Title
Quantitative probabilistic functional diffusion mapping in newly diagnosed glioblastoma treated with radiochemotherapy

Permalink
https://escholarship.org/uc/item/50d1z1sw

Journal
Neuro-Oncology, 15(3)

ISSN
1522-8517

Authors
Ellingson, BM
Cloughesy, TF
Lai, A
et al.

Publication Date
2013-03-01

DOI
10.1093/neuonc/nos314

Peer reviewed
Background. Functional diffusion mapping (fDM) is a cancer imaging technique that uses voxel-wise changes in apparent diffusion coefficients (ADC) to evaluate response to treatment. Despite promising initial results, uncertainty in image registration remains the largest barrier to widespread clinical application. The current study introduces a probabilistic approach to fDM quantification to overcome some of these limitations.

Methods. A total of 143 patients with newly diagnosed glioblastoma who were undergoing standard radiochemotherapy were enrolled in this retrospective study. Traditional and probabilistic fDMs were calculated using ADC maps acquired before and after therapy. Probabilistic fDMs were calculated by applying random, finite translational, and rotational perturbations to both pre- and posttherpay ADC maps, then repeating calculation of fDMs reflecting changes after treatment, resulting in probabilistic fDMs showing the voxel-wise probability of fDM classification. Probabilistic fDMs were then compared with traditional fDMs in their ability to predict progression-free survival (PFS) and overall survival (OS).

Results. Probabilistic fDMs applied to patients with newly diagnosed glioblastoma treated with radiochemotherapy demonstrated shortened PFS and OS among patients with a large volume of tumor with decreasing ADC evaluated at the posttreatment time with respect to baseline scans. Alternatively, patients with a large volume of tumor with increasing ADC evaluated at the posttreatment time with respect to baseline scans were more likely to progress later and live longer. Probabilistic fDMs performed better than traditional fDMs at predicting 12-month PFS and 24-month OS with use of receiver-operator characteristic analysis. Univariate log-rank analysis on Kaplan–Meier data also revealed that probabilistic fDMs could better separate patients on the basis of PFS and OS, compared with traditional fDMs.

Conclusions. Results suggest that probabilistic fDMs are a more predictive biomarker in terms of 12-month PFS and 24-month OS in newly diagnosed glioblastoma, compared with traditional fDM analysis.

Keywords: diffusion MRI, fDM, functional diffusion map, glioblastoma.
Functional diffusion imaging (fDM) is one method in a new class of imaging biomarkers that quantifies the voxel-wise changes in physiological parameters to assess tumor response to therapy. By quantifying voxel-wise changes in ADC (DADC) measured in the same patient over time, fDMs do not assume homogeneity in tumors. Thus, fDMs are a powerful tool for evaluating heterogeneous tumors, including malignant gliomas, such as glioblastoma. FDMs have been shown to be useful in monitoring patients with gliomatosis cerebri, predicting response to standard radiochemotherapy in newly diagnosed glioblastoma and predicting response to anti-angiogenic therapy in recurrent glioblastoma. Despite these promising findings, fDMs have not been implemented clinically in part because of challenges regarding reproducibility of image maps and quantitative information. The accuracy of image coregistration among subsequent ADC maps is the primary mechanism for inaccuracies and errors in fDM analysis and interpretation and the reason that fDM results may vary across different institutions and MR scanners. In the current study, we used a probabilistic approach to fDM quantification, in which finite translational and rotational perturbations were performed after linear registration of ADC maps. These probabilistic fDMs (prob-fDMs) were then applied to a large cohort of patients with newly diagnosed glioblastoma treated with standard radiochemotherapy (n = 143) to determine whether probabilistic fDMs were a better predictor of progression-free survival (PFS) and overall survival (OS), compared with traditional fDMs.

