor (NET) burden is an important, yet challenging part of brain tumor response assessment. The current study focuses on using dual echo turbo spin echo MRI as a means of quickly estimating tissue T2, which can be used to objectively define NET burden.

**Experimental Design:** A series of experiments were performed to establish the use of T2 maps for defining NET burden. First, variation in T2 was determined using ACR water phantoms in 16 scanners evaluated over 3 years. Next, sensitivity and specificity of T2 maps for delineating NET from other tissues was examined. Then, T2-defined NET was used to predict survival in separate subsets of glioblastoma patients treated with radiation therapy, concurrent radiation and chemotherapy, or bevacizumab at recurrence.

**Results:** Variability in T2 in the ACR phantom was 3-5%. In training data, ROC analysis suggested that 125ms < T2 < 250ms could delineate NET with a sensitivity >90% and specificity >65%. Using this criterion, NET burden after completion of radiation therapy alone, or concurrent radiation therapy and chemotherapy, was shown to be predictive of survival (Cox, P<0.05), and the change in NET volume before and after bevacizumab therapy in recurrent glioblastoma was also a predictive of survival (P<0.05).

**Conclusions:** T2 maps using dual echo data are feasible, stable, and can be used to objectively define NET burden for use in brain tumor characterization, prognosis, and response...
assessment. The use of effective T2 maps for defining NET burden should be validated in a randomized clinical trial.
INTRODUCTION

The use of contrast enhancing tumor burden has been the standard for brain tumor response assessment for more than 60 years; however, approximately 30-40% of patients are estimated to experience nonenhancing tumor progression prior to changes in contrast enhancement (1, 2). Although contrast enhancing tumor is thought to represent the most aggressive portion of the tumor (3, 4) and a large percentage of high grade gliomas have a significant enhancing component (5), malignant gliomas are known to contain proportions of both neovascularized and infiltrative tumor (6, 7). Since a substantial proportion of treated tumors can have nonenhancing tumor progression (1, 6), and progression of nonenhancing tumor can lead to neurologic decline, there is an emergent need for consideraton of nonenhancing tumor burden in an updated response criteria in neuro-oncology, which was further outlined in the Jumpstarting Brain Tumor Drug Development Coalition and Food and Drug Administration (FDA) Workshop in January 2014 (8, 9).

Damadian (10) first documented distinct differences in proton relaxation rates between normal and cancerous tissues as early as 1971, which was also confirmed by various groups in the subsequent years (11-15). Clinical diagnosis and monitoring of brain tumors with MRI, however, remains reliant on the use of $T_1$ or $T_2$ “weighted” images, which express qualitative changes in the MR signal. $T_2$ relaxometry can theoretically be performed using a simple modification to the standard $T_2$-weighted MRI sequence that results in obtaining multiple images at various echo times, which is essentially “free” information (i.e. no substantial added scan time). This information can in theory be used to create quantitative maps of $T_2$, which is specific for various tissue types including cancer but relatively independent of MR acquisition parameters such as echo time (TE). For example, an early study by Hoehn-Berlage et al. (16) correlated tumor tissue in a...
cat model with quantitative T2 maps at 4.7T and demonstrated that T2_{edema} > T2_{tumor} > T2_{normal WM}, consistent with the current clinical observation of moderately hyperintense lesions on T2-weighted images being associated with nonenhancing tumor (17). This difference in T2 relaxation rates between edema, tumor, and normal tissues has been characterized in both rodent models (18) and in human tumors (19-21). Additionally, studies have shown that T2 relaxation tends to be monoexponential (16), suggesting evaluation of nonenhancing tumor may be possible with a single T2 parameter. We hypothesize that “effective” T2 relaxometry performed using the turbo spin echo may be useful for objectively defining nonenhancing tumor tissue using a specific range of T2 values.

Accurate measurements of T2 relaxation times typically involve using relatively time inefficient multiecho Carr-Purcell-Meiboom-Gill (CPMG) sequences. Dual echo turbo spin echo (TSE) sequences are an attractive clinical alternative to CPMG sequences (22, 23), allowing for “effective” T2 estimation (T2_{eff}) in clinically feasible scan times and clinically useful spatial coverage for a variety of anatomical locations (24-28). In the current study we examined the stability of T2_{eff} estimates in a standard water phantom across a variety of scanners, used dual echo TSE data from 50 patients with a range of glioma grades scanned on a variety of scanners and field strengths to define T2_{eff} characteristics for various tissues, then applied this information to a cohort of newly diagnosed glioblastoma patients treated with radiation therapy, a cohort of newly diagnosed glioblastoma patients treated with concurrent radiation therapy and temozolomide, and a set of recurrent glioblastoma patients treated with bevacizumab to determine whether the volume of T2_{eff}-defined nonenhancing tumor can be used to predict progression-free (PFS) and overall survival (OS).
METHODS

Summary of Experiments

A series of experiments were performed in the current study to establish the use of $T_2^{\text{eff}}$ in defining nonenhancing tumor burden for use in quantification of tumor burden and response assessment. First, the normal variation expected in $T_2^{\text{eff}}$ estimates was determined on the American College of Radiology (ACR) standard water phantom using dual echo TSE images collected for 16 different scanners and field strengths. Next, the sensitivity and specificity of $T_2^{\text{eff}}$ in delineating nonenhancing tumor from edema, necrotic tissue, and normal-appearing brain tissue was estimated in a training set of 50 glioma patients in order to define cutoffs for defining nonenhancing tumor. Using thresholds defined by these training data, we examined an independent set of 25 newly diagnosed glioblastoma patients who were treated with radiation therapy, but no concurrent temozolomide and never were exposed to bevacizumab, to determine whether nonenhancing tumor burden evaluated at this post-radiation therapy time point could predict PFS and OS. Then, another independent set of 35 glioblastoma patients treated with concurrent radiation therapy and temozolomide, but were never exposed to bevacizumab, were evaluated at the post-radiation therapy time point to determine whether nonenhancing tumor burden prior to adjuvant therapy could predict PFS and OS. Lastly, we examined a set of 24 recurrent glioblastoma patients treated with bevacizumab to determine whether the change in nonenhancing tumor burden could be used to predict PFS and OS.

