Response Classification Based on a Minimal Model of Glioblastoma Growth Is Prognostic for Clinical Outcomes and Distinguishes Progression from Pseudoprogression

Maxwell Lewis Neal1, Andrew D. Trister2,6, Sunyoung Ahn1, Anne Baldock1,7,8, Carly A. Bridge1,7,8, Laura Guyman1,7,8, Jordan Lange1, Rita Sodt1, Tyler Cloke1, Albert Lai9, Timothy F. Cloughesy9, Maciej M. Mrugala3, Jason K. Rockhill2, Russell C. Rockne1,4,7,8, and Kristin R. Swanson1,4,5,7,8

Abstract

Glioblastoma multiforme is the most aggressive type of primary brain tumor. Glioblastoma growth dynamics vary widely across patients, making it difficult to accurately gauge their response to treatment. We developed a model-based metric of therapy response called Days Gained that accounts for this heterogeneity. Here, we show in 63 newly diagnosed patients with glioblastoma that Days Gained scores from a simple glioblastoma growth model computed at the time of the first postradiotherapy MRI scan are prognostic for time to tumor recurrence and overall patient survival. After radiation treatment, Days Gained also distinguished patients with pseudoprogression from those with true progression. Because Days Gained scores can be easily computed with routinely available clinical imaging devices, this model offers immediate potential to be used in ongoing prospective studies. Cancer Res; 73(10); 1–11. ©2013 AACR.

Introduction

Glioblastoma multiforme is a highly aggressive and invasive type of primary brain tumor. The current standard of care for glioblastoma includes multimodality therapy with surgery, radiation, and chemotherapy; yet, despite these interventions, newly diagnosed patients with glioblastoma have a dismal median survival time of 14.6 months (1). Throughout the period following standard therapy, decisions to change treatment are made based on the patient’s clinical symptoms and radiographic studies; however, the heterogeneity in the spatial distribution and growth dynamics of glioblastoma, combined with interpatient variability in treatment response, makes these decisions difficult. In particular, practical tools that incorporate these patient-specific differences into assessing the impact of treatment response on prognosis are lacking. In addition, patients frequently exhibit changes on imaging that mimic progression (pseudoprogression or radiation changes) in the 2 to 6 months following chemoradiation but that later stabilize or improve (2–4). Distinguishing these patients from those with true progression is difficult but crucial for providing the most appropriate therapy.

Motivated by these challenges, we developed a new response metric called "Days Gained" that uses patient-specific computational models of glioblastoma dynamics to estimate the amount of time a given therapy delayed tumor growth (Fig. 1). Days Gained differs from existing, commonly used response metrics such as the Macdonald (5), Response Evaluation Criteria in Solid Tumors (RECIST; ref. 6), and Response

Major Findings

A new response metric for glioblastoma that is based on a simple linear model of tumor growth is strongly prognostic of progression-free and overall survival and distinguishes pseudoprogression from true progression immediately following first-line radiotherapy.
Assessment in Neuro-Oncology (RANO) criteria (7) in that it incorporates, among other patient-specific data, a measure of the pretreatment tumor growth rate. This rate allows one to compute the trajectory of a simulated, untreated virtual control (UVC) and predict the tumor size at posttreatment time points as if it had remained untreated. Whereas current response criteria compare pretreatment with posttreatment tumor burdens on MRI to assess a patient’s response to therapy, our method compares the posttreatment tumor burden with the expected tumor burden predicted by our model at that same time point (Fig. 1). This provides a more complete, accurate, and personalized assessment of a therapy’s effectiveness. For example, a patient with a fast-growing tumor that responds robustly may show a similar pretreatment-to-posttreatment change in tumor burden on MRI as that of a patient with a slow-growing tumor that responds poorly. Accounting for the tumor’s rate of growth provides a means of avoiding such false negatives.

