Early post-bevacizumab progression on contrast-enhanced MRI as a prognostic marker for overall survival in recurrent glioblastoma: results from the ACRIN 6677/RTOG 0625 Central Reader Study

Jerrold L. Boxerman, Zheng Zhang, Yair Safriel, Mykol Larvie, Bradley S. Snyder, Rajan Jain, T. Linda Chi, A. Gregory Sorensen†, Mark R. Gilbert, and Daniel P. Barboriak

Department of Diagnostic Imaging, Rhode Island Hospital and Alpert Medical School of Brown University, (J.L.B.), and Center for Statistical Sciences, Brown University, Providence, Rhode Island (Z.Z., B.S.S.); Pharmascan Clinical Trials and Radiology Associates of Clearwater—University of South Florida, Clearwater, Florida (Y.S.); Department of Radiology, Massachusetts General Hospital, Boston, Massachusetts (M.L.); Departments of Radiology and Neurosurgery, Henry Ford Hospital, Detroit, Michigan (R.J.); Department of Diagnostic Radiology, The University of Texas MD Anderson Cancer Center, Houston, Texas (L.C.); A. A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charlestown, Massachusetts (A.G.S.); Department of Neuro-Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas (M.R.G.); Department of Radiology, Duke University Medical Center, Durham, North Carolina (D.P.B.)

†Present affiliation: Siemens Healthcare, Malvern, Pennsylvania.

Background. RTOG 0625/ACRIN 6677 is a multicenter, randomized, phase II trial of bevacizumab with irinotecan or temozolomide in recurrent glioblastoma (GBM). This study investigated whether early posttreatment progression on FLAIR or postcontrast MRI assessed by central reading predicts overall survival (OS).

Methods. Of 123 enrolled patients, 107 had baseline and at least 1 posttreatment MRI. Two central neuroradiologists serially measured bidimensional (2D) and volumetric (3D) enhancement on postcontrast T1-weighted images and volume of FLAIR hyperintensity. Progression status on all posttreatment MRIs was determined using Macdonald and RANO imaging threshold criteria, with a third neuroradiologist adjudicating discrepancies of both progression occurrence and timing. For each MRI pulse sequence, Kaplan-Meier survival estimates and log-rank test were used to compare OS between cases with or without radiologic progression.

Results. Radiologic progression occurred after 2 chemotherapy cycles (8 weeks) in 9 of 97 (9%), 9 of 73 (12%), and 11 of 98 (11%) 2D-T1, 3D-T1, and FLAIR cases, respectively, and 34 of 80 (43%), 21 of 58 (36%), and 37 of 79 (47%) corresponding cases after 4 cycles (16 weeks). Median OS among patients progressing at 8 or 16 weeks was significantly less than that among nonprogressors, as determined on 2D-T1 (114 vs 278 days and 214 vs 426 days, respectively; \( P < .0001 \) for both) and 3D-T1 (117 vs 306 days \( P < .0001 \) and 223 vs 448 days \( P = .0003 \), respectively) but not on FLAIR (201 vs 276 days \( P = .38 \) and 303 vs 321 days \( P = .13 \), respectively).

Conclusion. Early progression on 2D-T1 and 3D-T1, but not FLAIR MRI, after 8 and 16 weeks of anti-vascular endothelial growth factor therapy has highly significant prognostic value for OS in recurrent GBM.

