A Simple Automated Method for Detecting Recurrence in High-Grade Glioma


doi: https://doi.org/10.3174/ajnr.A4873
http://www.ajnr.org/content/37/11/2019
A Simple Automated Method for Detecting Recurrence in High-Grade Glioma


ABSTRACT

SUMMARY: Our aim was to develop an automated multiparametric MR imaging analysis of routinely acquired imaging sequences to identify areas of focally recurrent high-grade glioma. Data from 141 patients treated with radiation therapy with a diagnosis of high-grade glioma were reviewed. Strict inclusion/exclusion criteria identified a homogeneous cohort of 12 patients with a nodular recurrence of high-grade glioma that was amenable to focal re-irradiation (cohort 1). T1WI, FLAIR, and DWI data were used to create subtraction maps across time points. Linear regression was performed to identify the pattern of change in these 3 imaging sequences that best correlated with recurrence. The ability of these parameters to guide treatment decisions in individual patients was assessed in a separate cohort of 4 patients who were treated with radiosurgery for recurrent high-grade glioma (cohort 2). A leave-one-out analysis of cohort 1 revealed that automated subtraction maps consistently predicted the radiologist-identified area of recurrence (median area under the receiver operating characteristic curve = 0.91). The regression model was tested in preradiosurgery MRI in cohort 2 and identified 8 recurrent lesions. Six lesions were treated with radiosurgery and were controlled on follow-up imaging, but the remaining 2 lesions were not treated and progressed, consistent with the predictions of the model. Multiparametric subtraction maps can predict areas of nodular progression in patients with previously treated high-grade gliomas. This automated method based on routine imaging sequences is a valuable tool to be prospectively validated in subsequent studies of treatment planning and posttreatment surveillance.

ABBREVIATIONS: FSL = fMRI of the Brain Software Library; GBM = glioblastoma; GKRS = gamma knife radiosurgery; HGG = high-grade glioma; ROC = receiver operating characteristic

High-grade glioma (HGG) consists of World Health Organization grade III and IV gliomas, with glioblastoma (GBM) being the most common primary intracranial malignancy.1 Prognosis is particularly poor in GBM, with 5-year survival <10% and median time to recurrence of approximately 6–7 months.2,3 Most recurrences are within or adjacent to the primary tumor site,4 and multicentric disease is a very uncommon entity.5,6

Because posttreatment changes occur within areas at highest risk of tumor recurrence, discerning expected radiographic changes from evolving tumor presents a major challenge in neuroradiology.

There is often debate regarding the proper assessment of MR imaging in the routine surveillance of patients with HGG.7–10 Imaging assessment may be limited by tumor heterogeneity, radiation therapy–related changes, and the effects of systemic agents.11 Attempts to standardize imaging assessment have previously been by implementation of the criteria of Macdonald et al12 and now the Response Assessment in Neuro-Oncology criteria.13 However, there remains inherent subjectivity to tumor measurements, and the Response Assessment in Neuro-Oncology does not provide information regarding the spatial distribution of brain tumor progression. Approaches to improve prognostic imaging assessment include perfusion-weighted imaging,14,15 spectroscopy,16 highly diffusion-weighted imaging,17 and PET.18

The purpose of our study was to develop a method of quantifying changes in standard MR imaging parameters that is automated and does not rely on subjective user input. We applied digital subtraction maps by comparing images obtained across time (temporal subtraction maps) from multiple imaging se-
quences. A primary goal was to rely only on standard imaging acquisitions and to use accessible image analysis techniques to reduce barriers to general use. We hypothesized that standardized thresholds to define changes in gadolinium-enhanced T1WI, FLAIR hyperintensity, and diffusion restriction could provide an automated method to identify areas of nodular HGG recurrence.

We searched a large, single-institution data base for a highly selected group of patients with HGG with a well-defined area of recurrence in the setting of prior standard-of-care treatment for malignant glioma and other clinical signs of disease progression. These patients served as a training set to generate the pattern of change in MR imaging signal across gadolinium-enhanced T1WI, FLAIR, and apparent diffusion coefficient values that best correlated with tumor progression. This pattern was then validated in a separate group of patients treated with radiosurgery for recurrent HGG. Radiosurgery is very rarely used in the treatment of HGG, but this select group offers the ideal scenario to validate our measure.