Phantom Studies
A total of 16 MRI scanners spanning various field strengths and manufacturers were evaluated from 2010-2015 using the American College of Radiology (ACR) accreditation phantom (Figure 1). Dual echo TSE data were acquired according to ACR guidelines (www.acr.org/~/media/ACR/Documents/Accreditation/MRI/LargePhantomInstructions.pdf). In particular, dual echo TSE data were obtained with TE1 = 20ms, TE2 = 80ms, repetition time (TR) = 2000ms, field-of-view (FOV) = 25cm, 11 slices, 5mm slice thickness with interslice gap of 5mm, number of excitations (NEX) = 1, and matrix size of 256x256. Quantification of $T_2^{\text{eff}}$ was determined by

$$T_2^{\text{eff}}(x,y,z) = \frac{TE_2 - TE_1}{\ln\left(\frac{S_1(x,y,z)}{S_2(x,y,z)}\right)}$$

[Eq. 1]

where $T_2^{\text{eff}}$ is the estimated effective $T_2$ in the voxel at coordinate $(x,y,z)$, $TE_1$ and $TE_2$ are the echo times of the first and second MR image, respectively, and $S_1$ and $S_2$ are the MR signal intensity in the voxel at coordinate $(x,y,z)$ at the two echo times, respectively. The mean value of $T_2^{\text{eff}}$ was determined for the entire phantom at slice 6 using the standard ACR slice prescription, corresponding to a slice containing relatively homogeneous signal intensity and no internal geometric structures.

Patients

A total of 134 patients with gliomas who were evaluated using dual echo TSE between 2000-2008 were retrospectively evaluated in the current study. All patients gave informed written consent to be included in our institutional review board approved neuro-oncology database prior to this study. These 134 patients were broken into 4 groups for training and testing purposes: 1) 50
patients with various glioma grades used for training; 2) 25 newly diagnosed glioblastoma patients treated with radiation therapy and no concurrent chemotherapy, none of which ever were exposed to bevacizumab; 3) 35 newly diagnosed glioblastoma patients treated according to the current standard of care including radiation therapy with concurrent temozolomide (TMZ), followed by adjuvant TMZ; and 4) 24 recurrent glioblastoma patients treated with bevacizumab. Because this was a retrospective study, sample sizes were based solely on availability of data in patients that met the inclusion criteria outlined below.

*Training Set for Defining Nonenhancing Tumor.* A total of 50 glioma patients with obvious nonenhancing tumor who were evaluated using dual echo TSE between 2000-2008 were retrospectively evaluated. Patients were selected so the training set contained tumors with a variety of WHO grades, histological subtypes, and were at various stages of their disease. A total of 34/50 (68%) patients were male, approximately half the patients (26/50, 52%) were scanned on a 1.5T MR scanner and the other half (24/50, 48%) was scanned on a 3T scanner. A total of 32/50 (64%) patients had recurrent tumors. Approximately 15/50 tumors (30%) were World Health Organization (WHO) grade II astrocytomas, 15/50 tumors (30%) were WHO III anaplastic astrocytoma (9 of 15) or oligodendrogliomas (6 of 15), and 20/50 (40%) were WHO IV glioblastomas. To estimate $T_2^{\text{eff}}$ values in specific tissues, circular regions of interest (ROI) 1cm in diameter were placed in areas of obvious normal-appearing white matter (NAWM), edema, nonenhancing tumor (NET), and necrosis on T2-weighted images in these 50 patients. Regions of NET were defined as having a lower T2 signal intensity compared with CSF, architectural distortion of the gray-white matter junction, or significant mass effect, while edematous tissue was defined as having a higher T2 signal intensity approaching that of CSF, T2 hyperintensity respecting the
gray-white matter junction, no significant mass effect, and presence of “finger-like” projections of high intensity T2 signal within deep white matter, as defined previously by Pope et al. (17). The average T2\textsuperscript{eff} for these regions were retained and used to define field-strength-specific thresholds to accurately delineate nonenhancing tumor from both edema and NAWM using Receiver-Operator Characteristic (ROC) analysis.

**Newly Diagnosed Glioblastoma Patients Treated with Radiotherapy.** A total of 25 newly diagnosed glioblastoma patients (average age = 53.2, range = 24-78 years old; average Karnofsky Performance Status (KPS) = 83, range 70-100; 16/25 were male (64%)) treated with 60Gy intensity modulated radiation therapy (IMRT) with no concurrent chemotherapy who were evaluated following completion of IMRT using dual echo TSE between 2000-2008 were retrospectively evaluated (Figure 1A). None of the 25 patients were ever treated with bevacizumab during their clinical history. All 25 patients were treated with adjuvant chemotherapies, including isotretinoin (Accutane, Roche Pharmaceuticals, Nutley, NJ; 100 mg/m\textsuperscript{2} orally for 21 consecutive days followed by 7 drug-free days), temozolomide (150-200 mg/m\textsuperscript{2} intravenously or orally for 5 days of a 28 day cycle), lomustine (CCNU, Bristol-Myers Squibb, 110mg/m\textsuperscript{2} orally once every 6 weeks), or carboplatin (area under the plasma curve 4 mg/ml-min on day 1 of a 28 day cycle), until treatment failure.

