

Comparison of Voxel-Wise and Histogram Analyses of Glioma ADC Maps for Prediction of Early Therapeutic Change

Thomas L. Chenevert¹, Dariya I. Malyarenko¹, Craig J. Galbán¹, Diana M. Gomez-Hassan¹, Pia C. Sundgren², Christina I. Tsien³, and Brian D. Ross¹

¹Department of Radiology, University of Michigan Medical School, Ann Arbor, MI; ²Department of Clinical Sciences/Radiology Lund University, Lund, Sweden; and

³Department of Radiation Oncology, Washington University School of Medicine, St. Louis, MO

Corresponding Author:

Thomas L. Chenevert, PhD

University of Michigan Hospitals, 1500 E. Medical Center Dr.,

UHB2 Room A209; Ann Arbor, MI 48109-5030;

E-mail: tlchenev@med.umich.edu.

Key Words: glioma therapy response, apparent diffusion coefficient, functional diffusion map, voxel-wise analysis, quantitative response metric

Abbreviations: Apparent diffusion coefficient (ADC), functional diffusion mapping (fDM), diffusion-weighted imaging (DWI), parametric response map (PRM), region of interest (ROI), volume of interest (VOI), Kaplan–Meier (KM), pretreatment (preTx), midtreatment (midTx), cumulative distribution function (CDF), T1-weighted with gadolinium enhancement (T1Gd), number of averages (NAV), Meta-image Header (MHD), area under (AUC) receiver operating curve (ROC) (AU-ROC), mean cumulative probability difference (mCPD), confidence interval (CI)

ABSTRACT

Noninvasive imaging methods are sought to objectively predict early response to therapy for high-grade glioma tumors. Quantitative metrics derived from diffusion-weighted imaging, such as apparent diffusion coefficient (ADC), have previously shown promise when used in combination with voxel-based analysis reflecting regional changes. The functional diffusion mapping (fDM) metric is hypothesized to be associated with volume of tumor exhibiting an increasing ADC owing to effective therapeutic action. In this work, the reference fDM-predicted survival (from previous study) for 3 weeks from treatment initiation (midtreatment) is compared to multiple histogram-based metrics using Kaplan–Meier estimator for 80 glioma patients stratified to responders and nonresponders based on the population median value for the given metric. The ADC histogram metric reflecting reduction in midtreatment volume of solid tumor ($ADC < 1.25 \times 10^{-3} \text{ mm}^2/\text{s}$) by $>8\%$ population-median with respect to pretreatment is found to have the same predictive power as the reference fDM of increasing midtreatment ADC volume above 4%. This study establishes the level of correlation between fDM increase and low-ADC tumor volume shrinkage for prediction of early response to radiation therapy in patients with glioma malignancies.

INTRODUCTION

Clinical oncology trials actively seek robust radiological markers of early response to cancer therapy to noninvasively guide patient treatment plans. By measuring water mobility known to be altered by tissue cellular constituents (1–3), diffusion-weighted imaging (DWI) is able to provide information on changes in tumor cellular density related to cytotoxic therapy response (4–7). Growth of viable tumor leads to increased cell density and reduced water mobility, while effective therapy decreases cell density and increases water mobility. Higher water mobility independent of therapy is also observed for necrotic tissue (8, 9). DWI measurements are typically represented as quantitative parametric diffusion maps of the apparent diffusion coefficient (ADC) based on an assumed monoexponential DWI signal decay with increasing diffusion-weighting strength (denoted by *b*-value) (5–7, 10). The therapy-related changes in the ADC maps can be quantitatively characterized spatially by the functional diffusion map (fDM) method within the general class

of parametric response mapping (PRM). These approaches deal with tumor heterogeneity to display significant regional change of treatment responsive/resistant voxels, while supplying a global quantitative response metric (11–13). PRM fDM has been shown to allow earlier prediction of glioma therapy response and more accurate prediction of survival relative to conventional neuroimaging metric (12). To provide robust alternative to invasive biopsies, the predictive power of this promising method needs to be linked to changes in tumor histopathological properties.

The fDM method (13) generally requires robust spatial registration of tumor volumes between longitudinal scans, which is potentially dependent on specific registration algorithm parameters and thus may be prone to introducing additional repeatability errors due to variation in image registration workflow. The method also relies on precise tumor region/volume-of-interest (ROI/VOI) definition and on matching voxels during potentially rapid tumor growth or shrinkage. By virtue of the

underlying statistical assumptions (14), fDM analysis includes thresholding for significant change, which can be nonspecific to the ADC range and tumor density as was originally proposed in (13). Notwithstanding demonstrated promising predictive value of the fDM metrics (11, 12), its direct relation to the biophysical properties of dense versus necrotic tumor volumes has not yet been clearly established. In principle, significant changes of fDM may occur over the full range of ADC values (both for restricted and less restricted diffusion (1)).

An alternative approach that forfeits retention of spatial origin of voxels within tumor is to perform histogram analysis of ADC voxel values (6, 15). Intralesion heterogeneity is retained by the histogram, although direct spatial identification of responsive/resistant regions is lost. The histogram analysis approach has several benefits. First, this approach removes dependence on technical performance of an image volume registration step, as well as assumptions that regions of rapid tumor growth/shrinkage are adequately coregistered. Second, the ADC histogram inherently facilitates segmentation of tumor based on tissue density reflected by water mobility (6). Third, this also allows direct identification of naturally high water mobility within cystic necrotic tumor tissue present before initiation of treatment to potentially distinguish from additional necrosis (9) resultant from cytotoxic treatment.

The purpose of the present study was to evaluate predictive power of several histogram-based ADC metrics and their correlation to fDM using quantitative DWI data from a common cohort of patients with glioma treated by chemoradiation. Because the overall objective was a technical comparison of the metrics, image processing and image segmentation were held constant across metrics derivation, and “survival” was used as the sole clinical outcome.