Because radiosurgery delivers a highly conformal dose of radiation, the predictions of the model can be tested inside and outside the treatment field by reviewing posttreatment MR imaging.

MATERIALS AND METHODS

Patients
All patients included in this study either had their informed consent waived on the basis of the retrospective nature of the analysis (before March 2015) or signed informed consent for participation (beginning in March 2015). Data acquisition was performed in compliance with all applicable Health Insurance Portability and Accountability Act regulations and was approved by the institutional review board of Columbia University. The study spanned March 2013 to November 2015, and MR imaging data were collected at a single institution on one of a variety of scanners. We chose March 2013 as the start date for the analysis, when the institution began routinely considering patients for focal re-irradiation for recurrent glioblastoma on the basis of the opening of a multicenter study investigating the role of re-irradiation in recurrent GBM. All MR imaging data consisted of standard clinical scan acquisitions and were not obtained for research purposes.

We searched our intradepartmental data base for all patients with a pathologic diagnosis of HGG who also met all of the following criteria: 1) underwent gross total or subtotal resection followed by concurrent temozolomide and radiation therapy; 2) were treated with either hypofractionated radiation of 4005 cGy in 15 fractions or standard fractionation of 5940–6000 cGy in 30–33 fractions; 3) had at least 2 postradiation MR imaging scans with no findings suspicious for progression; 4) developed disease progression that was deemed unequivocal by a board-certified neuroradiologist; 5) had all MR imaging available for review, including gadolinium-enhanced T1WI, FLAIR, and ADC measures; and 6) was the subject of an interdisciplinary discussion (ie, tumor board) resulting in agreement that the patient had experienced a clinically meaningful focal recurrence amenable to re-irradiation.

From an initial group of 141 consecutive patients referred for radiation therapy for HGG (On-line Fig 1), these criteria yielded 22 patients with recurrent HGG who met the inclusion criteria above. Of these, 12 patients were included in the initial analysis (cohort 1), and a separate cohort of 4 patients was identified with gamma knife radiosurgery (GKRS) delivered to focal nodular recurrence (cohort 2). Disease recurrence was proved by pathology in 7 patients, all in cohort 1 (58.3%).

Image Processing and Statistical Analysis
Additional methods, including all image-processing routines, are provided in the On-line Appendix. All diffusion-weighted scans were acquired at 2 b-values of 0 and 1000 s/mm². Image processing was performed with the MR software library19 (FSL; http://www.fmrib.ox.ac.uk/fsl) and Matlab (MathWorks, Natick, Massachusetts). Brain extraction was performed for each imaging sequence, which was coregistered with a linear algorithm by using the most recently obtained T1WI gadolinium-enhanced scan.

All sites of disease recurrence were delineated by hand using FSL software on the T1WI gadolinium-enhanced scans, and inverted linear registration matrices based on a mutual information cost function were applied to the ROIs to perform the analyses in their native space. Areas of disease recurrence were identified on the basis of radiology reports. Only the report for the scan being contoured was available for review, and we were blinded to the ultimate outcome. Signal from the tumor ROI was normalized to signal from voxels outside the tumor volume plus a margin to allow for microscopic disease, to account for signal variability between scans and MR imaging scanners. Specifically, manually drawn tumor volumes were expanded by at least 1.5–2 cm outside any surgical cavity, contrast-enhancing area, and FLAIR hyperintensity (ie, similar to the standard approach to defining the radiation clinical target volume). Voxels outside this margin were considered free of infiltrating tumor and were considered the “normal” brain. The brain was then normalized, by division, to the mean signal obtained from the nontumor brain volume for further analysis.

A subtraction image was created on the basis of scans from 2 dates: 1) the MR imaging scan that demonstrated the radiologist-defined area of tumor recurrence, and 2) the scan before the recurrence that was last read as stable, with no evidence of disease recurrence.