**Newly Diagnosed Glioblastoma Patients Treated with Concurrent Radiotherapy and Temozolomide Followed by Adjuvant Temozolomide.** A total of 35 newly diagnosed glioblastoma patients (average age = 58, range = 27-77 years old; average KPS = 85, range = 70-100; 27/35 were male (77%)) treated concurrently with IMRT and temozolomide following by adjuvant temozolomide (TMZ) who were evaluated following completion of concurrent IMRT and TMZ
using dual echo TSE between 2000-2008 were retrospectively evaluated (Figure 1B). Patients were treated with 60Gy IMRT with concomitant TMZ (75 mg/m² orally or intravenously for 42 consecutive days) followed by adjuvant TMZ (150 mg/m² orally or intravenously for 5 consecutive days in the first 28 day cycle, followed by 200 mg/m² orally or intravenously for 5 consecutive days in the first 28 day cycle for a maximum of 6 cycles). Adjuvant TMZ was continued until disease progression, death, or completion of 6 cycles. A total of 8 of 35 patients (23%) were treated with bevacizumab during their clinical history, and 6 of the 8 patients were treated at first recurrence. Approximately 29 of the 35 patients (83%) were treated with chemotherapy at first recurrence, including isotretinoin (Accutane, Roche Pharmaceuticals, Nutley, NJ; 100 mg/m² orally for 21 consecutive days followed by 7 drug-free days, dose range across all patients = 160mg-200mg; N=9 patients), lomustine (CCNU, Bristol-Myers Squibb, 110mg/m² orally once every 6 weeks, dose range across all patients = 150mg-210mg; N=17 patients), or carboplatin (area under the plasma curve 4 mg/ml-min on day 1 of a 28 day cycle; N=3 patients).

Recurrent Glioblastoma Treated with Bevacizumab. An independent sample of 24 patients with recurrent glioblastoma (average age = 56, range = 27-91 years old; average KPS = 82, range 70-100; 12/24 were male (50%)) who failed IMRT and TMZ who also underwent evaluation at recurrence (baseline) and 2-4 weeks after the first administration of bevacizumab (10 mg/kg intravenously every 14 days) using dual echo TSE between 2000-2008 were retrospectively evaluated (Figure 1C). All patients were treated with bevacizumab every 14 days until disease progression or death.

Magnetic Resonance Imaging
Standard anatomical MR images including dual echo TSE images for patients included in the training dataset were collected on one of the 1.5T or 3T MR scanners listed in Figure 1. Standard anatomic images included an axial T1-weighted fast spin-echo sequence or magnetization-prepared rapid acquisition gradient-echo sequence (repetition time msec/echo time msec/inversion time msec, 400–3209/3.6–21.9/0–1238; section thickness, 3–6.5 mm; intersection gap, 0–2.5 mm; number of signals acquired, one to two; matrix size, 176–512 × 256–512; and field of view, 24–25.6 cm) and a fluid-attenuated inversion-recovery sequence. In addition, axial T1-weighted images enhanced with gadopentetate dimeglumine (Magnevist; Berlex), 0.1 mmol/kg, and acquired shortly after contrast material injection, were matched to nonenhanced T1-weighted images obtained with similar sequence parameters. Dual echo TSE images for the training dataset included both proton density/T2-weighted imaging pairs (e.g. TE₁=5-10ms, TE₂=80-120ms) or two sets of T2-weighted images (e.g. TE₁=80-120ms, TE₂=160-200ms). All therapeutic evaluations in the current study were collected on a 1.5T MR scanner (Avanto or Sonata, Siemens Healthcare, Erlangen, Germany; Excite HDx or LX, GE Medical Systems, Waukesha, WI) and all dual echo TSE images had full brain coverage with TE₁ = 9-15ms, TE₂=118-135ms, TR = 4000ms, matrix size = 256x256, FOV = 240mm, and slice thickness = 3mm with no interslice gap.

**Definition of Disease Progression**

Progression was defined prospectively by the treating neuro-oncologists using a method consistent with the current RANO criteria (29). In particular, progression was defined if subsequent scans showed an increase in imaging-evaluable tumor (≥ 25% increase in the sum of enhancing
lesions, new enhancing lesions > 1 cm², an unequivocal qualitative increase in nonenhancing tumor, or an unequivocal new area of noncontrast enhancing tumor). Patients were required to have stable or decreasing contrast agent dose before partial or complete response could be determined. Additionally, patients requiring increased dosage of steroids in order to maintain neurologic function, even in the absence of worsening on anatomical images, were considered to be stable, but required early reevaluation. Patients who experienced significant neurologic decline were also declared to have progressed at the time of irreversible decline. Landmark progression-free survival (PFS) was therefore defined as being the number of days between the post-treatment MRI scan (e.g. post-radiation therapy or post-bevacizumab) and declared progression. Landmark overall survival (OS) was defined as the number of days between the post-treatment MRI scan to death.

Automated Segmentation and Quantification of Nonenhancing Tumor Burden

Multiple studies have demonstrated that the T₂ characteristics for nonenhancing tumor fall between T₂ measurements of NAWM and edema (i.e. T₂ eff \(_{\text{NAWM}} \) < T₂ eff \(_{\text{tumor}} \) < T₂ eff \(_{\text{edema}} \)) (16-21). The particular thresholds used for tumor segmentation were determined empirically from the 50 glioma patients as per the experiment described above. After defining these thresholds, a 3 step automated procedure was performed to segment regions of nonenhancing tumor from surrounding tissues (Supplemental Figure A): 1) First, the skull was stripped and all tissue with T₂ eff greater than the threshold to delineate tumor and NAWM were isolated, effectively removing NAWM and retain tumor, edema, and necrosis. 2) Next, one image voxel surrounding these regions was eroded in order to remove tissues on the border of edema and NAWM, for example,
where a partial volume voxel may have $T_2^{\text{eff}}$ characteristics similar to tumor. A whole brain mask was then applied to confine all regions of interest (ROIs) inside the brain. 3) Lastly, nonenhancing tumor was segmented from regions of edema within the ROIs using the empirical thresholds defined previously and a cluster-based algorithm using a single voxel nearest-neighbor algorithm was used to isolate the largest contiguous clusters with $T_2^{\text{eff}}$ values consistent with nonenhancing tumor within the brain. Cluster sizes greater than 0.1cc confluent with the primary enhancing lesion were retained for subsequent analyses. A single investigator confirmed adequate automatic segmentation, including verification that clusters were isolated in regions that could contain tumor.