METHODOLOGY

This study analyzed Kaplan–Meier (KM) survival prediction for multiple ADC histogram metrics versus reference fDM-derived from quantitative DWI data including pretreatment (preTx) and 3-week midtreatment (midTx) imaging of a cohort of patients with high-grade glioma that underwent chemoradiotherapy treatment with longitudinal radiological surveillance (12). The baseline preTx scan was acquired postsurgery/biopsy before the start of treatment. The survival was assessed from the time of the diagnosis. All quantitative DWI and statistical analysis was performed using home-built routines developed in MATLAB 7 (MathWorks, Natick, MA). KM estimate of cumulative distribution function (CDF) for survival probability was generated using MATLAB built-in “*ecdf*” routine. The KM stair-step graphs for CDF censoring visualization were generated using MATLAB Central “*MatSurv*” function (16).

Patient Cohort

Details on patient cohort, treatment schedule, and diffusion scans are previously reported (12). Informed consent for images and medical record use for research was approved by institutional review board and renewed over the study period from 2000 to 2011. In total, 25 additional consented study subjects (scanned between 2007 and 2011) with grade 3 and 4 primary brain tumors were included into the present analysis and were

added to the 60 previously analyzed (2000 to 2006) (12). Overall patient demographics, pathology grade, treatment plans, response status, and imaging schedule were not significantly different from the original study and are not detailed here. Both patient survival (median months, 13.7 and 14.5) and pathology grade (3-to-4 ratios, 28% and 25%) were consistent between acquisition-date subgroups (Student’s *t*-test, $P > .7$), ensuring nominally unbiased clinical outcome measures of the combined group. Only preTx and 3-week midTx imaging were included in this study owing to previously demonstrated relevance for early response survival prediction by fDM (12). Only survival was used and no other clinical outcomes such as time-to-progression were considered.

Imaging Studies

Clinical MRI scans including quantitative diffusion MRI and standard MRI (fluid attenuation inversion recovery, T2-weighted, and T1-weighted with gadolinium enhancement [T1Gd] and without Gd enhancement) were performed for all imaging endpoints on 1.5 T MRI system (General Electric, Waukesha, WI; $n = 45$ patients) and on 3 T MRI scanner (Philips, Best, The Netherlands; $n = 40$ patients). The 75% of the initial (2000–2006) study scans were performed on 1.5 T, while 3 T scanner system was used exclusively for the (2007–2011) study subgroup. Consistent with the nominal independence on the acquisition-date, survival and pathology grade were not biased by the scanner subgroups ($P > .3$).

DWI protocol prescribed single-shot echo-planar imaging acquisition of three orthogonal-axial DWI scans with b -values = 0 and 1000 s/mm^2 using a 16-channel head-coil. On the 1.5 T system, 24 6-mm axial-oblique sections were acquired using a 22-cm field of view and 128 matrix (voxel size = 17.7 mm^3) repetition time = 10 000 ms; echo time = 71 to 100 ms, and number of averages (NAV) = 1. On the 3 T system, at least 28 4-mm axial-oblique sections were acquired through the brain using a 24-cm field of view and 128 matrix (voxel size = 14 mm^3 ; repetition time = 2.636 milliseconds; TE = 46 ms; NAV = 1 for $b = 0$, and NAV = 2 for $b = 1000 s/mm^2$). Parallel imaging (sensitivity-encoding factor = 3) was used at 3 T to reduce spatial distortion. PreTx and midTx scans for a given patient were performed on the same system.

ADC Parametric Map Generation

The diffusion images for the three orthogonal directions were combined into trace DWI to calculate an ADC map. All acquired data were stored and distributed in Digital Image Communication in Medicine (DICOM) format (17). ADC was fit as a slope of log-signal DWI as a function of b -value up to $b_{max} = 1000 s/mm^2$. For previously published data subset (12), image registration volumes and tumor segmentations were reused from prior analysis. For additional study subjects, the resulting low b -value, high b -value, and ADC maps were exported as Meta-image Header (MHD) format (18) for volumetric spatial registration to the anatomical pretreatment T1Gd images using the Elastix toolkit (19) with full-affine transformation. The low b -value DWI volume was used to drive image registration using the mutual information figure of merit, and the resultant spatial transformation was automatically applied to the corresponding high b -value and ADC volumes. Tumor-encompassing ROIs pre-