Keywords: bevacizumab and irinotecan, imaging biomarker, overall survival, progression, recurrent glioblastoma.
Glioblastoma multiforme (GBM), the most common and aggressive primary brain tumor, has a dismal prognosis with median survival of 12–15 months\(^1\) and a 5-year survival rate of 9.8\%.\(^2\) Surgical resection followed by chemoradiation with concomitant and adjuvant temozolomide (TMZ) is the standard of care offering prolonged overall survival (OS),\(^3\) but progressive GBM is uniformly fatal, with median survival <30 weeks.\(^4\) GBM typically overexpresses vascular endothelial growth factor (VEGF), yielding increased vascular permeability and tumor angiogenesis\(^5\) and prompting the use of antiangiogenic drugs in clinical treatment trials. Bevacizumab, a humanized monoclonal antibody to VEGF,\(^6\) is Food and Drug Administration–approved for second-line treatment of recurrent GBM,\(^7\) conferring progression-free survival (PFS) benefit in clinical trials,\(^8\) compared with historic controls.\(^3\) Combination therapy for recurrent GBM, including bevacizumab with irinotecan (a topoisomerase I inhibitor),\(^9\) may provide additional benefit,\(^10,11\) with relatively high objective response (OR), PFS, and OS.\(^12–15\) However, not all recurrent GBMs respond to antiangiogenic therapy, and response in this setting is poorly defined.\(^1\) Radiologic biomarkers that provide early indication of treatment failure and timely opportunity to select alternative treatments or trials are therefore important.

The MacDonald criteria\(^16\) applied to contrast-enhanced T1-weighted MRI has historically been the standard for determining treatment response of GBM, with the recent RANO criteria\(^17\) additionally recognizing the potential for FLAIR hyperintense tumor progression to precede progressive tumor enhancement. Although increased OR in recurrent GBM treated with antiangiogenic agents has revived interest in OR as a predictor of survival benefit,\(^18\) antiangiogenic agents create problems for the imaging evaluation of posttreatment GBM.\(^19\) VEGF blockade decreases vascular permeability and contrast enhancement but may not affect underlying tumor, and this pseudoresponse may limit the ability of OR to predict OS.\(^20,21\) Conversely, chemoradiation preceding anti-VEGF administration may induce pseudoprogressive enhancement\(^22\) persisting into early post–bevacizumab treatment imaging evaluation. The imaging assessment of response to antiangiogenic agents in recurrent GBM is therefore complicated, with an uncertain and controversial relationship between OR and OS. Whereas some studies have found that early posttreatment OR status predicts OS,\(^12,23\) others have not.\(^24\) The use of FLAIR in RANO\(^25\) and the relationship between isolated FLAIR progression and OS\(^24,26\) is also controversial.

RTOG 0625, a multicenter, randomized, phase II trial of bevacizumab with irinotecan or temozolomide in recurrent GBM, provided the opportunity to evaluate early posttreatment conventional MRI metrics for predicting OS using controlled, retrospective central reader methodology in a cohort of patients on antiangiogenic agents. ACRIN 6677 is the companion study that evolved, with an advanced MRI component to be reported separately. This study investigated whether progression on postcontrast 2D- or 3D-T1 or FLAIR MRI after 2 or 4 anti-VEGF chemotherapy cycles (8 or 16 weeks after therapy) is predictive of OS. We hypothesized that early posttherapy progression would be highly correlated with reduced OS and serve as a useful MRI biomarker for failed anti-VEGF therapy.

**Materials and Methods**

The Radiation Therapy Oncology Group (RTOG), in collaboration with the American College of Radiology Imaging Network (ACRIN), both funded by the National Cancer Institute, conducted a prospective, randomized, phase II multicenter trial comparing bevacizumab with either irinotecan or temozolomide treatment in recurrent GBM (RTOG 0625/ACRIN 6677). Twenty-three institutions participated, each obtaining institutional review board approval before subject accrual and conducting the trial with Health Insurance Portability and Accountability Act compliance. Informed consent was obtained for all subjects.