**Statistical Analyses**

*Phantom Measurements.* In order to determine the stability of $T_2^{\text{eff}}$ measurements in the ACR phantom over time we examined the average coefficient of variance for 1.5T and 3T scanners independently over 3 years. A $t$-test was used to compare average $T_2^{\text{eff}}$ measurement values between 1.5T and 3T scanners.

*Training Data.* For comparisons of $T_2^{\text{eff}}$ measurements in various tissue types, a one-way analysis of variance (ANOVA) was used to examine $T_2^{\text{eff}}$ measurements in NET across tumor grade, and a two-way ANOVA including interaction terms along with Tukey’s test for multiple comparisons was used to examine differences in $T_2^{\text{eff}}$ measurements across tissue types and field strengths. Additionally, receiver-operator characteristic (ROC) analysis was used to determine the sensitivity and specificity for delineating various tissue types including NET from edema and NET from NAWM for specific $T_2^{\text{eff}}$ thresholds.
Newly Diagnosed Glioblastoma Treated with Radiotherapy. For newly diagnosed glioblastoma patients treated with radiation alone followed by adjuvant chemotherapy, ROC analysis was used to determine whether the volume of NET burden (selected from ROC analysis of training data) could stratify patients who progressed 6 months as well as those who expired within 9 months following completion of radiation therapy. Log-rank univariate analysis applied to Kaplan-Meier data was used to determine whether patients with a large NET burden (from ROC analysis) had significantly shorter PFS and OS. A Cox multivariate proportional hazards model was then used to determine whether NET volume quantified using $T_2^{\text{eff}}$ was an independent predictor of PFS and OS after including covariates including patient age and KPS.

Newly Diagnosed Glioblastoma Treated with Concurrent Radiotherapy and Temozolomide Followed by Adjuvant Temozolomide. For newly diagnosed glioblastoma patients treated with concurrent radiation and temozolomide followed by adjuvant temozolomide, ROC analysis was used to determine whether the volume of NET burden (selected from ROC analysis of training data) could stratify patients who progressed 6 months as well as those who expired within 9 months following completion of radiation therapy. Log-rank univariate analysis applied to Kaplan-Meier data was used to determine whether patients with a large NET burden (from ROC analysis) had significantly shorter PFS and OS. A Cox multivariate proportional hazards model was then used to confirm whether NET burden quantified using $T_2^{\text{eff}}$ was an independent predictor of PFS and OS after including covariates including patient age and KPS.

Recurrent Glioblastoma Treated with Bevacizumab. For recurrent glioblastoma patients treated with bevacizumab, ROC analysis was used to determine whether the volume of NET burden (selected from ROC analysis of training data) could stratify patients who progressed 6
months as well as those who expired within 9 months following start of bevacizumab. Log-rank univariate analysis applied to Kaplan-Meier data was used to determine whether small changes in NET burden (from ROC analysis) had significantly shorter PFS and OS. A Cox multivariate proportional hazards model was then used to confirm that change in NET burden before and after bevacizumab was an independent predictor of PFS and OS after including covariates including patient age and KPS.
RESULTS

Evaluations of the ACR phantoms using the dual echo TSE sequence showed reasonable stability over time, averaging a coefficient of variance of approximately 3.4% for 3T scanners and 4.5% for 1.5T scanners when evaluated over the last 3 years. The coefficient of variance of $T_2^{\text{eff}}$ measurements across the 3T scanners was 3.1% and across all 1.5T scanners was 4.5%, suggesting long-term stability was similar to the stability across scanners at a single time point. $T_2^{\text{eff}}$ measurements on 3T MR scanners were significantly lower than 1.5T scanners (Figure 2; t-test, $P<0.0001$; $T_2^{\text{eff}}$ in 3T = 79.3 ± 2.5 ms vs. 1.5T = 98.8 ± 4.4 ms), confirming the expected decrease in $T_2$ relaxation times for increasing magnetic field strength.

A total of 50 patients with various tumor grades containing regions of obvious NET as training data for delineation of NET from other tissues based on specific $T_2^{\text{eff}}$ measurements. In general, $T_2^{\text{eff}}$ maps appeared relatively similar across field strengths (Figure 3), with areas of obvious edema or necrosis showing a higher $T_2^{\text{eff}}$, areas of NAWM showing the lowest $T_2^{\text{eff}}$, and regions of suspected NET adjacent or contiguous with enhancing tumor showing $T_2^{\text{eff}}$ measurements between edema and NAWM. No differences in $T_2^{\text{eff}}$ measurements were observed in NET components across tumor grade (ANOVA, $P=0.8678$). Measurement of $T_2^{\text{eff}}$ in the various tissues confirmed this observation (Figure 4A). A two-way ANOVA showed a significant difference between tissue types ($P<0.0001$), field strength ($P<0.0001$), and the interaction between tissue types and field strength ($P<0.0001$). Multiple comparisons tests showed significant differences between $T_2^{\text{eff}}$ values for each tissue type within both 1.5T and 3T data (Tukey’s test, $P<0.01$ for all comparisons between tissue types within 1.5T and 3T). Within each tissue type, $T_2^{\text{eff}}$ values were significantly different across field strengths in edema and necrotic tissues ($P<0.01$), but no
difference was observed between $T_2^{\text{eff}}$ values within NAWM or tumor. This suggests a single threshold may be used across field strengths to differentiate NAWM and tumor from other tissues.