**Methods**

**Patients**

All patients participating in this study signed institutional review board–approved informed consent to have their information in our neuro-oncology database. A total of 143 patients with histologically confirmed, newly diagnosed GBM with high-quality DWIs before and after initiation of radiochemotherapy (external beam radiation therapy and temozolomide) were included in the current retrospective study. Baseline (postsurgical, pretreatment) scans were obtained 1 week before therapy, and posttreatment scans were obtained 4–6 weeks after completion of radiochemotherapy. A total of 66 of the 143 patients was eventually treated with bevacizumab, at either the first or second recurrence. No patients were treated with bevacizumab during the periods used for fDM analysis. The same cohort of patients was analyzed as part of a previous fDM study involving traditional analyses to directly compare probabilistic fDM performance. More details regarding specific patient characteristics can be found in this previous study.

**MRI**

Data were collected on 1.5T MR systems (General Electric Medical Systems, Waukesha, WI; Siemens Medical Solutions, Erlangin, Germany) using pulse sequences supplied by the scanner manufacturer. Standard anatomical MRI sequences included axial T1-weighted (TE/TR = 15 ms/400 ms, slice thickness 5 mm with 1 mm interslice distance, number of excitations [NEX] = 2, matrix size = 256 × 256, and field-of-view [FOV] = 24 cm), T2-weighted fast spin-echo (TE/TR = 126–130 ms/2000 ms, slice thickness 5 mm with 1 mm interslice distance, NEX = 2, matrix size = 256 × 256, and FOV = 24 cm), and fluid-attenuated inversion recovery (FLAIR) images (TI = 2200 ms, TE/TR = 120 ms/4000 ms, slice thickness 5 mm with 1 mm interslice distance, NEX = 2, matrix size = 256 × 256, and FOV = 24 cm). DWIs were collected with TE/TR = 102.2 ms/8000 ms, NEX = 1, slice thickness = 5 mm with 1 mm interslice distance, matrix size = 128 × 128 (reconstructed images were zero-padded and interpolated to 256 × 256), and a FOV = 24 cm using a twice-refocused spin echo planar preparation. ADC images were calculated from acquired DWIs with b = 1000 s/mm² and b = 0 s/mm² images. In addition, gadopentetate dimeglumine–enhanced (Magnevist; Berlex, Wayne, NJ; 0.1 mmol/kg) axial and coronal T1-weighted images (T1 + C; coronal: TE/TR = 15 ms/400 ms, slice thickness 3 mm with 1 mm interslice distance, NEX = 2, a matrix size of 256 × 256, and FOV = 24 cm) were acquired after contrast injection.

**Initial Affine Registration**

All images for each patient were registered to their own pretreatment, postcontrast, T1-weighted image with use of a mutual information algorithm and a 12 degree of freedom transformation using FSL (FMRIB, Oxford, UK; http://www.fmrib.ox.ac.uk/fsl/). Fine registration (1–2 degrees and 1–2 voxels) was then performed using a Fourier transform-based, 6 degree of freedom, rigid body registration algorithm, followed by visual inspection to ensure adequate alignment. All images were interpolated to the resolution of baseline T1-weighted images with use of trilinear interpolation.

**Finite Perturbations**

After linear registration, finite perturbation of both baseline (pretreatment) ADC maps and posttreatment ADC maps was performed along 6 degrees of freedom (x, y, and z translation and u, f, and f rotation) with use of a uniform probability density function for both translation and rotation. Specifically, the probability density function for translation was defined as:

\[
p(D_x) = \begin{cases} 
\frac{1}{10} & -5 \text{ mm} \leq D_x \leq +5 \text{ mm} \\
0 & |D_x| > 5 \text{ mm} 
\end{cases}
\]
Individual voxels were stratified on the basis of the change in ADC relative to the baseline ADC map. Voxels having an ADC increase [ADC(+)] or decrease [ADC(−)] beyond a DADC threshold of 0.4 mm²/ms², representing the 95% confidence interval for DADC in a mixture of normal appearing gray and white matter, were labeled with red or blue, respectively, in traditional fDMs. For probabilistic fDMs, voxels having an increase or decrease in DADC threshold of 0.4 mm²/ms were categorized as such after each perturbation for a total of 1000 finite, random perturbations. The fraction of all perturbations resulting in significant increases or decreases in DADC was retained for further analysis, resulting in probabilistic maps p(ADC(+)dxdydz) and p(ADC(−)dxdydz). For traditional fDMs, the total relative hypercellular [ADC(2)] and hypocellular [ADC(+)] burden was quantified as volume fraction of each within the postsurgical, pretreatment FLAIR hyperintense ROIs (e.g., % ADC(+) volume of FLAIR with increasing ADC/volume of FLAIR hyperintensity). For probabilistic fDMs, the total relative hypercellular and hypocellular tumor burden was quantified as the sum of all probabilities within postsurgical, pretreatment FLAIR hyperintense ROI, divided by the volume of FLAIR abnormal regions:

\[
\text{Prob} - \text{fDM} % \text{ADC(+) = } \frac{p(\text{ADC}(+)dxdydz)}{\text{ROI volume}}
\]

\[
\text{Prob} - \text{fDM} % \text{ADC(−) = } \frac{p(\text{ADC}(−)dxdydz)}{\text{ROI volume}}
\]

The total number of perturbations was chosen on the basis of an examination of a subset of patients using different numbers of perturbations ranging from 10 to 1 million, for which probabilistic fDMs with 1000 perturbations resulted in probabilities similar to those of fDMs created with 1 million perturbations.

To gain better insight into the difference between results obtained from probabilistic and traditional fDMs, consider the following example. Probabilistic fDMs classify a group of 10 voxels as decreasing [ADC(2)] but at a probability of only 25% for each voxel (of 1000 perturbations, these voxels were considered to be significantly lower in 250 perturbations). Thus, the volume fraction for probabilistic fDMs is 10 voxels × 25% ÷ 2.5. Next, consider a grouping of only 3 voxels categorized on traditional fDMs. Although this is a smaller actual volume of fDM-classified tissue, fDM analysis results in a volume fraction of 3 × 100% ÷ 3.0, which is larger than the volume fraction obtained through probabilistic fDMs. As demonstrated by this example, it is not the entire volume extent of fDM-classified tumor that is important in probabilistic fDMs, but rather the combination of volume and the probability of being classified as abnormal.

**Regions of Interest (ROIs)**

In the current study, we used volumetric FLAIR ROIs to interpret fDM results, which have previously been used in nonenhancing and enhancing tumors, because tumor infiltration into normal brain parenchyma typically results in an increase in T2-weighted, or FLAIR, abnormal signal, and multiple investigations have suggested that T2 signal abnormalities should be routinely used to visualize the extent of malignant infiltrating tumor. ROIs were calculated by a single investigator and verified by a board-certified neuroradiologist using a semi-automated procedure consisting of (1) manually defining the relative region of tumor occurrence, (2) thresholding the FLAIR images in these regions using an empirical threshold combined with a region-growing algorithm, and (3) manually editing the resulting masks to exclude any obvious postsurgical cystic or necrotic areas. ROIs were drawn on every slice that contained tumor (i.e., 3D ROIs). ROIs were completed using custom scripts in Analysis of Functional NeuroImages software (http://afni.nimh.nih.gov/afni).
Receiver-Operator Characteristic (ROC) Analysis and Survival Analysis

ROC curves defining the sensitivity and specificity for predicting 12-month PFS and 24-month OS were constructed for both traditional and probabilistic fDMs. The area under the ROC curve (AUC) was used to evaluate ROC performance. For survival analysis, both probabilistic and traditional fDMs were stratified by the median volume fraction for traditional fDMs in all patients (20% and 20% of pretreatment FLAIR ROI). Log-rank statistical analysis was used to test the hypothesis that a large probabilistic fDM-classified volume fraction of decreasing ADC (prob-fDM % ADC[2] . 20%) is indicative of a nonresponsive, growing hypercellular tumor and will result in a significantly shorter PFS and OS. Log-rank analysis was also performed to test the hypothesis that a large probabilistic fDM-classified volume fraction of increasing ADC (prob-fDM % ADC[+] . 20%) is indicative of destruction of tumor cells, a favorable response to radiotherapy, resulting in a significantly longer PFS and OS. PFS was defined from the time of diagnosis to radiographic or neurological progression by the treating neuro-oncologists, as described in a previous publication.6 To decrease the likelihood of declaring progression in the context of pseudoprogression, the final treatment scan after radiation was considered to be the baseline scan for evaluating tumor progression. If the subsequent scans showed a definite increase in tumor (≥25% increase in the sum of enhancing lesions, new enhancing lesion . 1 cm2, an unequivocal qualitative increase in non–contrast-enhancing tumor, or unequivocal new area of non–contrast-enhancing tumor), progression was declared at that time. Change in steroid dose was taken into consideration before defining progression. Patients who did not meet these imaging criteria for progression but had significant neurological decline were declared progressed at the time of irreversible decline. Patients who died before evidence of imaging progression were defined as progressed on the date of death. OS was defined from the time of diagnosis until death.