**Study Subjects**

All patients had recurrent histologically proven GBM or gliosarcoma with progression on MRI within 14 days after registration, ≥42 days after completion of radiation/temozolomide therapy, ≥28 days after surgical resection or cytotoxic chemotherapy, as well as imaging or biopsy confirmation of true progressive disease rather than radiation necrosis after Gliadel placement or stereotactic radiosurgery. Detailed inclusion and exclusion criteria are available at http://www.acrin.org/Portals/0/Protocols/6677/RTOG062-ACRIN6677.pdf (Section 3.0). Bevacizumab was administered to all patients (10 mg/kg intravenously, days 1 and 15 of a 28-day cycle). In the first arm, patients received temozolomide (75 mg/m\(^2\) per os, days 1–21 during the first 28-day cycle; 100 mg/m\(^2\) for cycle 2 and beyond in the absence of myelotoxicity). In the second arm, patients received irinotecan (125 mg/m\(^2\) intravenously, days 1 and 15 of a 28-day cycle). Standard of care MRI occurred at baseline, after every 2 cycles of treatment (every 8 weeks), and after completion or termination of treatment. Patients demonstrating benefit (stable or responding tumor) were treated for 12 cycles with optional extension to 24 cycles in the presence of continued benefit and absence of severe toxicity.

**MRI Protocol**

Conventional MRI included precontrast T1-weighted, T2-weighted, FLAIR, and diffusion-weighted imaging. After intravenous injection of 0.1 mmol/kg of standard gadolinium-based agent, axial 2D spin-echo (2D-T1) and 3D volumetric (3D-T1) T1-weighted (post-Gd) images were acquired. Patients participating in the optional advanced component of the trial had dynamic contrast-enhanced MRI, dynamic susceptibility contrast perfusion-weighted MRI, and/or MR spectroscopy at baseline, week 2, and after every 2 cycles of treatment. Complete MRI
parameters for this protocol are listed on the ACRIN website (http://www.acrin.org/6677_protocol.aspx, Imaging Transmittal Worksheet in Imaging Materials link).

Central Reader Methods

All local imaging was retrospectively transmitted to ACRIN for central review. Two primary readers and one adjudicator, each with neuroradiology Certificates of Added Qualification and 8, 6, and 3 years of postfellowship experience, respectively, were trained via teleconference about 2D measurement and 3D contouring techniques. Each primary reader was assigned 2 similarly trained core laboratory technologists and conducted independent image assessments. 2D image analysis was performed using ClearCanvas Workstation (ClearCanvas, Inc., Toronto, ON). For each distinct contrast-enhancing target lesion as defined by Macdonald and RANO criteria (≥1 cm diameter, ≥1 cm from other enhancing lesions), the largest diameter of contrast enhancement and its maximum perpendicular diameter in the same plane were measured. 2D tumor area was computed by summing over all lesions the product of maximum perpendicular diameters. Segmentation of FLAIR and 3D-T1 images (pre-contouring by technologists and real-time review by neuroradiologists) was performed with MIMvista (MIM Software Inc., Cleveland, OH) using absolute thresholding, relative threshold of maximum, and region growing algorithms. FLAIR and 3D-T1 volumes were computed by summing the segmented voxel volumes. Pre- and post-Gd images were reviewed simultaneously to exclude blood products from 2D areas and 3D segmentations.

For all evaluable patients, images at each available time point were presented in random order to both central readers who independently made 2D-T1 measurements and approved or modified the 3D-T1 and FLAIR contours prepared by core laboratory technologists. After completing measurements and contours for all time points, readers were unblinded to the order of exams. Using the same thresholds for 2D-T1 prescribed by Macdonald and RANO criteria, each reader determined time of progression on 2D-T1, 3D-T1, and FLAIR when there was >25% increase with respect to nadir in maximal cross-sectional enhancing areas, volume of segmented enhancement, and sum of 2D contoured hyperintensity, respectively, or the appearance of any new enhancing tumor.16,17 Similarly, radiologic response was defined as ≥50% decrease with respect to baseline, confirmed on the subsequent time point. Steroid dosage and clinical status were unavailable to ACRIN readers for this study. The adjudicator settled discordant times to progression between primary readers on 2D-T1, 3D-T1, and FLAIR by selecting the times to progression that were most correct in his opinion.