Receiver-operator characteristic (ROC) analysis suggested that $T_2^{\text{eff}}$ could significantly stratify NET from edema (Figure 4B; 1.5T: ROC AUC = 0.9000 ± 0.0219, $P < 0.0001$; 3T: AUC = 0.9773 ± 0.01199, $P<0.0001$). A threshold of 250ms could be used to stratify NET from edema with a sensitivity of 91% and specificity of 71% at 1.5T, and a sensitivity of 100% and specificity of 65% at 3T. ROC analysis suggested that $T_2^{\text{eff}}$ could also stratify NET from NAWM (Figure 4B; 1.5T: AUC= 0.9934 ± 0.0043, $P<0.0001$; 3T: AUC = 0.9974 ± 0.0030, $P<0.0001$). A threshold of 125ms could be used to stratify NET from NAWM with a sensitivity of sensitivity of 96% and specificity of 97% at 1.5T, and 94% sensitivity and 97% specificity at 3T. Thus, we propose using a range of $T_2^{\text{eff}}$ values between 125ms < $T_2^{\text{eff}}$ < 250ms, on both 1.5T and 3T MR scanners, to objectively define NET for the purpose of tumor burden quantification. Figure 4C shows a patient with a recurrent glioblastoma containing a large, well-circumscribed NET appearing as slightly $T_2$ hyperintense on T2-weighted images at 3T. $T_2^{\text{eff}}$ maps clearly show that this NET falls largely within the defined range of 125ms < $T_2^{\text{eff}}$ < 250ms. Conversely, Figure 4D also shows a patient with recurrent glioblastoma scanned at 3T with a large extent of edematous tissue, as demonstrated by the “finger like” extensions of very bright $T_2$ hyperintensity on T2-weighted images. Consistent with these qualitative observations, most of this lesion had elevated $T_2^{\text{eff}}$ within the range associated with edema. Additionally, an area near the right insula with relatively lower signal intensity on $T_2$-weighted images (red arrow) also containing contrast enhancement (not shown) appears to have $T_2^{\text{eff}}$ values within the range defined as NET. MR spec-
trosopic imaging (MRSI) also confirmed that areas defined as NET on $T_2^{\text{eff}}$ maps were often associated with increased choline concentration in a subset of patients (Figures 4E-F), suggestive of freely mobile phosphocholine from the cell membrane turnover in actively proliferating tumor cells.

Using $125\text{ms} < T_2^{\text{eff}} < 250\text{ms}$ to objectively define NET, we aimed to test whether the volume of NET burden following radiation therapy alone was a prognostic predictor in patients with newly diagnosed glioblastoma (Figure 5A-C). The average volume of NET burden for the cohort was estimated at $39.6\text{cc} \pm 7.2\text{cc}$ (S.E.M.) and a range from $0.3\text{cc}$ to $119\text{cc}$. Results suggest that NET volume measured following radiation therapy could identify patients who progressed within 6 months following completion of radiation therapy (Figure 5A; ROC AUC = $0.8590 \pm 0.07248$, $P = 0.002$) as well as those who expired within 9 months following completion of radiation therapy (Figure 5A; ROC AUC = $0.8467 \pm 0.08341$, $P = 0.004$). A threshold of $32\text{cc}$, the median volume of NET, was shown to have a sensitivity of 69% and specificity of 79% for identifying patients who progressed within 6 months of radiation therapy and a sensitivity of 70% and specificity of 90% for identifying patients who expired within 9 months of completion of radiation therapy. Patients with NET volumes larger than $32\text{cc}$ had a significantly shorter PFS (Figure 5B; Log-rank, $P=0.0013$, $HR=2.219$, median survival for high vs. low NET volume = 132 vs. 294 days) and OS (Figure 5C; Log-rank, $P = 0.0071$, $HR = 2.982$, median OS for high vs. low NET volume = 228 vs. 681 days). Cox multivariate regression including clinical covariates of both age and KPS confirmed that NET volume after radiation therapy was a significant predictor of subsequent PFS (Table 1; Cox proportional hazards, $P=0.0216$, $HR = 2.8410\pm1.5756$ S.E.M) and OS (Cox proportional hazards, $P=0.0014$, $HR = 5.3475\pm1.6925$ S.E.M).
Age at diagnosis was also a prognostic factor in newly diagnosed glioblastoma patients treated with radiation therapy (age: $PFS, P = 0.0132, \text{HR}=1.0461\pm1.0184 \text{ S.E.M.}; \text{OS}, P = 0.0028, \text{HR}=1.0631\pm1.0207$), but KPS was not predictive ($KPS: PFS, P = 0.4433; \text{OS}, P = 0.1526$).