Temporal changes in signal between these 2 scan dates were analyzed by converting the difference image into z scores, which we term the “temporal subtraction map” (Fig 1). To identify the optimal combination of z scores, we used a hypothesis-driven approach as opposed to an evaluation of all possible combinations of z scores. We reasoned that the latter approach may identify a mathematically better fit to the data but could generate a clinically meaningless pattern of change. Specifically, we hypothesized that tumor recurrence would follow a pattern of increased T1WI contrast enhancement along with a concurrent increase in FLAIR signal intensity with a corresponding decrease in ADC values. In cohort 1, a linear regression with these restricted ranges of z scores was performed to generate a least-squares fit to the criterion standard radiologists’ definition of the recurrent area. This identified the optimal relative weighting for each imaging sequence at the individual patient level. The least-squares intercept and parameter estimates from this model were then averaged to the group level to provide the common pattern across the test group of patients. The model was tested by using a leave-one-out approach.
within cohort 1, and receiver operating characteristic (ROC) curves were generated with the area under the curve computed. The mean parameter estimates for each imaging sequence were multiplied by the respective imaging sequence (ie, T1WI postcontrast, FLAIR, and ADC) for each patient in the GKRS cohort, and the mean least-squares intercept was added. This process resulted in a composite map based on the multiparametric data generated in cohort 1 to be validated in cohort 2.

Cohort 2 comprised all patients with a focal recurrence of malignant glioma that was treated with GKRS. Four patients met these criteria: Two patients had multifocal disease, and all potential lesions were evaluated. Radiosurgery is not typically a consideration in patients with recurrent HGG, but 4 exceptional patients experienced a local recurrence, in whom surgical intervention was not deemed appropriate. Across the 4 patients, 8 lesions were considered to potentially represent disease progression. Temporal subtraction maps were generated by registering the patient’s MR imaging scan that last showed stable findings to the first MR imaging scan with detectable disease progression. The recurrence seen in this latter scan was the clinical motivation to treat the patient with GKRS in all 4 patients. The regression model generated in cohort 1 was then applied to the contrast-enhanced T1WI, FLAIR, and ADC temporal subtraction maps from each patient in cohort 2. This resulted in a single z statistic image for each patient in the validation cohort that represented the prediction of areas of the model representing active disease recurrence. A stringent statistical threshold was applied to each z statistic image so that voxels whose correlation to the model was <5 SDs above the mean were filtered.

An additional cluster filter was applied so that only clusters in the top 1% by size were retained. The cluster filter removed numerous individual and small clusters of voxels particularly seen at anatomic boundaries where image coregistration may create artifacts. These thresholds were chosen to eliminate false-positives of the subtraction method, such as small clusters at anatomic boundaries or where differences in timing of gadolinium injection caused differences in vessel filling (On-line Appendix). The full range of unthresholded images is also provided in the On-line Appendix. The product z statistic image was then overlaid with the patient’s actual GKRS treatment plan and post-GKRS MR imaging scans to evaluate the accuracy of the predictions of the model.

Accurate image registration is critical in the generation of temporal subtraction maps. We demonstrate the accuracy of our image registration for all 48 coregistrations (ie, 3 images for each of 16 patients) by using the methods described by Klein et al,20 with details provided in the On-line Appendix and On-line Fig 2.

**RESULTS**

Patient characteristics from both patient groups are outlined in the Table. Cohort 1 consisted of 8 female and 4 male patients with a median age of 61.5 years. Nine patients had a primary diagnosis of GBM (World Health Organization grade IV), and 3 had anaplastic astrocytoma (World Health Organization grade III). Of the latter, 2 had a pathology-proved recurrence with GBM and the remaining patient was retreated on the basis of radiographic progression alone.