Statistical Methods

The primary objective of this study was to compare OS between patients who progressed on imaging by 8 weeks after treatment and those who did not. The analysis was done separately for 2D-T1, 3D-T1, and FLAIR and repeated for the imaging progression status at 16 weeks after treatment. We used Kaplan-Meier survival estimates and the log-rank test for comparisons and analyzed pooled data regardless of treatment assignment. Survival time was calculated from the trial enrollment date.

We further classified patients who did not progress by week 8 or 16 (nonprogressors) for each MRI method into 2 groups: those with partial or complete radiologic response (responders) and those without (non-responders and nonprogressors [NR-NPs]). Patients were grouped as responders only if both readers rated partial or complete response by the specified time point. We compared OS between responders and NR-NPs.

We stratified patients who had not progressed on T1 at 8 weeks into those who progressed on FLAIR (isolated FLAIR progressors) and those who did not (nonprogressors) and compared OS between isolated FLAIR progressors and FLAIR nonprogressors. The analysis was done separately for 2D-T1 and 3D-T1 and repeated for 16-week data.

Statistical computations were performed using SAS, version 9.2 (SAS Institute, Cary, NC), with P values <.05 considered to be statistically significant.

Results

Study Cohort

One hundred twenty-three patients were enrolled (71 men, 52 women; age range, 23–87 years; median age, 56 years). Excluding 4 ineligibles, 1 lost to follow-up, and 11 with only baseline imaging, we included 107 evaluable patients with baseline and at least 1 posttreatment MRI. Interpretable FLAIR, 2D-T1, and 3D-T1 images were available for 107, 103, and 76 patients, respectively.

Adjudication Rate

The adjudication rates for time of progression were 43% (45 of 105 patients) for 2D-T1, 42% (32 of 76) for 3D-T1, and 39% (42 of 107) for FLAIR. Excluding patients missing relevant interpretable scans or with precedent death, there was adjudicated radiologic progression at 8 weeks in 9 of 97 (9%), 9 of 73 (12%), and 11 of 98 (11%) evaluable 2D-T1, 3D-T1, and FLAIR cases, respectively, and in 34 of 80 (43%), 21 of 58 (36%), and 37 of 79 (47%) corresponding cases at 16 weeks.

Relationship between Radiologic Progression and OS

Thirteen (12%) of the 107 included patients were alive at the time of analysis. The estimated median survival was 270 days (95% confidence interval [CI], 217–309 days). Fig. 1 compares Kaplan-Meier survival curves for patients by progression status on 2D-T1, 3D-T1, and FLAIR at 8 weeks and 16 weeks after initiation of anti-VEGFR therapy. At both 8 and 16 weeks, there was a significant difference between survival curves on 2D-T1 (P < .0001 for both) and 3D-T1 (P < .0001 and
but not on FLAIR (P = .38 and P = .13). The median survival among patients with progression at 8 or 16 weeks was significantly less than that among patients without progression on 2D-T1 (114 vs 278 days and 214 vs 426 days, respectively) and 3D-T1 (117 vs 306 days and 223 vs 448 days, respectively), but not on FLAIR (201 vs 276 days and 303 vs 321 days, respectively). Table 1 summarizes these results.

Substratification of Radiologic Nonprogressors

Table 2 summarizes overall survival results by each MRI method after further stratifying patients who had not progressed at 8 or 16 weeks. There was no statistically significant survival benefit among responders, compared with NR-NPs, at 8 or 16 weeks on any pulse sequence. Fig. 2 compares Kaplan-Meier survival curves for progressors, responders, and NR-NPs on 2D-T1 and 3D-T1 at 8 and 16 weeks after initiation of anti-VEGF therapy. Although not statistically significant, there was better visual separation of Kaplan-Meier curves between responders and NR-NPs for 3D-T1 than for 2D-T1.