Next, we aimed to test whether the volume of NET burden following radiation therapy combined with temozolomide followed by adjuvant temozolomide was a prognostic predictor in patients with newly diagnosed glioblastoma (Figure 5D-F). The average volume of NET burden for the cohort was estimated at 34.1cc ± 4.7cc (S.E.M.) and a range from 0.6cc to 122cc. Results suggest that NET volume evaluated following combined radiation therapy and temozolomide could not identify patients who progressed within 6 months following completion of radiochemotherapy ($Figure \ 5D; \text{ROC AUC} = 0.5906 \pm 0.0963, P = 0.3850$) or patients at risk for expiring within 9 months following completion of radiochemotherapy ($Figure \ 5D; \text{ROC AUC}=0.6733 \pm 0.0925, P = 0.08304$). Patients with NET volumes larger than the median NET volume of 23cc for this cohort following radiochemotherapy were at significantly higher risk for shorter PFS ($Figure \ 5E; \text{Log-rank}, P=0.035, \text{HR} = 2.217, \text{median PFS for high vs. low NET volume} = 67 \text{ vs. 149 days}$) and OS ($Figure \ 5F; \text{Log-rank}, P=0.0335, \text{HR} = 1.1660, \text{median OS for high vs. low NET volume} = 181 \text{ vs. 301 days}$). Cox multivariate regression including clinical covariates confirmed this trend (Table 2), showing that NET volume following completion of radiochemotherapy was a significant independent predictor of both PFS (Cox proportional hazards, $P=0.0323, \text{HR} = 2.3708\pm1.4690 \text{ S.E.M.}$) and OS (Cox proportional hazards, $P=0.0004, \text{HR} = 2.7907\pm1.5173 \text{ S.E.M.}$). Age and KPS at diagnosis were not prognostic factors for PFS or OS in
newly diagnosed glioblastoma patients treated with radiochemotherapy \( (age: \text{PFS}, P = 0.7711, \text{OS}, P = 0.3520; \text{KPS: PFS}, P = 0.2745; \text{OS}, P = 0.2450) \). Lastly, we tested whether early NET response measured as the change in NET volume before and after bevacizumab therapy could be used to predict survival in patients with recurrent glioblastoma \( (\text{Figure 5G-I}) \). The average volume of NET burden for the cohort prior to and after bevacizumab therapy was estimated at \( 30.3cc \pm 3.7cc \) (S.E.M.) and \( 15.8cc \pm 3.5cc \) (S.E.M.), respectively. The average change in NET volume was \( -14.5cc \pm 2.7cc \) (S.E.M.). Results demonstrated that changes in NET volume before and after bevacizumab therapy could be used to identify patients who progressed within 6 months of completion of the first dose of bevacizumab \( (\text{Figure 5G; ROC AUC} = 0.7786 \pm 0.0967, P = 0.022) \), as well as patients who expired within 9 months of completion of the first dose of bevacizumab \( (\text{Figure 5G; ROC AUC} = 0.8681 \pm 0.07399, P = 0.002) \). A decrease in NET volume greater than 50% was found to have a sensitivity of 70% and specificity of 71% for identifying patients that will progress within 6 months of starting bevacizumab therapy, along with a sensitivity of 83% and specificity of 67% for identifying patients that will expire within 9 months of starting bevacizumab therapy. Patients showing a “response”, or more than 50% decrease in NET volume after bevacizumab therapy, had a significantly longer PFS \( (\text{Figure 5H; Log-rank, } P=0.00070, HR = 2.565, \text{median PFS for responders vs. nonresponders} = 215 \text{ vs. 84 days}) \) and OS \( (\text{Figure 5I; Log-rank, } P=0.0105, HR = 2.00, \text{median OS for responders vs. nonresponders} = 375 \text{ vs. 187 days}) \). Cox multivariate regression confirmed this observation \( (\text{Table 3}) \), showing longer PFS \( (\text{Cox proportional hazards, } P=0.0071, HR = 1.0189\pm1.0070) \) and OS \( (\text{Cox proportional hazards, } P=0.0377, HR = 1.0122\pm1.0058) \) in patients demonstrating more than a 50% decrease in NET volume after the
first dose of bevacizumab. Both age and KPS were not significant prognostic factors for recurrent glioblastoma patients treated with bevacizumab (age: PFS, \( P = 0.1090 \), OS, \( P = 0.1220 \); KPS: PFS, \( P = 0.5283 \), OS, \( P = 0.7142 \)).
DISCUSSION

Results from the current study support the hypothesis that T₂ maps quickly obtained using dual echo TSE images can be used to objectively define and quantify NET burden. Results support the concept that quantification of NET is clinically meaningful and prognostic, as demonstrated both by the prognostic significance of NET tumor burden following radiation therapy with or without temozolomide as well as the change in NET volume as a simple response measure during bevacizumab therapy in recurrent glioblastoma. This is a particularly timely question as this concept has come into serious debate recently (30). In particular, studies have suggested limited value of including challenging and costly qualitative assessment of nonenhancing tumor response when evaluating drug efficacy, as response rates and PFS estimates seem comparable with and without the use of nonenhancing tumor assessment (1) and in most instances, nonenhancing tumor progression results in contrast enhancing tumor progression by the subsequent follow-up examination (31). Additionally, recent results from ACRIN 6677/RTOG 0625, a prospective, randomized, phase II multicenter trial comparing bevacizumab with either irinotecan or temozolomide in recurrent glioblastoma, demonstrated that response rate of contrast enhancing tumor was predictive of OS, whereas nonenhancing tumor progression rates did not predict OS (32). Regardless, a large proportion of patients are expected to experience nonenhancing tumor progression prior to changes in contrast enhancement and the vast majority of low-grade gliomas do not have contrast enhancement, reinforcing the need for an objective definition of nonenhancing tumor using quantitative MRI.

Although ROC analysis demonstrated that NET burden could be used to quickly identify newly diagnosed glioblastomas at high risk for progressing by 6 months or expiring before 9
months when treated with radiation therapy alone, this same time point was of limited predictive value in patients treated with radiation therapy and concurrent temozolomide, the current standard of care for newly diagnosed glioblastoma. These results may be influenced by other molecular characteristics of the tumor, such as MGMT promoter methylation status or gene expression subtype. Future studies with larger patient cohorts are necessary to identify the roles of these additional prognostic factors.

Results demonstrate that $T_2^{\text{eff}}$ measurements using dual echo TSE are relatively stable, varying approximately 3-5% over time and across scanners of similar field strength when evaluated in phantoms. This is consistent with previous reports showing variation in normal tissues of between 5-15% (23, 33, 34), on par with errors due to the specific pulse sequence and vendor specific errors (35). $T_2$ estimation using dual echo TSE techniques, however, has been shown to slightly overestimate $T_2$ by as much as 10% (26, 36-38). This bias in $T_2$ estimation arises from multiple sources, including the use of only a few echoes for estimation, spurious signals from stimulated echoes, radiofrequency pulse imperfections, and different phase encoding profile orders (26, 39, 40). Despite this inaccuracy, $T_2$ mapping using dual echo TSE acquisition has proven very useful for identification of various pathologies, including Alzheimer’s and other neurodegenerative diseases (23, 37, 41, 42). Techniques to overcome this slight bias have been developed (39), although they remain relatively cumbersome to implement and are not often used in clinical practice. Importantly, the Alzheimer’s disease neuroimaging initiative (ADNI) chose to implemented dual echo $T_2$-weighted TSE for estimation of pathology-specific changes in quantitative $T_2$; thus, a standardized dual echo TSE sequence is openly available from all major MRI vendors, which further supports the view of using dual echo T2-weighted TSE to quantify “ef-
fective $T_2$” ($T_2^{\text{eff}}$) for use in objectively defining NET burden within clinically realizable scan times.