The creation of existing response metrics has been largely motivated by the need to standardize the clinical discourse around response to therapy, and their strength lies more in their definitions of response and progression than in their predictive capabilities. Our purpose in creating the Days Gained metric was to develop a quantitative measure of response that is prognostic for patient outcomes, one that can aid clinical decision making by translating measurable treatment response to survival benefit. Our development of this metric was guided by the intuition that, given the heterogeneity of glioblastoma dynamics, maximizing a response metric’s prognostic power will require incorporating as much

### Quick Guide to Main Model Equations

The minimal "linear" model presented here assumes that an untreated tumor’s spherically equivalent radius on T1 gadolinium-enhanced (T1Gd) imaging (the radius assuming a spherical geometry of the measured tumor volume) increases linearly with time:

\[ y = mx + y_{PRE} \]

Here, \( y \) is the simulated spherically equivalent T1Gd radius, \( x \) is the amount of time since the final T1Gd MRI taken before surgical resection and chemoradiation treatment, \( m \) is the rate of untreated T1Gd radial growth computed from sequential pretreatment scans, and \( y_{PRE} \) is the spherically equivalent radius computed from the volumetric tumor burden observed on the second pretreatment T1Gd MRI.

To compute Days Gained scores from the linear model, we solve for \( x \) in the equation above when \( y \) is the measured posttreatment T1Gd spherically equivalent radius (\( y_{POST} \)), then subtract this value from the number of days that have passed since the final pretreatment scan (\( x_{POST} \)).

\[ \text{Days Gained} = x_{POST} - \frac{y_{POST} - y_{PRE}}{m} \]

Days Gained scores are only applicable for positive values of \( m \); our method assumes positive tumor growth during the period immediately before treatment.

The creation of existing response metrics has been largely motivated by the need to standardize the clinical discourse around response to therapy, and their strength lies more in their definitions of response and progression than in their predictive capabilities. Our purpose in creating the Days Gained metric was to develop a quantitative measure of response that is prognostic for patient outcomes, one that can aid clinical decision making by translating measurable treatment response to survival benefit. Our development of this metric was guided by the intuition that, given the heterogeneity of glioblastoma dynamics, maximizing a response metric’s prognostic power will require incorporating as much

**Figure 1.** Computing the Days Gained metric of response with models of varying complexity. We use patient-specific simulations of untreated tumor growth to estimate tumor burden on T1Gd MRI imaging at posttreatment time points. 4D-anatomic, 4D-spherical, and linear models simulate glioblastoma growth with decreasing levels of computational complexity. Days Gained scores can be computed from each of the 3 model types using the simulated T1Gd spherically equivalent radius curve. Top, model solutions from a 4D-anatomic UVC for a 53-year-old patient with a left temporal lobe glioblastoma. Bottom, concept of Days Gained score as previously published (8).
patient-specific data about tumor kinetics as possible. In our initial examination of the prognostic capabilities of our metric, we used spatially resolved, highly individualized models of glioblastoma growth to compute Days Gained scores (8). These models, referred to here as "4D-anatomic" models, account for the full, complex geometry of a patient's tumor and simulate tumor growth within the unique anatomic confines of the patient's brain. These models simulate tumor dynamics in 4 dimensions: length, breadth, height, and time. From this initial research, we found that patients with higher Days Gained scores had improved overall survival (OS) and progression-free survival (PFS) outcomes compared with lower-scoring patients. However, these 4D-anatomic simulations are computationally intensive, requiring several hours or days to complete. To maximize the ease of clinical translation of our approach, we conducted the current study to determine how much detail we could remove from the model and still retain the ability to discriminate OS and PFS.

In addition, we investigated whether Days Gained might provide a means of addressing the challenges posed by pseudoprogression. Pseudoprogression occurs frequently (4), obscures true patient response, and delays clinical decision making at a time point when a patient's course of therapy is often reevaluated. We therefore examined whether Days Gained scores computed from the first MRI studies conducted after radiation could discriminate pseudoprogression from true progression.