Additional Contribution of FLAIR beyond 2D-T1 and 3D-T1

By using the progression status on FLAIR to further classify the nonprogressors on 2D-T1 into nonprogressors
and isolated FLAIR progressors, we found no statistically significant survival time reduction among isolated FLAIR progressors, compared with nonprogressors, at both 8 ($P = .763$) and 16 ($P = .318$) weeks. Similar results were obtained for 3D-T1 ($P = .919$ and $P = .126$).

Table 3 summarizes these results. Fig. 3 compares Kaplan-Meier survival curves for nonprogressors and isolated FLAIR progressors on 2D-T1 and 3D-T1 at 8 and 16 weeks after initiation of anti-VEGF therapy.

**Table 3.** Comparison of median survival time by progression status on MRI performed 8 and 16 weeks after initiation of anti-VEGF therapy

<table>
<thead>
<tr>
<th>MRI Method</th>
<th>$n$</th>
<th>Progression by week 8</th>
<th>Progression by week 16</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No (n (%)) Median survival (95% CI)</td>
<td>Yes (n (%)) Median survival (95% CI)</td>
</tr>
<tr>
<td>2D-T1</td>
<td>105</td>
<td>88 (84)</td>
<td>278 (234, 327)</td>
</tr>
<tr>
<td>3D-T1</td>
<td>76</td>
<td>64 (84)</td>
<td>306 (243, 409)</td>
</tr>
<tr>
<td>FLAIR</td>
<td>107</td>
<td>87 (81)</td>
<td>276 (223, 327)</td>
</tr>
</tbody>
</table>

Median survival time is given in days with 95% confidence interval in parentheses. N, baseline analyzable cases. Bold $P$ values are statistically significant.

**Table 2.** Comparison of median survival time by response status in patients who did not progress on MRI performed 8 and 16 weeks after initiation of anti-VEGF therapy

<table>
<thead>
<tr>
<th>MRI Method</th>
<th>Response Status, Week 8</th>
<th>Response Status, Week 16</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$ (%) Median survival (95% CI)</td>
<td>$n$ (%) Median survival (95% CI)</td>
</tr>
<tr>
<td>2D-T1</td>
<td>23 (26)</td>
<td>65 (74)</td>
</tr>
<tr>
<td>3D-T1</td>
<td>21 (33)</td>
<td>43 (67)</td>
</tr>
<tr>
<td>FLAIR</td>
<td>13 (15)</td>
<td>74 (85)</td>
</tr>
</tbody>
</table>

Median survival time is given in days with 95% confidence interval in parentheses. R, responder (patients with complete or partial response according to both central readers); NR-NP, nonresponder, non-progressor (patients who did not progress and who did not meet criteria for responder).

Discussion

Our results demonstrate that radiologic progression of recurrent GBM on post-Gd T1-weighted MRI after 2 (8 weeks) or 4 (16 weeks) cycles of anti-VEGF therapy (bevacizumab plus irinotecan or temozolomide) is a prognostic marker of poor survival. Progressive enhancement after early anti-angiogenic treatment, despite the potential for late pseudoprogession from initial chemoradia- tion, is associated with poorer survival relative to stable or regressed enhancement, despite the potential for pseudoresponse.

Pseudoprogession and pseudoresponse present challenges for the imaging assessment of posttreatment glioma.19 Pseudoprogession, a common local inflammatory reaction after radiotherapy and temozolomide marked by progressive enhancement on early postradiotherapy imaging and subsequent radiologic improvement or stabilization without modified therapy,22,25–33 may confound the interpretation of progressive enhancement. Reported peak observation times range from 1–6 months (60% within 3 months) after treatment completion.30,31,33 –35 Because at least 42 days passed between chemoradiation and enrollment in our study, >3 months should have elapsed between radiation and the 8-week MRI and at least 5 months before the 16-week MRI. At 8 weeks, all patients with progressive 2D-T1 (9 of 9) and 3D-T1 (9 of 9) enhancement had completed initial radiation 180 days before progression, as had 33 of 34 and 20 of 21 patients at 16 weeks on 2D-T1 and 3D-T1, respectively. Therefore, our progressive enhancement likely reflects true tumor progression rather than late-onset pseudoprogession. Because pseudoprogession is associated with increased survival,30,33 –35 inclusion of pseudoprogessors with true progressors may artificially increase the median survival time in this group. To the extent that this occurs, we may underestimate the actual OS disparity between true imaging progressors and nonprogressors.