Results

Probabilistic fDMs, quantifying the probability of obtaining fDM classification per voxel during random, finite perturbations, showed a range of regions with significant increasing and decreasing ADCs. As the number of random perturbations increased, an increasing number of voxels within pretreatment FLAIR abnormal regions were subject to fDM classification, albeit often with a lower probability (Fig. 1A and B). These results suggest that traditional fDMs using only a single registration step may, in fact, introduce misregistration error. Although the spatial extent of classified tumor was larger in probabilistic, compared with traditional fDMs, tumor burden on probabilistic fDMs is dependent on both the probability of classification and the spatial extent of classified voxels. In general, regions of highest probability of increasing and decreasing ADC corresponded well spatially with traditional fDMs (Fig. 1C and D); however, probabilistic fDMs appeared to identify regions suspected of subtle tumor growth not identified by traditional fDMs (Fig. 1D). Qualitatively, patients with a large volume of tumor with decreasing ADC on probabilistic fDMs after radiochemotherapy, suggestive of an increase in tumor cell density, appeared to be more likely to progress and die sooner (Fig. 2A–E). Alternatively, patients with a larger tumor volume exhibiting an increasing ADC appeared to be more likely to progress later and live longer (Fig. 2F–J).

ROC Analysis

Both probabilistic and traditional fDMs performed modestly on ROC analysis with respect to 12-month PFS and 24-month OS. A high volume fraction of decreasing ADC on probabilistic fDMs, or high prob-fDM % ADC[2], was a significant predictor of 12-month PFS (Fig. 3A; AUC 0.6624, P < 0.001), showing a sensitivity of 71% and a specificity of 69% for predicting short-term progression-free survivors with use of a threshold of prob-fDM % ADC[2] . 20%. Similarly, traditional fDMs (% ADC[2]) in FLAIR-abnormal regions were also a significant predictor of 12-month PFS (Fig. 3A; AUC 0.6423, P = 0.0107).

Probabilistic fDMs were also a significant predictor of 12-month PFS when evaluating the volume fraction of increasing ADC (Fig. 3B; AUC 0.6186, P = 0.0352), resulting in a 54% sensitivity and 75% specificity for identifying patients with 12-month PFS with use of a threshold of prob-fDM % ADC[+] of 20%. Traditional fDM volume fraction of increasing ADC (% ADC[+]) was not a significant predictor of 12-month PFS (Fig. 3B; AUC 0.6046, P = 0.0631).

Similar to trends with PFS, probabilistic fDMs consistently outperformed traditional fDMs with respect to 24-month OS. The volume fraction of pretreatment FLAIR abnormal tissue with decreasing ADC on traditional fDMs (% ADC[2]) did not predict 24-month OS according to ROC analysis (Fig. 3C; AUC 0.5788, P = 0.1511); however, the volume fraction of decreasing ADC on probabilistic fDMs was a significant predictor (Fig. 3C; AUC 0.6117, P = 0.0419), showing 58% sensitivity and 70% specificity for 24-month OS with use of a threshold of 20%. Similarly, the volume fraction of pretreatment FLAIR-abnormal tissue with increasing ADC on traditional fDMs (% ADC[+]) was not a significant predictor of 24-month OS (Fig. 3D; AUC 0.5507, P = 0.3558), whereas probabilistic fDMs could predict 24-month OS with a 56% sensitivity and 69% specificity with use of a threshold of 20% (Fig. 3D; AUC 0.6516, P = 0.0074).