The ultimate goal of our research is to validate and translate our response metric in a prospective trial to provide a clearer and more quantitative assessment of patients' responses to therapy. However, the introduction of complex computational methods into the clinical setting can be disruptive and burdensome to workflow, so our approach was to identify a minimal model of glioblastoma growth that could be used to calculate Days Gained and still retain discriminatory power. We applied 2 simplified models that reduce the computational cost of simulation: a 4D model that treats the tumor as an isotropic sphere (referred to here as the "4D-spherical" model), and a second, purely linear model of growth (referred to as the "linear" model). The 4D-spherical model is simpler and computationally faster than the 4D-anatomic model but still requires software that can solve partial differential equations, such as MATLAB (The MathWorks). Generally, Days Gained scores from the 4D-spherical model can be computed for 1 patient on the order of seconds to minutes on the 3.5-GHz desktop computer used for this study. Scores from the simpler linear model can be computed almost instantaneously using 1 algebraic equation (see Materials and Methods). In theory, they can be generated using any software that processes algebraic relationships or even from article-based nomograms or look-up tables.

For this study we computed Days Gained scores from the first MRI scans taken after standard-of-care radiotherapy for 63 newly diagnosed patients with glioblastoma. We computed scores using both the 4D-spherical and linear models and compared them against scores generated using the more computationally intensive 4D-anatomic model that we collected from a 31-patient subset of our cohort. We used linear regression and Bland–Altman analyses (9) to compare the scores and conducted Kaplan–Meier analyses (10) to determine their ability to discriminate PFS and OS outcomes. We used Days Gained scores calculated from the minimal model in attempting to distinguish pseudoprogression from true progression among the 36 patients who had radiographic evidence of progression within the first 180 days after completion of radiotherapy.

Materials and Methods

Patients

We collected written informed consent from a total of 63 newly diagnosed patients with glioblastoma at the University of Washington (Seattle, WA) Medical Center and the University of California, Los Angeles (UCLA) Medical Center on a protocol approved by the University of Washington and UCLA Institutional Review Boards. Patients began treatment between September 2001 and September 2011 (mean, October 2006). Inclusion in the study required the existence of 2 pretreatment MRIs, 1 pretreatment T2-weighted sequence (T2), at least 1 gadolinium contrast-enhanced, T1-weighted (T1Gd) MRI taken following radiotherapy, a positive difference between the later and earlier pretreatment T1Gd tumor volumes, and a pathologic diagnosis of primary glioblastoma. The 2 pretreatment T1Gd and T2 imaging sequences are routine in most patients with glioblastoma because day-of-presentation (diagnostic) MRI scans and day-of-operation (neuro-navigational) guidance scans are typically taken several days apart, thus providing a temporal interval to measure the tumor's growth rate. Among our cohort, this pretreatment interval ranged from 2 to 266 days with a median of 11.

We computed Days Gained scores for these patients using the 4D-spherical and linear models and compared them against the 4D-anatomic model scores, our gold standard (11, 12). Our studies with the 4D-anatomic model, which requires a significant investment in time and human resources to parameterize and execute, included 31 of 63 patients in the current study. Therefore, when comparing the 4D-anatomic Days Gained scores with the other models’ scores, we used this 31-patient subset. This subset is similar to the 33-patient set used in the preliminary study that identified the prognostic power of the 4D-anatomic model (8), but it excludes patients with a hedged pathologic assessment of glioblastoma. Table 1 presents the clinical characteristics of the 63 patients used to compute 4D-spherical and linear Days Gained scores and the characteristics of the subset used to compute the 4D-anatomic scores. We observed no significant statistical differences between the 63- and 31-patient groups among the categories listed in the table.

A.D. Trister assessed all 63 patients for imaging evidence of progression in the first 180 days following radiation. Pseudoprogression was defined in those patients with early changes in enhancement consistent with progression that showed subsequent radiographic stabilization or improvement (3). The reviewer examined the entire clinical record of each patient to identify true progression versus pseudoprogression; this included radiology reports and all MRI images taken from the time of diagnosis through the time of the last clinical follow-up. Table 1 shows the clinical characteristics of the progression and pseudoprogression patient groups. We observed no statistically significant differences in their clinical characteristics.
Table 1. Clinical characteristics of patients in the primary study cohort, patients included in the 4D-anatomic simulation subset, and those who had progression and pseudoprogression on imaging.