Pseudoresponse is observed with anti-angiogenic agents, such as bevacizumab, that reduce contrast agent extravasation and discernible tumor enhancement independent of cytotoxic or cytostatic effect.36,37 FLAIR
hyperintensity may progress after anti-angiogenic therapy, despite decreased or even resolved tumor-related contrast enhancement.\textsuperscript{20,21} The possibility that progressive FLAIR abnormalities represent nonenhancing tumor growth could explain why tumor enhancement-based OR may be incongruous with and a poor predictor of OS.\textsuperscript{38} The effect of pseudoresponse in our study remains unclear, because distinguishing patients with diminished enhancement because of anti-VEGF effects alone from those with both anti-VEGF and tumor cytotoxic effects is not possible. This study demonstrates that, without pseudoprogression, progressive enhancement strongly implicates true tumor progression, thereby reflecting either lack of effective anti-VEGF response and/or failure of treatment to inhibit tumor growth. Analogous to anti-VEGF effect, steroid-induced decreased enhancement could similarly mask true progressive disease. Classification of pseudoresponders from bevacizumab or steroids with nonprogressors may actually reduce an otherwise more striking OS.

Fig. 2. Kaplan-Meier survival curves stratified by radiologic progression status on 2D-T1 (top row) and 3D-T1 (bottom row) imaging at 8 weeks (left column) and 16 weeks (right column) after initiation of anti-VEGF therapy, with substratification of patients who had not progressed (nonprogressors) into those showing partial or complete response (responders) and those without response (nonresponder, nonprogressors [NR-NPs]). Listed $P$ values are between responders and NR-NPs. At both 8 and 16 weeks, there was no statistically significant difference between the survival curves for responders (dot-dash curves) and NR-NPs (solid curves) on 2D-T1 or 3D-T1. Although not statistically significant, there is better visual separation of Kaplan-Meier curves between the responders and NR-NPs for 3D-T1 than for 2D-T1.

Table 3. Comparison of median survival time for T1 nonprogressors by progression status on FLAIR MRI performed 8 and 16 weeks after initiation of anti-VEGF therapy

<table>
<thead>
<tr>
<th>T1 Method</th>
<th>Progression Status, Week 8 (T1, FLAIR)</th>
<th>Progression Status, Week 16 (T1, FLAIR)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NP, NP $n (%)$</td>
<td>Median survival (95% CI)</td>
</tr>
<tr>
<td>2D-T1</td>
<td>82 (78)</td>
<td>6 (6)</td>
</tr>
<tr>
<td></td>
<td>283 (232, 330)</td>
<td>255 (119, 971)</td>
</tr>
<tr>
<td>3D-T1</td>
<td>60 (79)</td>
<td>4 (5)</td>
</tr>
<tr>
<td></td>
<td>306 (243, 409)</td>
<td>375 (119, 971)</td>
</tr>
</tbody>
</table>

Median survival time is given in days with 95% confidence interval in parentheses. NP, nonprogression; P, progression.
discrepancy between progressors and nonprogressors, although our observed separation is still highly statistically significant.