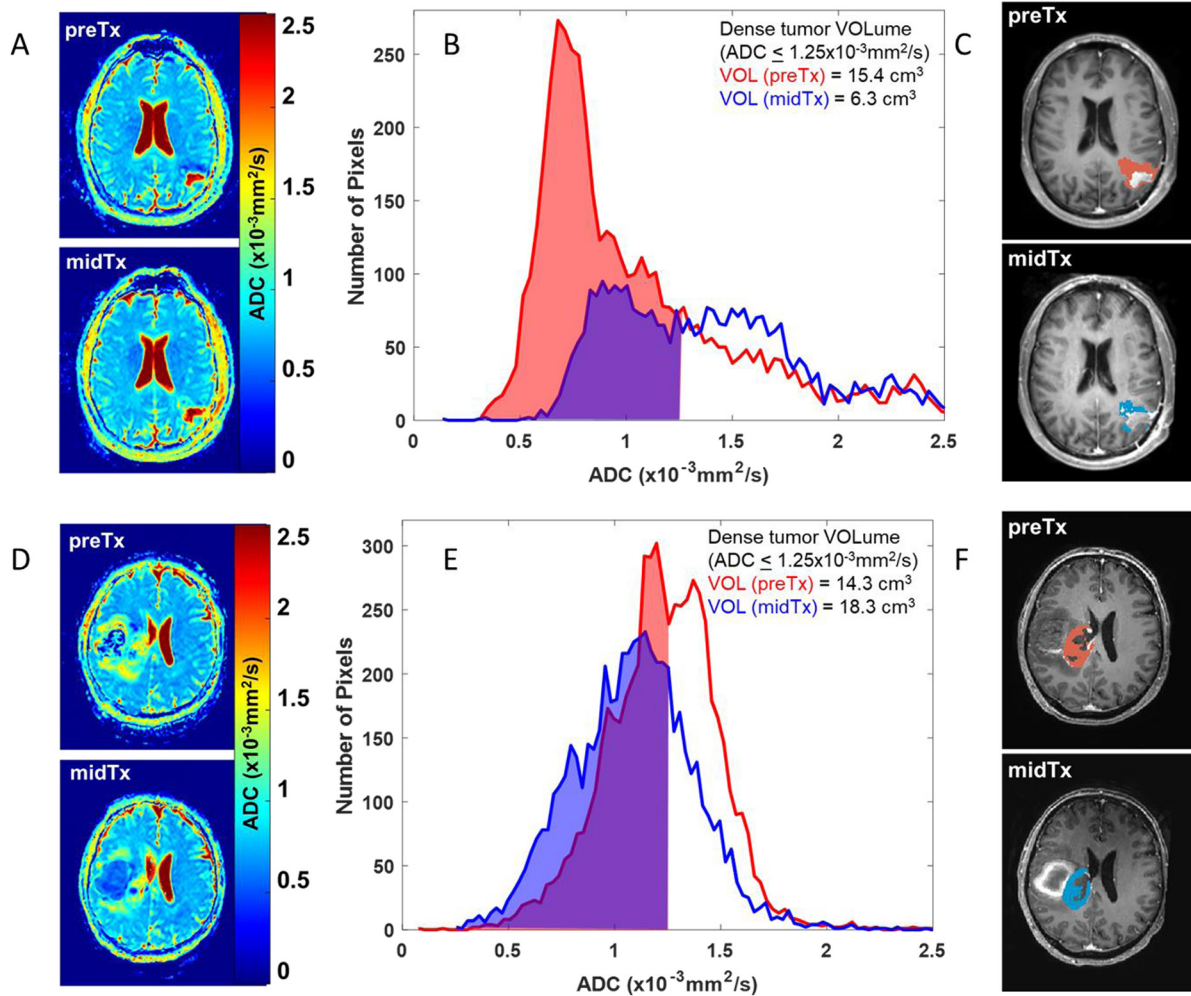


Figure 1. Left vertically arranged images (A, D) show ADC maps for preTx and midTx imaging time-points of 2 patients with glioma that responded favorably (A) and did not responded (D) to chemoradiation therapy. Common scale for the ADC maps is indicated by the color bar. The center panes (B, E) illustrate the corresponding tumor volume ADC histograms (preTx: red, and midTx: blue) and tumor voxel volumes (filled) below ADC threshold of $1.25 \times 10^{-3} \text{ mm}^2/\text{s}$. The corresponding integrated volumes of the dense tumor are listed in the legend. The spatial location of thresholded histogram voxels is overlaid in red and blue on a single representative slice of each patient preTx and midTx T1Gd images on the right in (C, F), used as a reference for tumor ROI definition.

viously defined by two experienced (>20 years) radiologists on the T1Gd images (coregistered to ADC maps) were imported into 3D Slicer (20) and converted to MHD ROI labels. These MHD ROI masks were then imported to MATLAB and applied to ADC maps to generate histograms of voxel ADC values within the defined tumor ROI (Figure 1). Additional ROIs (median volume, 5.4 cm^3 ; range, $3.6\text{--}7.6 \text{ cm}^3$) were defined on 3 slices for frontal normal-appearing white matter (contralateral to tumor) to confirm negligible system-specific ADC bias (21, 22) in two scanner subgroups [median ADC ($\times 10^{-3} \text{ mm}^2/\text{s}$): 0.785 (1.5 T) and 0.789 (3 T); $P = .19$].

ADC Histogram Metrics

Histogram “volume” metrics (in cubic centimeter units) were generated by numerically integrating the voxels up to specified ADC thresholds (without reference to spatial location other than being within the specified tumor ROI) and multiplying by the

known image voxel volume. The upper thresholds for low-ADC histogram portion (presumably reflecting more cellular-dense tumor) were sampled from 0.25 to 1.5 in steps of 0.25 ($\times 10^{-3} \text{ mm}^2/\text{s}$). The upper sampling bound of 1.5 ($\times 10^{-3} \text{ mm}^2/\text{s}$) was set to the previously published ADC value for necrotic tumor tissue (8). The standard whole-tumor histograms metrics, including ADC mean, median, and standard deviation were likewise evaluated for preTx and midTx imaging points separately and for their fraction-change with respect to preTx. The thresholds for survival-based therapy response prediction of each ADC histogram metric were dichotomized by population median values.

fDM Reference Metrics and KM Analysis

fDM analysis was performed as previously described (12). Only voxels present both in preTx and midTx tumor ROIs were stratified according to their change in ADC value (Figure 2, A and B)

into significantly increased (Vi, red, ADC change $> 0.55 \times 10^{-3} \text{ mm}^2/\text{s}$), decreased (Vd, blue, $< 0.55 \times 10^{-3} \text{ mm}^2/\text{s}$), and the remainder unchanged (Vo, green, within the $0.55 \times 10^{-3} \text{ mm}^2/\text{s}$ 95% confidence interval [CI]). The total percentage of tumor with significant increase in diffusion value was calculated as $100\% \times Vi/(Vi + Vo + Vd)$ and used as the reference fDM biomarker.