<table>
<thead>
<tr>
<th>Clinical characteristics</th>
<th>4D-spherical and linear simulations (n = 63)</th>
<th>4D-anatomic simulations (n = 31)</th>
<th>Patients with progression on imaging (n = 24)</th>
<th>Patients with pseudoprogression on imaging (n = 12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y (median, range)</td>
<td>55, 19–89</td>
<td>58, 40–89</td>
<td>57, 37–76</td>
<td>54, 19–74</td>
</tr>
<tr>
<td>Gender</td>
<td>44% Female</td>
<td>39% Female</td>
<td>42% Female</td>
<td>33% Female</td>
</tr>
<tr>
<td>KPS at diagnosis (median, range)</td>
<td>90, 50–100</td>
<td>90, 60–100</td>
<td>90, 60–100</td>
<td>90, 50–100</td>
</tr>
<tr>
<td>RTOG RPA classification</td>
<td>III (21%)</td>
<td>IV (52%)</td>
<td>IV (50%)</td>
<td>IV (67%)</td>
</tr>
<tr>
<td>Resection treatment</td>
<td>Biopsy only 24%</td>
<td>26%</td>
<td>37.5%</td>
<td>8%</td>
</tr>
<tr>
<td></td>
<td>Subtotal resection 27%</td>
<td>26%</td>
<td>20.8%</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>Gross-total resection 49%</td>
<td>48%</td>
<td>41.7%</td>
<td>67%</td>
</tr>
<tr>
<td>Radiotherapy dose (cGy; median, range)</td>
<td>6,000, 2,900–6,900</td>
<td>6,080, 5,000–6,900</td>
<td>6,000, 2,900–6,900</td>
<td>6,000, 5,940–6,120</td>
</tr>
<tr>
<td>Chemoradiation with temozolomide</td>
<td>42 (67%)</td>
<td>20 (65%)</td>
<td>14 (58%)</td>
<td>8 (67%)</td>
</tr>
<tr>
<td>Chemoradiation with BCNU</td>
<td>5 (8%)</td>
<td>4 (13%)</td>
<td>2 (8%)</td>
<td>2 (17%)</td>
</tr>
<tr>
<td>Radiation without</td>
<td>16 (25%)</td>
<td>7 (22%)</td>
<td>8 (34%)</td>
<td>2 (16%)</td>
</tr>
<tr>
<td>temozolomide or BCNU</td>
<td>Days between end of radiotherapy and next MRI scan (mean, range)</td>
<td>27, 0–117</td>
<td>29, 1–72</td>
<td>21, 1–56</td>
</tr>
</tbody>
</table>

Abbreviations: BCNU, 1,3-bis(2-chloroethyl)-1-nitrosourea; KPS, Karnofsky performance status; RPA, recursive partitioning analysis; RTOG, Radiation Therapy Oncology Group.

Mathematical models

The 3 models used in this study simulate the growing tumor at different levels of detail. The 4D-anatomic model incorporates the patients’ neuroanatomy and the complex geometry of their tumor to simulate temporal changes in tumor cell densities throughout the brain. The 4D-spherical model is a simplified version of the 4D-anatomic model and treats the tumor as an expanding sphere, unconstrained by anatomy. This model simulates the temporal dynamics of tumor cell densities within 3 spatial dimensions but does not account for the tumor’s complex geometry nor the anatomic confines of the brain. The third model is a purely linear, empirical model of radial expansion of the T1Gd abnormality. Rather than simulate cell densities within a continuum, it models the growth of the spherically equivalent radius of the tumor, which is the radius of the measured tumor volume, assuming a spherical distribution as measured from the tumor burden visualized on T1Gd MRI. (For multifocal tumors this radius is computed from the aggregated volume of the disparate foci.)

\[
\text{Spherically equivalent radius} = \sqrt[3]{\frac{3 \times \text{volume}}{4 \times \pi}} \quad (1)
\]