At 8 weeks, most patients were classified as nonprogressors on 2D-T1 and 3D-T1; a more equal split between progressors and nonprogressors was observed at 16 weeks. Using Macdonald and RANO imaging thresholds, we further stratified the nonprogressors into responders and NR-NPs, to determine whether imaging response had prognostic value for OS among those patients who were not shown to progress. We found no significant survival benefit among responders, compared with NR-NPs, at 8 or 16 weeks on any pulse sequence. Whereas progressive contrast enhancement appears to be useful for identifying patients likely to do poorly with bevacizumab treatment regardless of clinical status or steroid use, regressive enhancement appears to be ineffective for subselecting those nonprogressors who are likely to do particularly well, although survival curves were better visually separated for 3D-T1 than for 2D-T1. This raises the possibility that volumetric technique may suffer less from the pitfalls of pseudoresponse than bidimensional technique. Adjunctive diffusion, perfusion, or spectroscopic MRI, combined with 3D-T1, may more favorably separate the nonprogressors and overcome the apparent shortcomings of conventional contrast-enhanced imaging.

The lack of survival benefit for responders in our study would seem to contradict the results of Prados et al., who demonstrated that early OR status was a statistically significant predictor of survival among similarly treated patients with recurrent GBM. However, it is important to clarify differences in the use cases for these studies with correspondingly different statistical design. Whereas Prados et al. primarily correlated OR with OS to identify treatment successes and compared OS between responders and all nonresponders (NR-NPs plus progressors), we primarily correlated radiographic progression and OS to identify treatment failures and secondarily compared OS between responders and NR-NPs only. Because progressors have significantly different OS than nonprogressors, the addition by Prados et al. of progressors to the NR-NPs likely accounts for the conflicting results concerning significance of OR status. We are unaware of other reports with a similar analysis.

We determined progression and response on postcontrast 2D-T1 and 3D-T1 images with use of the thresholds for 2D tumor measurements adopted by the Macdonald and RANO criteria, the most widely used criteria for evaluating tumor response in clinical trials. The anti-permeability effect of bevacizumab has prompted redefinition of response criteria, with RANO recognizing the nonspecificity of changed enhancement because of pseudoprogression and pseudoresponse and the potential...
for FLAIR progression to precede progressive tumor enhancement. To investigate the potential contribution of FLAIR to modified response criteria, we evaluated the ability of isolated FLAIR progression to predict OS. For simplicity and consistency, we applied the same threshold for T1 progression (25% increase, compared with nadir) to our determination of FLAIR progression, although lesion progression on FLAIR is not quantified by RANO and there is evidence that lower thresholds are more sensitive for isolated FLAIR progression.25 At 8 or 16 weeks, survival curves for isolated FLAIR progressors are not significantly different from those for FLAIR and T1 non-progressors, and thus, isolated FLAIR progression did not portend a significant survival disadvantage compared with nonprogression. These findings support previous literature24,43 and question the prognostic value of early posttreatment isolated FLAIR progression, although different thresholds or heuristics for determining FLAIR progression may yield different conclusions. Similarly, compared with total nonprogressors, we found less separation of OS for progressors on T1 or FLAIR (mimicking inclusion of FLAIR in the modified RANO criteria) than for isolated T1 progressors (data not shown). At 8 weeks, for example, we found 6 FLAIR progressors who had not progressed on T1, in addition to the 9 T1 progressors. Because survival time for isolated FLAIR progressors is not significantly different than that for nonprogressors (Table 3), mixing these patients with T1 progressors reduces OS separation. Although the modified RANO criteria identify more imaging progressors by including FLAIR, our results suggest that, when using this scheme, the specificity of imaging progression for predicting outcome may be diminished.

It should be emphasized that, although we used the Macdonald and RANO threshold criteria for defining imaging progression and response, ACRIN readers were blinded to the clinical status and steroid administration, and we therefore report on pure imaging biomarkers relating to OS. Although our study, in principle, provides an opportunity to compare the Macdonald and RANO criteria and to possibly validate the RANO criteria, we begin to address this issue by evaluating the potential additional value of including changes in FLAIR in defining progression and acknowledge that this comparison would be fertile ground for a secondary analysis in a follow-up article.