It is important to point out that there remains significant overlap between $T_2^{\text{eff}}$ measurements within regions believed to contain mixtures of edema and NET, thus quantification of NET burden in the current study may not be completely accurate. During the training phase we chose to specifically identify regions containing obvious, pure NET with little partial volume contamination from edema or necrosis. It is both conceivable and highly likely that some regions classified as edema and not included in the quantification of NET burden contain active tumor, whereas some regions classified as NET may in fact contain substantial proportions of edema. Future studies aimed at histological validation should be performed in order to further aid in defining NET burden using $T_2^{\text{eff}}$ measurements.

Additional considerations and study limitations should also be addressed. For example, white matter changes after radiation therapy may have confounded our interpretation of various tissue types; however, we strategically included recurrent glioma patients in our training cohort in order to provide a more robust range of T2 values that might include treatment-related changes. Lastly, the current study lacks separate, independent validation of the primary findings. Future studies aimed at independent validation in a randomized clinical trial, along with comparison to other imaging biomarkers, is warranted in order to establish effective $T_2$ maps as a surrogate for quantifying NET burden in patients with gliomas.

CONCLUSION
Dual echo TSE can be used to objectively define NET burden for use in brain tumor characterization, prognosis, and response assessment. $T_2^{\text{eff}}$ measurements obtained using dual echo TSE has relatively low variability, is feasible at both 1.5T and 3T, and can be used to predict response and survival in patients with gliomas.
REFERENCES


Figure Captions:

Figure 1: Treatment paradigms for testing the prognostic and predictive utility of $T_2^{\text{eff}}$-defined nonenhancing tumor (NET) volumes. A) Newly diagnosed glioblastoma patients treated with radiation therapy were evaluated 10 weeks following the start of radiation. B) Newly diagnosed glioblastoma patients treated with concurrent radiation and temozolomide therapy followed by adjuvant temozolomide were evaluated 10 weeks following concurrent radiation and temozolomide, prior to start of adjuvant temozolomide. C) Recurrent glioblastoma patients treated with bevacizumab were evaluated before and after the first dose of bevacizumab therapy.

Figure 2: Dual echo turbo spin echo (TSE) estimates of effective $T_2$ ($T_2^{\text{eff}}$) obtained in an American College of Radiology (ACR) phantom for 16 different MRI scanners. Results show a significant difference in $T_2^{\text{eff}}$ measurements in 3T (mean = 79.3ms) versus 1.5T (mean = 98.8ms) scanners.

Figure 3: Post-contrast T1-weighted images, dual echo TSE images, and $T_2^{\text{eff}}$ maps of five patients with glioblastoma scanned on different MRI scanners with slightly different acquisition parameters.

Figure 4: Training data used for delineating nonenhancing tumor (NET) from other tissues. A) Mean $T_2^{\text{eff}}$ measurements for normal-appearing white matter (NAWM), nonenhancing tumor (NET), edema, and necrosis for both 1.5T and 3T. Error bars represent standard error about the mean (S.E.M.) across all 25 patients scanned at either 1.5T or 3T. B) Receiver operator
characteristic (ROC) curves showing delineation between edema and necrosis, NET and edema, and NET and NAWM for both 1.5T and 3T. C) An example of a glioblastoma patient with a large NET mass as evidenced by the well-circumscribed, slightly T2-hyperintense lesion adjacent to the resection cavity. $T_2^{\text{eff}}$ maps show that this mass falls within the range defined as NET in the current study ($125 \text{ms} < T_2^{\text{eff}} < 250 \text{ms}$). D) An example of a glioblastoma patient with a large T2-hyperintense lesion consistent with a mostly edematous mass as evidenced by the high T2-hyperintense signal and “finger like” white matter extensions. An area consistent with $T_2^{\text{eff}}$-defined NET is shown (red arrow), which also is colocalized to the area of contrast enhancement (not shown). E) A patient with a low-grade glioma shown to have a NET mass within the splenium of the corpus callosum on $T_2^{\text{eff}}$ maps. F) MR spectroscopic imaging (MRSI) choline metabol-ic maps show high choline concentration in this area, consistent with metabolically active tumor.

Figure 5: Receiver-operator characteristic (ROC) curves, progression-free (PFS), and overall survival (OS) using $T_2^{\text{eff}}$-defined nonenhancing tumor (NET) volumes in newly diagnosted glioblastoma patients treated with radiation therapy or radiation therapy and concurrent temozolomide, and recurrent glioblastoma treated with bevacizumab. A) ROC analysis performed using NET volume estimates following radiation therapy to predict 6 month PFS (PFS6) and 9 month OS (OS9). B) Kaplan-Meier curves showing that the median volume of NET (32cc) could stratify long and short-term PFS in newly diagnosed glioblastoma patients treated with radiation therapy ($P=0.0013$). C) Kaplan-Meier curves showing the median volume of NET (32cc) could also stratify long and short-term OS in newly diagnosed glioblastoma pa-tients treated with radiation therapy ($P=0.0071$). D) ROC analysis performed using NET volume
estimates following concurrent radiation therapy and temozolomide to predict PFS6 and OS9. Results did not show statistical significance for either PFS6 or OS9 in this patient cohort. E) Kaplan-Meier curves illustrating that the median volume of NET (23cc) could stratify long and short-term PFS in newly diagnosed glioblastoma patients treated with concurrent radiation therapy and temozolomide followed by adjuvant temozolomide (\(P=0.0350\)). F) Kaplan-Meier curves showing that median NET volume (23cc) could stratify long and short-term OS (\(P=0.0335\)). G) ROC analysis performed using the change in NET volume before and after bevacizumab treatment in recurrent glioblastoma to predict PFS6 and OS9. H) Kaplan-Meier curves demonstrating that recurrent glioblastoma patients with more than 50% decrease in NET volume following bevacizumab therapy (“responders”) have a longer PFS compared with patients demonstrating less than a 50% decrease in NET volume (“nonresponders”) (\(P=0.0070\)). I) Kaplan-Meier curves showing that recurrent glioblastoma patients showing “response” to bevacizumab on NET have a significantly lower OS compared with “nonresponders” (\(P=0.0105\)).
Table 1: Cox Proportional Hazards Model Results for Newly Diagnosed Glioblastoma Treated with Radiation Therapy. Error bars represent ± S.E.M.