4D-anatomic model. The general methods for generating our 4D UVC simulations are well documented (13–21) and are followed here. Two patient-specific parameters control the model’s dynamics: the tumor’s net diffusive capacity (\(D\)) and the tumor’s net proliferation rate (\(\rho\)). We compute \(D\) and \(\rho\) using measures of tumor volume on T1Gd and T2 MRIs at 2 pretreatment time points. To solve these terms, we first compute tumor growth rate by dividing the difference of the 2 pretreatment T1Gd volumes by the amount of time between their measurements. We then compute the invasiveness index of the tumor, \(D/\rho\), using a nonlinear relationship that maps a concomitant pair of T1Gd and T2 spherically equivalent radii to \(D/\rho\) values (14–16, 18, 20, 21). Using this relationship and Fisher approximation for traveling wave velocity \((2\sqrt{D/\rho})\), we identify \(D\) and \(\rho\) separately. We then simulate the growth of the tumor to the posttreatment time point, accounting for its unique diffusive and proliferative dynamics and the complex architecture of the brain with the following reaction-diffusion partial differential equation:

\[
\frac{\partial c}{\partial t} = \nabla \cdot (D(\nabla c)) + \rho c \left(1 - \frac{c}{k}\right) \quad (2)
\]

The equation describes the density of glioma cancer cells \(c\) in terms of \(D\), \(\rho\), and the tumor cell carrying capacity of the tissue \(k\) \((10^8 \text{ cells/cm}^3\); ref. 16).

In our 4D-anatomic simulations, we simulate tumor growth within a spatial discretization that is based on the patient’s day-of-operation, neuro-navigational, high resolution T1Gd MRI. We use the Statistical Parametric Mapping (22) toolbox in MATLAB to automatically segment the gray matter, white matter, and cerebrospinal fluid (CSF) regions of...
the brain. Each voxel in our discretization is then assigned a value of \( D \) based on this segmentation. To constrain the tumor growth to neuroanatomic boundaries, we set \( D \) to 0 in the CSF. To simulate the increased diffusive capacity of gliomas in white matter, we set \( D \) in areas of gray matter to a value 1,000 times less than that in white matter. From our simulations, we are able to visualize the tumor’s volumetric growth within the patient’s brain as shown in the serial, 3D images of Fig. 1.

\textbf{4D-spherical model.} Applying Equation 2 with growth only in the radial direction generates a full 4D simulation, creating a spherical tumor. We conducted simulations for this simplified geometry using the built-in \texttt{pdepem} function in MATLAB. We parameterize these simulations with \( D \) and \( \rho \) values that are computed in the same way as the 4D-anatomic simulations but execute them within a homogenous spatial domain that, unlike the 4D-anatomic models, is not informed by patient-specific anatomy.

\textbf{Linear model.} The linear model assumes that the T1Gd spherically equivalent radius of the tumor increases linearly from the date of diagnosis onward. This trajectory is represented by the equation

\[ y = mx + y_{\text{PRE}} \]  

where \( y \) is the simulated spherically equivalent T1Gd radius, \( x \) is the amount of time since the final T1Gd MRI taken before cytotoxic treatment, \( m \) is the pretreatment rate of T1Gd radial growth, and \( y_{\text{PRE}} \) is the spherically equivalent radius computed from the tumor volume observed on the final pretreatment T1Gd MRI.

To compute \( m \), we find the difference between the tumor volumes on 2 pretreatment T1Gd scans and divide it by the amount of time between the scans. For this study, we used the first and final pretreatment scans to compute \( m \), regardless of how many additional pretreatment scans were conducted. As an alternative, \( m \) can also be computed using pretreatment T2 scans. We repeated all our analyses using T2-derived velocities where available, and the implications of our results remained consistent with the T1Gd-derived velocities (data not shown).

Days Gained scores are only applicable for positive values of \( m \), as our method assumes positive tumor growth during the period immediately before treatment. Scores will be meaningless within the context of clinical treatment if pretreatment scan (\( x_{\text{PRE}} \))

\[ \text{Days Gained} = x_{\text{POST}} - \frac{y_{\text{POST}} - y_{\text{PRE}}}{m} \]  

For all 3 model implementations tested, there were some patients who had a posttreatment tumor burden that was larger than the model-predicted burden. Days Gained scores are negative in these cases, indicating that the burden increased beyond what would be expected in the absence of treatment.