Radiologic assessment of bevacizumab-treated recurrent GBM is impacted by suboptimal delineation of irregularly shaped and poorly enhancing tumors, interobserver variability, and lack of guidelines for assessment of FLAIR and multifocal disease.48 We used blinded central imaging review with structured reader training to reduce bias and subjectivity in tumor measurement and adjudication of discrepant interpretations to address interreader variability. Nonetheless, we still report relatively high adjudication rates, although slightly improved from prior adjudicated bevacizumab trials, emphasizing the inherent subjectivity of bidimensional and 3D measurements in post-bevacizumab treatment GBMs with heterogeneous, ill-defined contrast enhancement patterns. High adjudication on FLAIR may relate to poorly defined inclusion criteria and superimposed leukoaraiosis and posttreatment effects.49 The use of an adjudicator is likely to be very important and may partially explain some discrepancies in the literature among studies not using adjudication of central interpretations.

Because of the geometric complexity of contrast enhancement and FLAIR hyperintensity associated with recurrent GBMs, volumetric evaluations may improve accuracy, and computerization of such measures may lessen interobserver variability.48 However, previous studies demonstrated concordance between 2D-T1 and 3D-T1 methods for determining radiologic response of newly diagnosed and recurrent high-grade gliomas,10,31 and our results show essentially similar performance of 2D-T1 and 3D-T1 at both 8 and 16 weeks. For purposes of OS prognostication based on progression status, both methods would appear to suffice, although our results suggest, but do not conclusively establish, that volumetric T1 technique may be better suited for prognostication based on response status of nonprogressors.

In conclusion, early progression on post-Gd 2D-T1 and 3D-T1, but not FLAIR MRI, after 2 and 4 cycles (8 and 16 weeks) of anti-VEGF therapy had highly significant prognostic value for OS. Because of the controversies about imaging criteria in brain tumor trials, these results provide clear evidence supporting identification of treatment failures with use of contrast-enhanced T1-weighted images in patients with recurrent GBM, even when anti-angiogenic agents are used in treatment. When attempting to further stratify the nonprogressors, we found no statistically significant survival benefit for responders, compared with NR-NPs, at 8 or 16 weeks on any pulse sequence. Although FLAIR is included in the modified response criteria for treated gliomas to address the issue of pseudoresponse, we found no statistically significant OS reduction for isolated FLAIR progressors, compared with nonprogressors.

**Funding**

ACRIN receives funding from the National Cancer Institute through grants U01-CA080098 and U01-CA079778.

**Acknowledgments**

We thank Drs. Benjamin Corn, Deborah Blumenthal, Steven Chmura, R. Jeffrey Lee, Joseph Simpson, and Linda Grossheim, accrual efforts; Drs. Tom Mikkelsen, Orna Aizenstein, Michael Vannier, Duane Blatter, Robert McKinstry, and Scott Rand, for imaging coordination and data acquisition efforts; Bernadine Dunning, James Gimpel, Kesha Warren, Sandy Toland-Cary, Cyndi Price, and Maxine Crooks from ACRIN, for administrative and technologist support; and all study patients, for their cooperation.

**Conflicts of interest statement.** D.B. has received laboratory support from Eisai Pharmaceuticals, Siemens.
Healthcare, and Adnexus Therapeutics; has an advisory affiliation with GE Medical Systems and neuro-MRI advisory board; and has served as a consultant for GE Medical Systems and neuro-MRI advisory board. G.S. has stock ownership in Siemens, has received research funding from ACRIN, has advisory affiliations with Genentech and Siemens Healthcare N.A., and has received honoraria from Genentech.

M.G. has received research funding from GlaxoSmithKline, Merck, and Genentech; has advisory affiliations with Genentech and Merck; and has received honoraria from Genentech, Merck, and Abbott.

R.J. has received research funding from Lantheus Medical Imaging, Inc. and has served on the MRI Steering Committee for Medscape Education.

References