The KM survival probability analysis was then performed for the choice metrics with predetermined (population-median)

thresholds and the corresponding log-rank *P*-values (P_{KM}). Median fDM threshold was $Vi > 4\%$ ($P_{KM} = 0.0008$; Figure 2C; magenta KM line), which reasonably agreed with the optimized fDM threshold of 4.7% from the previous study (12) corresponding to maximum area under (AUC) receiver operating curve (ROC). Note that compared to the typical stair-step graphical representation (Figure 2C), the actual KM CDF curves would terminate before the last “stair-step” to exclude (unchanging) probability from the last censored patients (eg,

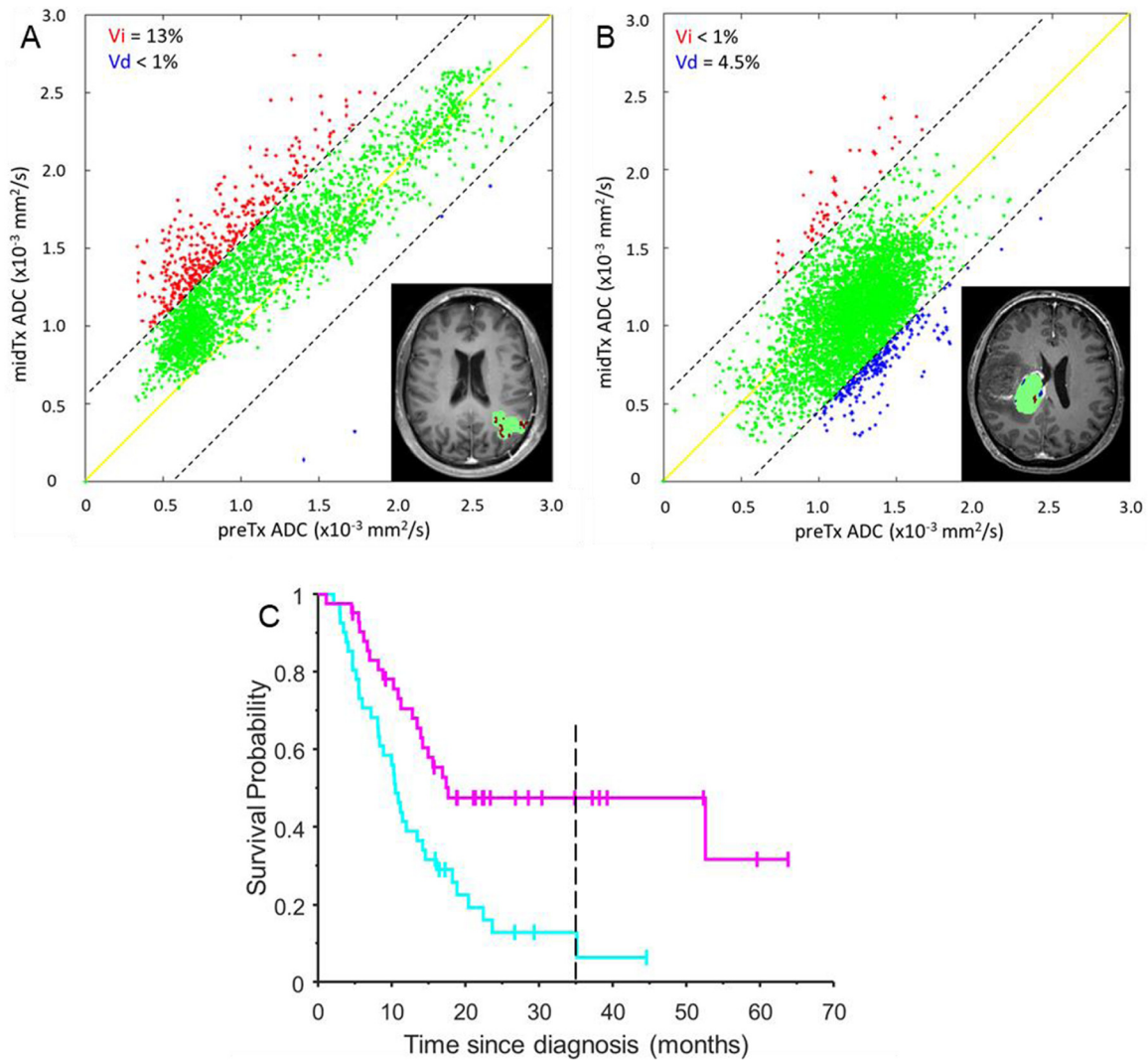


Figure 2. fDM metrics determined from midTx versus preTx ADC PRM scatter plots is overlaid on the T1Gd image inserts for the same two patients [responder (A) and nonresponder (B)] as in Figure 1 histograms. The dashed diagonal lines indicate 95% CI for the change encompassing green voxels corresponding to tumor regions not altered by therapy. The solid yellow line corresponds to the perfect fDM correlation. Red and blue areas mark tumor voxels with respective significant increase and decrease in ADC midTx versus preTx (summarized in the legends). (C) shows stair-step graph for reference fDM KM survival analysis of responders (magenta) and nonresponders (cyan) based on a median response threshold of 4% fDM-increase (magenta KM stair-step trend) for the whole glioma study population. Magenta and cyan KM trends correspond to the tumor fDM, respectively, above and below median response threshold. Vertical tick-marks along KM trends indicate individual patients whose survival times have been censored. Dashed vertical line corresponds to the minimal survival time included into the corresponding KM cumulative distribution function (CDF) probability analysis (excluding survival for the late censored patients).

at minimum CDF probability values of 0.07 and 0.3 for Figure 2C cyan and magenta trends, respectively).

Predictive power of each KM estimator was quantified by the mean cumulative probability difference (mCPD) between KM CDF curves (0.21 for reference fDM in Figure 2C). The KM curves for each sampled ADC metric were linearly interpolated to the common time-since-diagnosis axis corresponding to the fDM reference. The time-dependent survival probability differences between KM responder and nonresponder curves were correlated to that of the fDM reference to determine metrics with maximum KM “alignment” to the fDM. Pearson correlation, R_{fDM} , with $P_R < .05$ was considered significant. KM-length was determined as the minimal length of the two survival CDF curves for each metric. Similarity index was assessed by product of R_{fDM} and KM-length ratio, L_R , with respect to the fDM non-responder reference (Figure 2C; vertical dashed line marks the end of the corresponding CDF at 35 months).