\textbf{Computing Days Gained scores from untreated virtual controls} To generate Days Gained scores from the simulation results of the 4D-spherical and 4D-anatomic models, we first compute the T1Gd spherically equivalent radius time curve from the volume of voxels with tumor cell densities at or more than 80% of the total cell-carrying capacity of the tissue (14, 16), and then determine the spherically equivalent radius using Equation 1. To compute a Days Gained score, we find the time point on the T1Gd spherically equivalent curve where it best matches the actual posttreatment tumor size. The Days Gained score is the amount of time between this time point and the final (posttreatment) time point on the curve (Fig. 1). In some 4D-anatomic simulation studies, therapy reduced the tumor to a size smaller or larger than what was computed on the curve. In these cases, we used a linear interpolation between the T1Gd spherically equivalent radii values at the last pretreatment time point and the posttreatment time point to compute the Days Gained score. This method is consistent with the radial expansions predicted by the model and already observed in a spectrum of gliomas (17, 23).

Our 4D UVC simulations provide a number of outputs for quantifying patient response to treatment. Each UVC can be used to estimate the tumor volume as it would be observed on T1Gd and T2 imaging at any time between the pre- and posttreatment time points along with the total number of tumor cells distributed throughout the brain. Thus, we have several metrics available for computing Days Gained scores. We have found that the T1Gd spherically equivalent radius discriminates OS most significantly (24) and also discriminates PFS (8); therefore, our focus here is on the predictive value of Days Gained scores based on this output.

To compute Days Gained from the linear model, we use Equation 3 to solve for \( x \) when \( y \) is the measured posttreatment T1Gd spherically equivalent radius (\( y_{\text{POST}} \)), then subtract this value from the number of days that have passed since the final pretreatment scan (\( x_{\text{POST}} \)):

\[ \text{Days Gained} = x_{\text{POST}} - \frac{y_{\text{POST}} - y_{\text{PRE}}}{m} \]  

\textbf{Image processing} Computing a Days Gained score requires a volumetric segmentation of the tumor on 2 pretreatment and 1 posttreatment T1Gd MRI scan. Spherically equivalent radii, which are used to parameterize each UVC and to compute Days Gained scores, can be computed from these volumes using Equation 1. We measured tumor volumes on T1Gd and T2 patient MRIs using semiautomated segmentation software developed in MATLAB. This software facilitates tumor volume measurement by using a background subtraction algorithm that helps automate the tumor segmentation process by allowing the user to identify the tumor region with a polygon of interest. An automated edge detection algorithm then defines the imaging abnormality within the enclosed area. Although our segmentation pipeline relies on custom software not currently available to clinicians, several commonly used clinical visualization tools are readily available for conducting this same kind of segmentation, including ImageJ (25), OsirIX (26), and MIPAV (27).

All images were segmented by at least 2 independent reviewers who received training in segmenting tumor volumes from various MRI modalities. All measurements were reviewed by a third, independent reviewer who ensured all scans were
accurately measured. All measurers and reviewers received training in neuroanatomy and MRI abnormalities and were familiar with common MRI signals, such as resection, stroke, CSF, and normal anatomic variants and features.

Statistical analyses
To quantify the agreement between Days Gained scores computed from the 3 different models, we conducted linear regression and Bland–Altman analyses (9). We computed the coefficient of regression statistic in R version 2.14.1 and computed Bland–Altman biases and limits of agreement using MATLAB.

For each set of Days Gained scores that we computed from our models, we conducted iterative Kaplan–Meier analyses (10) to find an optimal Days Gained threshold that would maximally discriminate patients based on PFS and OS. We computed PFS as the interval between the patient’s start of cytotoxic therapy and the confirmation of progression using standard criteria as gathered from radiology reports. We computed OS as the interval between the patient’s date of diagnosis and his or her date of death. We censored observations at the time of last follow-up if the outcome in question was not observed. To ensure that our analyses included a substantial number of patients with and without the observed outcome, we limited it to cases in which both groups discriminated by the Days Gained threshold contained at least 20% of the total patient number. We used the "survival" package for R software version 2.14.1 (R Foundation for Statistical Computing) to conduct the Kaplan–Meier analyses (log-rank test).