Survival Analysis

Log-rank analysis applied to Kaplan–Meier data suggested that patients with a probabilistic fDM-classified...
Fig. 1. Construction of probabilistic fDMs. (A) Probability of voxels exhibiting a decrease in ADC beyond $\geq 0.4 \text{mm}^2/\text{ms}$ between pre- and posttreatment time points as a function of perturbation (iteration) number. (B) Probability of voxels exhibiting an increase in ADC beyond $+0.4 \text{mm}^2/\text{ms}$ between pre- and posttreatment time points as a function of perturbation (iteration) number. (C) The traditional fDM constructed from only a single registration step. (D) Probabilistic fDMs for the same slice after 1000 random, finite translational, and rotational perturbations. Arrow in (D) shows a region suspected of subtle tumor growth not identified by traditional fDMs.

Fig. 2. M R images and probabilistic fDMs for a short-term (top row) and long-term survivor (bottom row). (A and F) Postsurgical, pretherapy, postcontrast T1-weighted image. (B and G) Posttherapy, postcontrast T1-weighted image. (C and H) Postsurgical, pretherapy FLAIR image. (D and I) Posttherapy FLAIR image. (E) Probabilistic fDMs showing a larger volume fraction of tumor with decreasing ADC in a short-term survivor, and (J) a long-term survivor exhibited a larger volume fraction of tumor with increasing ADC.
volume fraction of decreasing ADC in 20% of pretreatment FLAIR-abnormal regions (prob-fDM [%ADC(-)] 20%), indicative of increased tumor cellularity, had a significantly shorter PFS, compared with patients with a lower volume fraction (Fig. 4A; log-rank, P = .0001; hazard ratio [HR] 3.3). Median PFS among patients with probabilistic fDM-classified %ADC(2) 20% was 239 days, compared with 602 days for patients with %ADC(2) 20%. The same trends were apparent with traditional fDMs (Fig. 4A; log-rank, P = .0004; HR 2.2), although the HR was significantly lower, compared with probabilistic fDMs (95% confidence interval for probabilistic fDMs 1.7, HR 1.3), suggesting that decreasing ADC on probabilistic fDMs may provide better separation of PFS, compared with traditional fDMs, with use of the same criterion. Patients with a probabilistic fDM-classified volume fraction of increasing ADC 20% (prob-fDM [%ADC(+)] 20%), indicative of decreased cell density, had a significantly longer PFS, compared with patients with a lower volume fraction (Fig. 4B; log-rank, P = .0001, HR 2.5). Traditional fDMs using the same ROI and threshold criteria did not show significant separation of patient groups (Fig. 4B; log-rank, P = .3366, HR 1.3). Patients with a volume fraction of pretreatment FLAIR-abnormal tissue with a decreased ADC on probabilistic fDMs 20% (prob-fDM [%ADC(2)] 20%) were more likely to die sooner than patients with a lower volume fraction (Fig. 4C; log-rank, P = .0001; HR 3.8). Similar to trends in PFS, a large volume fraction of decreasing ADC on traditional fDMs was also predictive of shorter OS (Fig. 4C; log-rank, P = .0026; HR 1.3). Patients with a volume fraction of pretreatment FLAIR-abnormal tissue exhibiting an increase in ADC 20%, suggestive of more tumor destruction, were more likely to live longer when evaluated using probabilistic fDMs (Fig. 4D; log-rank, P = .0105, HR 1.9), but not when evaluated using traditional fDMs (Fig. 4; log-rank, P = .6516, HR 1.0).
Discussion

Voxel-based temporal analysis of physiological imaging data constitutes a new class of imaging biomarkers. These biomarkers include fDMs; cell invasion, motility, and proliferation level estimate maps; functional perfusion maps or perfusion parametric response maps; and differential quantitative T2 maps. fDMs constitute the oldest and most tested of these imaging biomarkers, having been extensively validated and showing sensitivity to predict response to a variety of treatments. The lack of a measure of uncertainty for image registration, however, is a significant limitation for fDMs and one shared among all voxel-based biomarkers. The current study focused on overcoming this limitation by providing a probabilistic approach to estimating misregistration uncertainty.