The initial course of radiation therapy was delivered at a dose of 6000 cGy in 8 patients, 5940 cGy in 2 patients, and 4005 cGy in 2 patients. The number of days between the end of radiation therapy and the MR imaging showing clear disease progression was a

![Temporal subtraction maps calculated for a representative patient from cohort 1. Column 1 is the MR imaging that was last read as “stable” by the interpreting radiologist. Column 2 is the next follow-up MR imaging demonstrating a nodular recurrence. A focal recurrence is visible in the right posterior frontal lobe with an area of contrast enhancement, increased FLAIR anteromedially, and a small focus of diffusion restriction medially. Automated digital subtraction maps displayed in column 3 demonstrate areas of significant change across the 2 time points for each imaging sequence.](https://www.ajnr.org/content/37/11/2019)
median of 298 days (range, 79–3150 days) measured from the end of radiation therapy. When measured from the last MR imaging scan with “normal” findings (ie, the temporal subtraction interval), the interval was a median of 63 days (range, 28–98 days). The median volume of recurrent tumor was 1.97 mL (range, 0.39–16.41 mL).

In the second cohort (patients A–D in the Table), all 4 patients were men with a median age of 57 years (range, 35–84 years). Three patients were diagnosed with GBM at recurrence, and 1 patient had a primary diagnosis of anaplastic astrocytoma with an aggressive multicentric recurrence that was not amenable to biopsy.

Two patients received 6000 cGy in their initial radiation course, 1 received 5940 cGy, and 1 received 4005 cGy. The recurrence interval was a median of 505 days (range, 233–2961 days), and the temporal subtraction interval was a median of 90 days (range, 50–98 days). The median volume of tumor recurrence was 2.23 mL (range, 0.22–8.88 mL).

ROC curves were generated on the basis of the entire brain, as opposed to a ROI analysis, for each of the 12 patients in cohort 1 and are depicted along with the area under the curve in Fig 2. The mean area under the curve for the leave-one-out analysis was 0.88 with an SD of 0.1. The regression model was sensitive and specific, with an area under the curve of $0.8$ in all except 2 patients.

The regression model was applied to temporal subtraction maps for each of the 4 patients in cohort 2. Results from individual patients are presented to demonstrate the spatial resolution of the model predictions and without depiction of ADC images for brevity.

In patient A, 2 foci of recurrent disease (sum volume of 8.88 mL) were identified. A right inferior temporal lesion was identified as likely recurrent tumor (Fig 3A–B), and voxels within the lesion were highly correlated with the model (Fig 3C). The lesion was treated with GKRS (Fig 3D) and became increasingly necrotic but was stable in size at 35-day (Fig 3E–F) and 66-day (Fig 3G–H) follow-up. Similarly, a left inferior temporal lesion was observed (Online Fig 3A–B), and the model confirmed this as an area of likely recurrent tumor (Online Fig 3C). GKRS to this area was performed (Online Fig 3D), and follow-up imaging demonstrated that the lesion remained stable in size at 35 days (Online Fig 3E–F) and 66 days (Online Fig 3G–H) posttreatment. However, the primary lesion in the right posterior frontal lobe was imaged (Fig 4A–B) and predicted by the model to have active tumor within a circumferential area of the enhancing rim (Fig 4C). This lesion was not treated by GKRS (Fig 4D) and progressed at 35 days (Fig 4E–F). By 66 days (Fig 4G–H), there was a dramatic increase in size with

![FIG 2. ROC curves from the leave-one-out cross-validation. Each curve and its area were computed by generating a model based on 11 patients and applying it to an individual patient.](image-url)
the tumor crossing to the contralateral hemisphere and causing mass effect with a midline shift. Unthresholded axial images are also provided for patient A in On-line Fig 4.

Further demonstrations of the application of the method in 3 patients are provided in the On-line Appendix and On-line Figs 3–17.

**DISCUSSION**

Nearly all patients with a diagnosis of HGG experience local recurrence of their disease despite aggressive tri-modality therapy. Surveillance MR imaging for patients with HGG typically involves a qualitative review of relevant images with careful attention paid to abnormal enhancement, changes in T2 or fluid-attenuated in-
version recovery patterns, and areas of restricted diffusion. A limitation of this qualitative approach is the subjectivity in the relative value (ie, weighting) of each imaging sequence reviewed and the interreader variability that may result. To improve on this qualitative approach, Ellingson et al.21 have applied digital subtraction of pre- from postadolinium T1WI to more accurately isolate areas of enhancing tumor. The investigators demonstrate both the qualitative value of the technique in visualizing tumor and the quantitative relationship between subtraction maps and survival. Subtraction maps were first developed 2 decades ago,22,23 but their application in HGG has not been adopted. More recent advances, such as PWI, have improved surveillance and continue to be investigated, with multiparametric approaches showing particular promise.24,25 PWI was not acquired routinely in patients included in this study, but future analyses would benefit from the addition of this technique.