Once we identified the simplest model that significantly discriminated outcomes according to these Kaplan–Meier analyses, we validated its overall prognostic power using bootstrap validation (28) in R (version 2.15.1). We created training sets by randomly resampling the original dataset with a replacement, then computed the optimal Days Gained threshold that discriminated the outcome in question. We then applied this threshold to a Kaplan–Meier analysis of the original dataset and recorded the \( \chi^2 \) statistic. After 10,000 iterations of this procedure, we averaged the resulting \( \chi^2 \) statistics, then computed the corresponding \( P \) value to validate our metric for discriminating PFS and OS outcomes.

We conducted a receiver–operator curve (ROC) analysis to find the optimal Days Gained threshold that would distinguish our true progression and pseudoprogression cohorts. After identifying this cutoff, we conducted a Welch \( t \) test to determine if the 2 groups had significantly different Days Gained scores and also assessed the positive and negative predictive value of the cutoff. To validate our approach more generally, we conducted leave-one-out cross-validation (LOOCV; ref. 28), a validation method more suited to dichotomous tests and reduced datasets. From the LOOCV, we computed the overall positive and negative predictive value to validate our method’s discriminatory power.

We conducted uni- and multivariate Cox proportional hazard analyses on Days Gained, age, RTOG RPA classification, enrollment in clinical trial, and gender to identify potential confounders in the study and to determine the value of Days Gained independent of other clinical factors.

We conducted the ROC, \( t \) test, LOOCV, predictive value, and Cox proportional hazards analyses in R version 2.15.0. We used the pROC package for ROC analysis (29). For all statistical tests, we considered \( P \) values less than or equal to 0.05 to be significant.

Results
Agreement between Days Gained scores from 4D-anatomic, 4D-spherical, and linear models
Figure 2 shows the agreement between Days Gained scores computed from T1Gd MRI scans at the first postradiotherapy time point for the 3 model implementations. Figure 2B, D, and F show the Bland–Altman biases and limits of agreement (9) across the model comparisons. We observed very strong correlation and agreement between scores generated by the 3 models: \( R^2 \) values were greater than 0.9 and absolute Bland–Altman biases were less than 14 days.

Progression-free and overall survival analyses
Figure 3 shows the results of our Kaplan–Meier analysis in which we identified the Days Gained thresholds that would most significantly discriminate PFS and OS among the linear, 4D-spherical, and 4D-anatomic model studies. We found that the 3 sets of Days Gained scores all have a maximal discriminating threshold that separates the patient population into groups with significantly different PFS and OS outcomes. For PFS, this level of significance was slightly higher for the linear and 4D-spherical model scores (\( P<0.001 \) for both) than for the 4D-anatomic scores (\( P = 0.004 \)). For OS, the 4D-anatomic scores showed a slightly higher level of significance (\( P = 0.004 \)) than the linear (\( P = 0.005 \)) and 4D-spherical (\( P = 0.006 \)) models. Maximally discriminating Days Gained thresholds and associated \( P \) values from the Kaplan–Meier analyses are shown in Fig. 3 and Table 2. Table 2 also shows the differences in median PFS and OS between patient groups above and below the optimal Days Gained thresholds that we identified. These optimal thresholds differentiated patients into cohorts with statistically significant median PFS and OS. Across the 3 model implementations, the PFS of high-scoring patients was 102% to 154% greater than that of lower-scoring patients. OS of high-scoring patients was 65% to 85% greater than that of lower-scoring patients. Figure 4 illustrates the complete set of \( P \) values from our Kaplan–Meier analyses for PFS and OS across model implementations. The figure shows that our OS and PFS results are robust across a range of Days Gained thresholds. Bootstrap validation of the minimal, linear model confirmed its significant prognostic power for discriminating both PFS (\( P < 0.001 \)) and OS (\( P = 0.011 \)).