RESULTS

Figure 1 illustrates ADC histogram analysis for the representative responder and nonresponder tumors using a low-ADC volume threshold of $1.25 \times 10^{-3} \text{ mm}^2/\text{s}$ (ie, only counting voxels within VOI having an ADC below this value) to favor inclusion of dense tumor while excluding necrotic regions. The corresponding ADC maps (Figure 1, A and D) depict quantitative regional diffusion changes in response to therapy, more pronounced for the responder (Figure 1, A–C) (survival, >27 months), relative to the nonresponder in Figure 1, D–F (survival, <9 months). The low ADC tumor component between midTx and preTx is quantified by a 9 cm^3 decrease of integrated dense tumor volume for the responder (Figure 1B) versus a 4 cm^3 increase for nonresponder (Figure 1E). That is, the fractional change in the low-ADC component of the histogram (59% decrease) owing to an upward shift, and shape change is enhanced by exclusion of the high ADC contribution that attenuates whole-tumor volumetric change (32% decrease) and whole-tumor mean ADC (30% increase). The low-ADC histogram voxel overlays on T1Gd images (Figure 1, C and F) further illustrate

how influence of the preexisting necrotic portion of the tumor is reduced by this analysis. Conversely, the nonresponder had an increase in dense tumor volume (by +28%) despite a reduction in whole-tumor volume (–6%). Although only central-tumor slices are shown in Figure 1, the histogram VOI analysis included all tumor slices.

Figure 2 illustrates fDM analysis for the same 2 subjects with diagnostic changes related to tumor response metrics (Figure 2A: $V_i = 13\%$, red, and Figure 2B: $V_d = 4.5\%$ blue voxels) observed predominantly toward lower ADC values ($<1.5 \times 10^{-3} \text{ mm}^2/\text{s}$). The red or blue fDM voxels marking regions with respective significant increase or decrease in ADC are evidently clustered in the lower half of midTx versus preTx values for a responder (Figure 2A, red) and nonresponder (Figure 2B, blue). The voxels with significantly higher midTx ADC for responder are distributed more uniformly across the ADC range of dense and necrotic tumor ($[1.25 - 2.25] \times 10^{-3} \text{ mm}^2/\text{s}$). However, the necrotic portion of the tumor does not significantly contribute to V_i in fDM analysis owing to high baseline ADC. Much lower red fDM volume shifted toward higher (necrotic) midTx ADC ($>1.5 \times 10^{-3} \text{ mm}^2/\text{s}$) is observed for nonresponder in Figure 2B with a noticeable increase in blue fDM voxel areas corresponding to lower (dense-tumor) ADC ($<1.25 \times 10^{-3} \text{ mm}^2/\text{s}$) for midTx. As in Figure 1, fDM difference overlays are on a single slice (Figure 2, inserts), whereas the fDM analysis spans the full tumor volume.

The responder versus nonresponder KM thresholds for the select test histogram characteristics based on population-wise median values are summarized in Table 1 along with their KM mCPD and percent-similarity index to the fDM CDF reference (Figure 2C). These median thresholds were used for the corresponding KM survival analysis shown in Figure 3. Other histogram metrics (not included) has shown <50% absolute similarity to fDM KM reference. Low predictive power was observed for all preTx metrics (median response threshold, $P_{KM} > .1$, mCPD < 0.06), reflecting dependence of response on the therapy administration. As expected, the corresponding KM CDF (Figure 3, A, D, and G) have shown low absolute similarity (<35%) to refer-

Table 1. Population-wise Median KM Response-Threshold, mCPD, and Similarity to Reference KM fDM for Select ADC Histogram Metrics

Metric	Median KM Threshold (P_{KM}^a)	mCPD	Similarity Index (%)
preTx Mean ADC ($10^{-3} \text{ mm}^2/\text{s}$)	1.19 (0.36)	0.06	20
midTx Mean ADC ($10^{-3} \text{ mm}^2/\text{s}$)	1.25 (0.0033)	0.2	13
% Change ^b Mean ADC	1.83 (0.05)	0.17	51
preTx Volume (cm^3)	32.5 (0.75)	0.05	35
midTx Volume (cm^3)	27.6 (0.38)	0.1	13
% Change ^b Volume	–0.8 (0.011)	0.18	–87
preTx LowADC Vol ^c (cm^3)	17.6 (0.51)	0.04	–18.6
midTx LowADC Vol ^c (cm^3)	15 (0.047)	0.14	–86
% Change ^b LowADC Vol ^b	–7.8 (0.0006)	0.22	–92.5

^a P -value of population-wise median KM response-threshold.

^b % Change = $100\% (\text{midTx} - \text{preTx})/\text{preTx}$.

^c Volume of tumor with $\text{ADC} \leq 1.25 \times 10^{-3} \text{ mm}^2/\text{s}$.