<table>
<thead>
<tr>
<th></th>
<th>PFS Hazard Ratio</th>
<th>OS Hazard Ratio</th>
<th>P-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.0461 ± 1.0184</td>
<td>1.0631 ± 1.0207</td>
<td>0.0132</td>
<td>0.0028</td>
</tr>
<tr>
<td>KPS</td>
<td>1.0184 ± 1.0240</td>
<td>0.9624 ± 1.0271</td>
<td>0.4433</td>
<td>0.1526</td>
</tr>
<tr>
<td>T&lt;sub&gt;2&lt;/sub&gt;&lt;sup&gt;eff&lt;/sup&gt; Defined NET &gt; 32cc</td>
<td>2.8410 ± 1.5756</td>
<td>5.3475 ± 1.6925</td>
<td>0.0216</td>
<td>0.0014</td>
</tr>
</tbody>
</table>

Table 2: Cox Proportional Hazards Model Results for Newly Diagnosed Glioblastoma Treated with Concurrent Radiation Therapy and Temozolomide. Error bars represent ± S.E.M.

<table>
<thead>
<tr>
<th></th>
<th>PFS Hazard Ratio</th>
<th>OS Hazard Ratio</th>
<th>P-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.9939 ± 1.0211</td>
<td>1.0251 ± 1.0270</td>
<td>0.7711</td>
<td>0.3520</td>
</tr>
<tr>
<td>KPS</td>
<td>1.0227 ± 1.0207</td>
<td>1.0270 ± 1.0232</td>
<td>0.2745</td>
<td>0.2450</td>
</tr>
<tr>
<td>T&lt;sub&gt;2&lt;/sub&gt;&lt;sup&gt;eff&lt;/sup&gt; Defined NET &gt; 23cc</td>
<td>2.3708 ± 1.4969</td>
<td>2.7907 ± 1.5173</td>
<td>0.0323</td>
<td>0.0138</td>
</tr>
</tbody>
</table>

Table 3: Cox Proportional Hazards Model Results for Recurrent Glioblastoma Treated with Bevacizumab. Error bars represent ± S.E.M.

<table>
<thead>
<tr>
<th></th>
<th>PFS Hazard Ratio</th>
<th>OS Hazard Ratio</th>
<th>P-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.9703 ± 1.0190</td>
<td>0.9746 ± 1.0168</td>
<td>0.1090</td>
<td>0.1220</td>
</tr>
<tr>
<td>KPS</td>
<td>0.9840 ± 1.0259</td>
<td>0.9909 ± 1.0254</td>
<td>0.5283</td>
<td>0.7142</td>
</tr>
<tr>
<td>Reduction in NET &gt; 50%</td>
<td>1.0189 ± 1.0070</td>
<td>1.0122 ± 1.0058</td>
<td>0.0071</td>
<td>0.0377</td>
</tr>
</tbody>
</table>
Figure 1
Figure 2

**P < 0.0001**

- 3T Siemens Skyra #1
- 3T Siemens Skyra #2
- 3T Siemens Skyra #3
- 3T Siemens Trio #1
- 3T Siemens Verio #1
- 3T Siemens Verio #2
- 3T Siemens Allegra
- 1.5T Siemens Avanto #1
- 1.5T Siemens Avanto #2
- 1.5T Siemens Avanto #3
- 1.5T Siemens Sonata #1
- 1.5T Siemens Sonata HDX
- 1.5T GE Signa LX

**T2_{eff} [ms]**

- 79.3 ms
- 98.8 ms

**T2_{eff} [ms]**

- 0
- 50
- 100
- 150
Figure 4
Figure 5

RADIOThERAPY IN NEW GBM

(A) ROC Analysis

(B) Progression-Free Survival
- Volume < 32cc (Median)
- Volume > 32cc

P=0.0013** (Mantel-Cox)

(C) Overall Survival
- Volume < 32cc
- Volume > 32cc

P=0.0071** (Mantel-Cox)

RADOThERAPY + CONCURRENT TEMOZOLOMIDE IN NEW GBM

(D) ROC Analysis

(E) Progression-Free Survival
- Volume < 23cc (Median)
- Volume > 23cc

P=0.0350* (Mantel-Cox)

(F) Overall Survival
- Volume < 23cc
- Volume > 23cc

P=0.0335* (Mantel-Cox)

BEVACIZUMAB IN RECURRENT GBM

(G) ROC Analysis

(H) Progression-Free Survival
- Non-Responders
- Responders

P=0.0070** (Mantel-Cox)

(I) Overall Survival
- Non-Responders
- Responders

P=0.0105* (Mantel-Cox)
Quantification of nonenhancing tumor burden in gliomas using effective T2 maps derived from dual-echo turbo spin-echo MRI


Clin Cancer Res  Published OnlineFirst April 21, 2015.

Updated version  Access the most recent version of this article at: doi:10.1158/1078-0432.CCR-14-2862

Supplementary Material  Access the most recent supplemental material at: http://clincancerres.aacrjournals.org/content/suppl/2015/04/22/1078-0432.CCR-14-2862.DC1

Author Manuscript  Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.