Probabilistic fDMs applied to pretreatment FLAIR-abnormal regions were predictive of tumor response to radiochemotherapy and followed trends consistent with previous diffusion MR and traditional fDM studies. ADC is expected to increase as tumor cells are destroyed by cytotoxic therapy. Consistent with this theory and previous fDM analyses, results from the current study suggest that a large volume fraction of tumor that exhibits an increase in ADC is characteristic of more favorable PFS and OS. Alternatively, if tumors are not responding to therapy, a decrease in ADC is expected because of increasing tumor cell density. Results from the current and previous studies support this hypothesis, revealing that patients with a large volume fraction of tumor showing a decrease in ADC had a significantly shorter PFS and OS. Despite similar trends between the 2 techniques, results from the current study show superior performance of probabilistic fDMs, when compared with traditional fDMs, in terms of their ability to predict PFS and OS in patients with newly diagnosed glioblastoma treated with standard radiochemotherapy. In particular, the volume fraction of both increasing and decreasing ADC on probabilistic fDMs had a significantly higher sensitivity and specificity, compared with traditional fDMs, with respect to predicting 12-month PFS and 24-month OS evaluated using ROC analysis. In addition, log-rank analyses clearly showed greater separation between patient cohorts stratified by similar criteria.
between the 2 techniques, supported by statistically significant differences in HRs observed between the 2 techniques. Together, these results support the use of a probabilistic approach to analyze voxel-wise changes in physiological parameters through the use of finite registration perturbations.

**Study Limitations**

The current study used standard, clinically available DWI sequences that may have posed some potential limitations, including lack of mid-range b values for estimation of perfusion-insensitive ADC according to the recommendations of the National Cancer Institute along with using a single-shot echoplanar acquisition, which may have resulted in geometric distortions, leading to increased inaccuracies in image registration. Although previous fDM studies have used similar image acquisition strategies, these limitations may have decreased the clinical sensitivity of both traditional and probabilistic fDMs. In addition, studies have suggested that an additional nonlinear registration step may be advantageous for increasing clinical sensitivity; however, the current study did not directly compare with nonlinear fDMs. It is conceivable that nonlinear registration applied before finite, random perturbations may further increase sensitivity to infiltrating tumor. Although relatively difficult to apply in the context of ill-defined structures, such as neoplasms, further analysis involving the use of the Dice Coefficient after each perturbation may be beneficial for additional quantitative comparisons between alignment techniques. In addition, changes in ADC attributable to other pathologies are well documented and are a potential confounder; therefore, clinical implementation of fDMs should also involve interpretation from clinicians to rule out contributions from other factors.

**Conclusion**

Probabilistic fDMs provide a robust alternative to the use of traditional fDMs with respect both to visualization and to quantification of patient response to treatment. Probabilistic fDMs allow for a relatively registration-independent estimate of fDM parameters, using the sum of all probabilities of fDM classification after random perturbations as a new measure of tumor burden. Probabilistic fDMs are a better prognostic biomarker than traditional fDMs. Future studies aimed at examining the histological features of regions identified as suspected subtle tumor growth on probabilistic fDMs are warranted.

**Conflict of interest statement.** None declared.

**Funding**

This work was supported by National Institutes of Health, National Cancer Institute (R21CA167354 to B. M. E.); UCLA Institute for Molecular Medicine Seed Grant (to B. M. E.); UCLA Radiology Exploratory Research Grant (to B. M. E.); University of California Cancer Research Coordinating Committee Grant (to B. M. E.); ACRIN Young Investigator Initiative Grant (to B. M. E.); Art of the Brain (to T. F. C.); Ziering Foundation in memory of Sigi Ziering (to T. F. C.); Singleton Family Foundation (to T. F. C.); and Clarence Klein Fund for Neuro-Oncology (to T. F. C.).

**References**