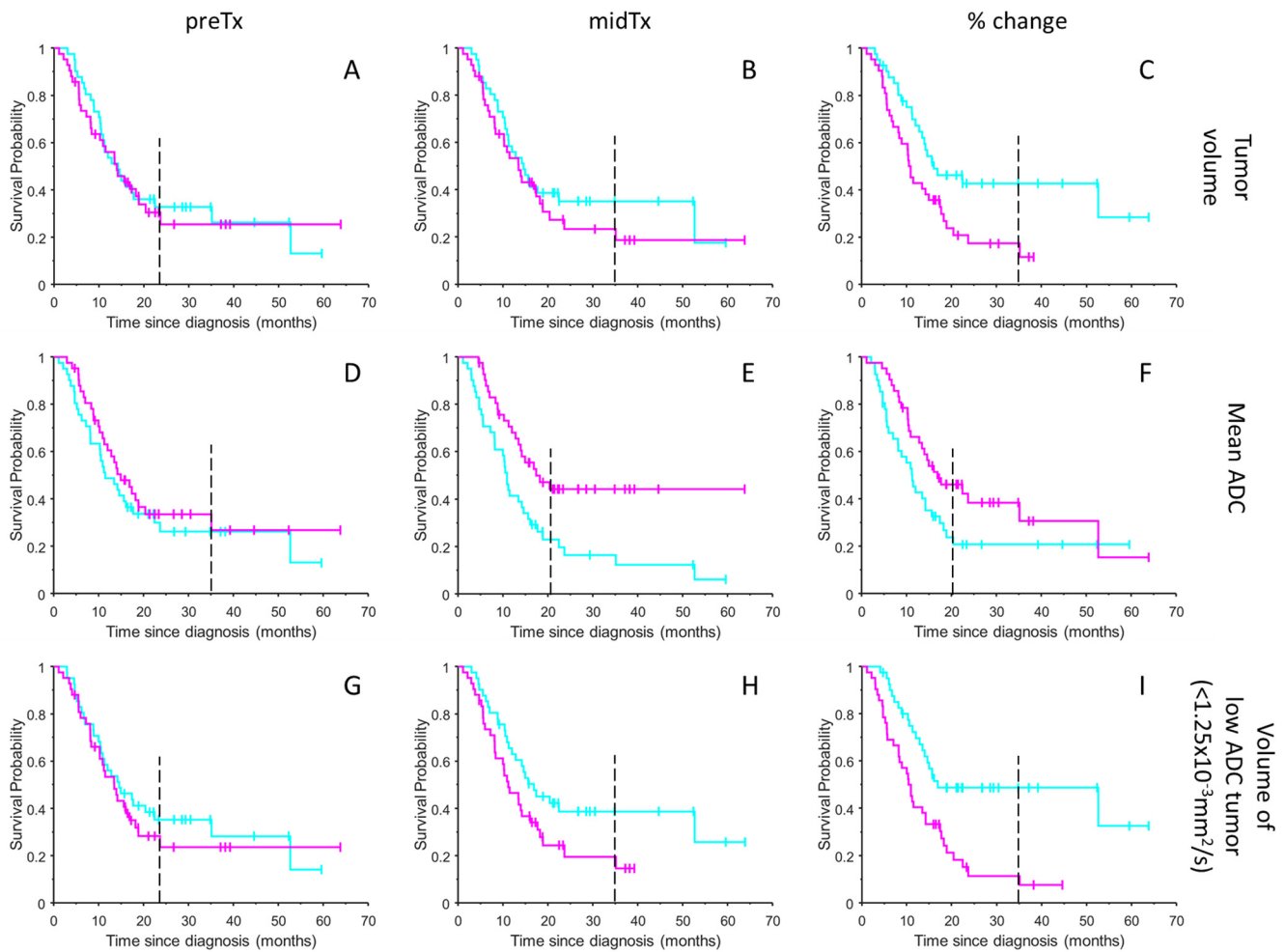


Figure 3. KM survival probability analysis results are summarized as stair-step graphs for conventional histogram metrics of total T1Gd tumor volume in (A–C), mean ADC in (D–F), and low ADC ($<1.25 \times 10^{-3} \text{ mm}^2/\text{s}$) histogram volume in (G–I). Magenta and cyan KM trends correspond to the tumor characteristics, respectively, above and below median response threshold for the studied ADC histogram metrics. The color flip from cyan to magenta for responder KM trends (with higher probability of survival) between mean ADC (D–F) and volume-based metrics (A–C, G–I) reflects negative change in tumor volume versus positive change in ADC metrics. Time-dependent distance between KM curves reports on predictive power of the studied histogram metrics. Vertical tick-marks along KM trends indicate individual patients whose survival times have been censored. Dashed vertical line corresponds to the minimal survival time included into the corresponding KM CDF probability analysis (excluding survival for the late censored patients).

ence KM fDM (Figure 2C) that was based on changes between midTx and preTx. Significant enhancement of KM CDF separation ($P_{KM} = 0.003\text{--}0.05$, $mCPD = 0.17\text{--}0.2$) was observed for midTx ADC (Figure 3E) above a median response threshold of $1.25(\times 10^{-3} \text{ mm}^2/\text{s})$, as well as for change in whole-tumor mean ADC and total tumor-volume differences above versus below 1%–2% (Figure 3, C, E, and F). However, a notably high number (fourteen) of censored patients (Figure 3E, magenta ticks) made CDF estimate for midTx ADC metric unreliable beyond 21-months survival (Figure 3E, dashed). The similarity of the fractional volume KM to reference fDM was -87% , notably higher than that for significant (midTx and fractional change) ADC metrics, consistent with volumetric nature of the fDM analysis. This is also consistent with observation of high KM similarity

(-86%) for low-ADC volume midTx (Figure 3H). The general color “flip” for responder KM trends based on volume metrics (Figure 3, A–C, G–I, cyan) versus ADC metrics (Figure 3, D–F, magenta) reflected negative change in tumor volume versus positive change in ADC metrics related to higher probability of survival.

The best KM survival probability CDF estimator in Figure 3I (with maximum $mCPD = 0.22$ and minimum $P_{KM} < 0.001$) was based on the fraction low-ADC volume shrinkage (cyan KM trend). This estimator used combined tumor volume change and tumor density (ADC-threshold $< 1.25 \times 10^{-3} \text{ mm}^2/\text{s}$) information. The fractional low-ADC volume metric clearly showed similar predictive power (relative distance between KM CDF) as reference fDM KM (Figure 2C, $mCPD = 0.21$) based on the

increased fDM PRM midTx (“magenta” trend). The reliable CDF estimate for both reference (Figure 2C) and fractional low-ADC volume (Figure 3I) was confirmed by a small number (two) of patients censored beyond minimal CDF values of the corresponding KM trends (at survival probabilities of 0.3 and 0.07). The bulk of the KM differences between responders and nonresponders was evidently related to the low ADC volume midTx (Figure 3H), rather than preTx volume (Figure 3G), confirming that the functional response was triggered by treatment. The decreasing low-ADC volume midTx versus preTx (less than -8%, $P_{KM} < 0.001$) in Figure 3I, was significantly (negatively) correlated to increasing fDM (>4%, $P_{KM} < 0.001$) in Figure 2C and Table 1 (-92.5%), confirming fDM relation to shrinking tumor volume.