In this study, we sought to develop an analytic approach that is quantitative, requires minimal subjective user input, makes spatial predictions to guide decisions regarding local therapies, and is based on readily available imaging tools. To this end, we identified a homogeneous group of patients with a well-defined spatial recurrence and a comprehensive clinical assessment leading to the recommendation for retreatment with radiation therapy.

Analysis of these data revealed a consistent pattern of change in T1WI, FLAIR, and ADC signal intensities. When the regression model was applied to a separate cohort of patients who were treated with conformal radiosurgery, spatially discernable areas at risk of tumor progression were identified with high specificity. A leave-one-out validation was performed in the 12 patients used to develop the regression model, which confirmed its high sensitivity and specificity. We then performed a diagnostic test of the model in a cohort of 4 patients treated with GKRS. Results were not sent to the PACS because of the investigational nature of the study, but results can be stored in a format that would permit using PACS in the future.

GKRS is an atypical method to treat HGG because this disease tends to recur diffusely but can occasionally recur focally. This unique group of patients is an ideal cohort in which to test the application of temporal subtraction maps because disease progression within and outside the highly conformal GKRS dose distribution can be monitored with time.

Across the 4 patients in cohort 2, the model predicted that 5 lesions were at high risk of recurrence and these were treated with GKRS. Follow-up imaging demonstrated the stability of all 5 lesions for at least 64 days (ie, 66 days in patient A, 248 days in patient B, 95 days in patient C, and 64 days in patient D). The model also predicted active tumor in the enhancing rim or adjacent to the primary lesion in 2 patients. These areas were not treated with GKRS and were found to have progressed at the next available imaging follow-up (35 days in patient A and 53 days in patient C).

Limitations of this study are its retrospective nature, small patient size due to strict inclusion criteria, and the analytic methods needing to be validated prospectively before being applied to clinical practice more generally. A further limitation is the reliance on image registration to develop temporal subtraction maps. Poor coregistration leads to artifacts in the z statistic map, particularly at anatomic boundaries such as the brain outline. We addressed this limitation by creating 2 stringent statistical filters, including a z score threshold of 5 SDs above the mean and a spatial cluster threshold leaving only the top 1 percentile. While this feature serves to confirm the robust performance of the regression model on temporal subtraction maps within the brain parenchyma, there may be increased type II errors, particularly if the suspected recurrence is small or at the brain surface. Our assessment of registration accuracy demonstrates that even with the generally high agreement with linear registration in FSL, there were 3 outliers (of 48 registrations) in which the registration was poor. While this automated analytic tool may help reduce subjective bias, the potential for false-negative results and reliance on accurate image registration highlight its role as a supplement to clinical judgment rather than a fully automated method of treatment planning.

An additional limitation is the reliance on the radiologist’s opinion as the “ground truth” in our analysis. Although more than half of the recurrences in cohort 1 were proved by pathology, a pathologic diagnosis was not obtained in the 4 patients in cohort 2, and this omission may increase susceptibility to false-positive results. While this caveat is important to consider, our inclusion criteria required that patients be selected for re-irradiation on the basis of a comprehensive clinical assessment. This decision was made as part of institutional policy and was not related to the present study, which should mitigate the presence of false-positive results.

CONCLUSIONS

We applied a data-driven model to temporal subtraction maps of gadolinium-enhanced T1WI, FLAIR, and ADC images in patients with recurrent HGG. This automated method could be used to discern tumor regions that may benefit from additional local therapy, such as conformal radiosurgery. Future work will validate this approach prospectively to facilitate broader clinical application.


*Money paid to the institution.

REFERENCES

5. Barnard RO, Geddes JF. The incidence of multifocal cerebral