DISCUSSION

The decrease in low-ADC volume was found to be a good predictor of KM survival (treatment response) most similar to the fDM reference. The strong alignment between KM curves for fDM and low-ADC volume metrics confirms that the early response prediction power of increasing fDM likely stems from decreasing volume of shrinking dense tumor observed as early as 3 weeks after radiation therapy for glioma tumors. Interestingly, the fDM population-median KM threshold for responders versus nonresponders of 4% was still close to 4.7% that maximized AU-ROC as previously determined (12) despite the additional 25 subjects. Another supporting observation is that the population-median response threshold for mean ADC-based KM survival probability midTx corresponded to the dense tumor low-ADC integration limit of $1.25 \times 10^{-3} \text{ mm}^2/\text{s}$. The proximity of median thresholds for fractional ADC and tumor volume changes to 0% likely reflected KM sensitivity to the sign of the effect (increasing ADC and decreasing volume) rather than absolute metric value. The fact that no significance was observed for preTx low-ADC volume itself, suggested that midTx volume change was indeed reflective of the therapy efficacy. This specific relation to reduction of the dense tumor ADC volume and treatment option provided independent evidence for the biophysical origin of the fDM predictive power. Our analysis effectively revealed that fDM portions with low-ADC midTx report on the therapy response.

The main limitation of this study was that the data analysis was restricted to only two imaging end points, precluding evaluation of relative longitudinal changes in the histogram metrics over the full course of radiological surveillance. Furthermore, the KM thresholds were not optimized by AU-ROC analysis or cross-validation. These restrictions were intentional for the largely technical aims of this study to determine the ADC histogram metrics that had early response prediction power similar to the reference fDM, as shown by previous work (12), and to maximize method consistency across histogram and fDM analyses, reducing dependence on any residual study bias. For this reason, ADC histograms were derived from the same coregistered image sets and the same tumor segmentations as used to generate the reference fDM metrics, even though ADC histogram analysis can be performed on non-coregistered images. This study design precluded evaluation of sensitivity of low-ADC

histogram-based segmentation to image registration-related errors. For ADC histogram threshold method, the specific voxel locations are less important, and hence higher immunity is potentially expected to coregistration errors. This should be a topic of a future study.

Others have applied alternative ADC histogram-based analyses in the context of newly diagnosed (6, 10, 15) and recurrent (23) glioblastoma to predict response to antivascular chemotherapy used alone or in combination with radiation treatment. Technical aspects of histogram analysis varied. Bimodal mixed normal distribution fitting of the whole tumor ADC histogram into means of the low-ADC curve and high-ADC curve was performed by Pope et al. (10, 15, 23). In contrast, Wen et al. (6) analyzed specific percentile points of the ADC histogram. However, both methods consistently found greater predictive content in the low-ADC regime. Prediction metrics in both of these alternative histogram approaches were expressed in physical diffusion units (ie, square millimeter per second), whereas the method presented in this study focused on volume (ie, in cubic centimeter units) of ostensibly dense tumor defined by an ADC below a specified value, $1.25 \times 10^{-3} \text{ mm}^2/\text{s}$.

The low-ADC volume approach presented here parallels similar logic used to assess traditional response metrics based on tumor shrinkage assessed by conventional neuroimaging (24-26), although it exploits tumor density segmentation qualities inherent in diffusion mapping. A common feature in these various diffusion histogram approaches and fDM (or PRM) is a framework to deal with tumor heterogeneity and to avoid inclusion of preexisting cystic/necrotic portions of the tumor that can attenuate sensitivity to therapeutic changes in viable tumor. Response to treatment (or tumor progression) can be spatially nonuniform as well, and fDM/PRM provides means to map responsive/resistant/progression regions (11, 12, 27).

The current study design amplified ADC measurement sensitivity to the therapeutic effect by performing longitudinal patient surveillance scans on the same MRI system. Although desirable, this level of control may be challenging in the clinical setting. When multiple scanners are used, systematic biases may increase between-scan variability (eg, due to spatial *b*-value bias for anatomy at different offsets from isocenter (21, 22)). For longitudinal studies, these errors may potentially increase the population histogram noise and attenuate the absolute ADC measurement sensitivity to the therapeutic effect. In principle, such systematic errors should be monitored similar to normal-appearing white matter analysis in this study [or using phantoms with known ADC (21, 22)] and, when present, corrected using MRI system gradient characteristics before population ADC histogram analysis.

In conclusion, fDM changes diagnostic of early therapy response for high-grade glioma tumors are confirmed using comprehensive analysis of multiple ADC histogram metrics. Reduction in solid (non-necrotic) tumor volume correlates with low-ADC fDM changes. Histogram-based ADC segmentation facilitates elimination of high-mobility (necrotic) tissue, allowing for focusing on shrinkage of low-mobility (cellular-dense) tumor regions.

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health Grants: U01CA166104, R44CA210825, and P01CA085878, and by the Swedish Cancer Society CAN 2016/365.

REFERENCES

- Ellingson BM, Malkin MG, Rand SD, Connelly JM, Quinsey C, LaViolette PS, Bedekar DP, Schmainda KM. Validation of functional diffusion maps (fDMs) as a biomarker for human glioma cellularity. *J Magn Reson Imaging*. 2010;31:538–548.
- Le Bihan D. Molecular diffusion, tissue microdynamics and microstructure. *NMR Biomed*. 1995;8:375–386.
- Squillaci E, Manenti G, Cova M, Di Roma M, Miano R, Palmieri G, Simonetti G. Correlation of diffusion-weighted MR imaging with cellularity of renal tumours. *Anticancer Res*. 2004;24:4175–4179.
- Chenevert TL, Stegman LD, Taylor JM, Robertson PL, Greenberg HS, Rehemtulla A, Greenberg HS, Rehemtulla A, Ross BD. Diffusion magnetic resonance imaging: an early surrogate marker of therapeutic efficacy in brain tumors. *J Natl Cancer Inst*. 2000;92:2029–2036.
- Nagane M, Kobayashi K, Tanaka M, Tsuchiya K, Shishido-Hara Y, Shimizu S, Shiokawa Y. Predictive significance of mean apparent diffusion coefficient value for responsiveness of temozolomide-refractory malignant glioma to bevacizumab: preliminary report. *Int J Clin Oncol*. 2014;19:16–23.
- Wen Q, Jalilian L, Lupo JM, Molinaro AM, Chang SM, Clarke J, Prados M, Nelson SJ. Comparison of ADC metrics and their association with outcome for patients with newly diagnosed glioblastoma being treated with radiation therapy, temozolomide, erlotinib and bevacizumab. *J Neurooncol*. 2015;121:331–339.
- Qu J, Qin L, Cheng S, Leung K, Li X, Li H, Dai J, Jiang T, Akgoz A, Seethamraju R, Wang Q, Rahman R, Li S, Ai L, Jiang T, Young GS. Residual low ADC and high FA at the resection margin correlate with poor chemoradiation response and overall survival in high-grade glioma patients. *Eur J Radiol*. 2016;85:657–664.
- Chenevert TL, McKeever PE, Ross BD. Monitoring early response of experimental brain tumors to therapy using diffusion magnetic resonance imaging. *Clin Cancer Res*. 1997;3:1457–1466.
- Higano S, Yun X, Kumabe T, Watanabe M, Mugikura S, Umetsu A, Sato A, Yamada T, Takahashi S. Malignant astrocytic tumors: clinical importance of apparent diffusion coefficient in prediction of grade and prognosis. *Radiology*. 2006;241:839–846.
- Pope WB, Kim HJ, Huo J, Alger J, Brown MS, Gjertson D, Sai V, Young JR, Tekchandani L, Cloughesy T, Mischel PS, Lai A, Nghiemphu P, Rahmanuddin S, Goldin J. Recurrent glioblastoma multiforme: ADC histogram analysis predicts response to bevacizumab treatment. *Radiology*. 2009;252:182–189.
- Ellingson BM, Malkin MG, Rand SD, LaViolette PS, Connelly JM, Mueller WM, Schmainda KM. Volumetric analysis of functional diffusion maps is a predictive imaging biomarker for cytotoxic and anti-angiogenic treatments in malignant gliomas. *J Neurooncol*. 2011;102:95–103.
- Hamstra DA, Galban CJ, Meyer CR, Johnson TD, Sundgren PC, Tsien C, Lawrence TS, Junck L, Ross DJ, Rehemtulla A, Ross BD, Chenevert TL. Functional diffusion map as an early imaging biomarker for high-grade glioma: correlation with conventional radiologic response and overall survival. *J Clin Oncol*. 2008;26:3387–3394.
- Moffat BA, Chenevert TL, Lawrence TS, Meyer CR, Johnson TD, Dong Q, Tsien C, Mukherji S, Quint DJ, Gebarski SS, Robertson PL, Junck LR, Rehemtulla A, Ross BD. Functional diffusion map: a noninvasive MRI biomarker for early stratification of clinical brain tumor response. *Proc Natl Acad Sci U S A*. 2005;102:5524–5529.
- Hastie T, Tibshirani I, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer; 2001.
- Pope WB, Lai A, Mehta R, Kim HJ, Qiao J, Young JR, Xue X, Goldin J, Brown MS, Nghiemphu PL, Tran A, Cloughesy TF. Apparent diffusion coefficient histogram analysis stratifies progression-free survival in newly diagnosed bevacizumab-treated glioblastoma. *AJNR Am J Neuroradiol*. 2011;32:882–889.
- Berglund AE. *MatSurv*.m 2017. <https://github.com/aeberg/MatSurv>.
- Clunie DA. DICOM structured reporting and cancer clinical trials results. *Cancer Inform*. 2007;4:33–56.
- MHD: image metadata format Public Wiki2014. <https://itk.org/Wiki/ITK/MetalO/Documentation>.
- Klein S, Staring M, Murphy K, Viergever MA, Pluim JP. elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging*. 2010;29:196–205.
- Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M, Buatti J, Aylward S, Miller JV, Pieper S, Kikinis R. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging*. 2012;30:1323–1341.
- Malyarenko D, Galban CJ, Londy FJ, Meyer CR, Johnson TD, Rehemtulla A, Ross BD, Chenevert TL. Multi-system repeatability and reproducibility of apparent diffusion coefficient measurement using an ice-water phantom. *J Magn Reson Imaging*. 2013;37:1238–1246.
- Mulkern RV, Ricci KI, Vajapeyam S, Chenevert TL, Malyarenko DI, Kocak M, Poussaint TY. Pediatric brain tumor consortium multisite assessment of apparent diffusion coefficient z-axis variation assessed with an ice-water phantom. *Acad Radiol*. 2015;22:363–369.
- Pope WB, Qiao XJ, Kim HJ, Lai A, Nghiemphu P, Xue X, Ellingson BM, Schiff D, Aregawi D, Cha S, Puduvali VK, Wu J, Yung WK, Young GS, Vredenburg J, Barboriak D, Abrey LE, Mikkelsen T, Jain R, Paleologos NA, Rn PL, Prados M, Goldin J, Wen PY, Cloughesy T. Apparent diffusion coefficient histogram analysis stratifies progression-free and overall survival in patients with recurrent GBM treated with bevacizumab: a multi-center study. *J Neurooncol*. 2012;108:491–498.
- Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D, Verweij J. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45:228–247.
- Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, Ohgaki H, Wiestler OD, Kleihues P, Ellison DW. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol*. 2016;131:803–820.
- Wen PY, Macdonald DR, Reardon DA, Cloughesy TF, Sorensen AG, Galanis E, Degroot J, Wick W, Gilbert MR, Lassman AB, Tsien C, Mikkelsen T, Wong ET, Chamberlain MC, Stupp R, Lamborn KR, Vogelbaum MA, van den Bent MJ, Chang SM. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. *J Clin Oncol*. 2010;28:1963–1972.
- Ellingson BM, Cloughesy TF, Lai A, Mischel PS, Nghiemphu PL, Lalezari S, Schmainda KM, Pope WB. Graded functional diffusion map-defined characteristics of apparent diffusion coefficients predict overall survival in recurrent glioblastoma treated with bevacizumab. *Neuro Oncol*. 2011;13:1151–1161.

Disclosures: TLC, CJG, and BDR are coinventors on intellectual property assigned to and managed by the University of Michigan licensed by Imbio for histogram and fDM analysis.