Discriminating pseudoprogression
Fifty-eight of 63 patients had more than 1 posttreatment imaging study available for review, with 12 showing evidence of pseudoprogression and 24 with true progression in the first 180 days following radiotherapy (the remaining 22 showed no evidence of progression). ROC analysis of Days Gained within the population of patients with progression revealed an optimal score threshold of 90 that detects pseudoprogression with a sensitivity of 1.0, specificity of 0.58, positive predictive value of
0.55, negative predictive value of 1.0, and area under the ROC curve of 0.81. Figure 5A shows the significant difference in Days Gained scores between the true progression and pseudoprogression cohorts ($P = 0.003$; Welch $t$ test). To anticipate results from the same study with a higher patient number, we applied a binormal smoothing algorithm (30) to our ROC curve. At a threshold of 90 Days Gained, this smoothed curve gives a sensitivity of 0.83, specificity of 0.63, positive predictive value of 0.53, and negative predictive value of 0.88. Area under the curve is 0.79. Figure 5B presents the original and smoothed versions of the ROC curve. LOOCV of our method for discriminating pseudoprogression returned similar predictive value results, with a sensitivity of 0.79, specificity of 0.59, positive predictive value of 0.5, and negative predictive value of 0.93.

### Cox proportional hazard analyses

Univariate Cox proportional hazard analysis revealed significant correlations between survival and Days Gained score ($P = 0.009$), enrollment in clinical trial ($P = 0.006$), age ($P = 0.027$), and RTOG RPA classification. Gender was not significantly correlated with survival ($P = 0.35$). Multivariate Cox proportional hazard analysis only showed a significant correlation between survival and Days Gained score ($P = 0.03$) and enrollment in clinical trial ($P = 0.001$).

### Discussion

Patient-specific modeling is currently an active area of research, but few examples of modeling applications have been translated into the clinic for routine use (31). One of the challenges in making the step from the computational biology laboratory to the bedside is model complexity. Most patient-specific modeling research focuses on the creation of 4D, finite element or finite volume models. These models, including our 4D-anatomic model, require significant computational overhead to simulate and analyze, and interpretation of simulation results often requires expertise with the model. The few patient-specific modeling applications that have shown clinical impact are based on simple models that do not necessitate...
the use of a computer (31, 32). Because our ultimate goal is to improve the outcomes of glioblastoma patients by applying patient-specific modeling approaches, we sought to find a minimal model of glioblastoma growth that can be readily computed and analyzed. Our results here indicate that Days Gained scores based on our minimal growth model are strongly predictive of scores from more detailed models (Fig. 2) and are able to discriminate PFS and OS outcomes with similar statistical power (Figs. 3 and 4; Table 2). We are therefore confident that the purely linear model of tumor growth can serve as a minimal model for calculating patients’ Days Gained scores at posttreatment time points. In addition, we applied Days Gained scores from this model to address a current, difficult clinical problem: discriminating pseudoprogression in patients with disease progression following radiation (Fig. 5).

We anticipate that the linear model’s simplicity will help accelerate the translation of our Days Gained response metric into the clinic. Days Gained scores can be computed from this model using nomograms or mobile device applications or desktop software, whichever is most clinically convenient.

![Figure 3. Kaplan–Meier analyses comparing the PFS and OS of patients classified according to Days Gained score. Scores from the minimal linear model discriminated outcomes with a level of significance similar to the other, more detailed, model implementations. Statistics for these analyses are shown in Table 2.](image-url)

<p>| Table 2. Statistics for Kaplan–Meier analyses of PFS and OS outcomes using Days Gained scores from the 3 model implementations |
| --- | --- | --- | --- | --- | --- | --- | --- |</p>
<table>
<thead>
<tr>
<th>Model</th>
<th>n</th>
<th>Optimal Days Gained threshold for PFS</th>
<th>P value for PFS</th>
<th>Difference in median PFS (months, % improvement)</th>
<th>Optimal Days Gained threshold for OS</th>
<th>P value for OS</th>
<th>Difference in median OS (months, % improvement)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>63</td>
<td>93</td>
<td>&lt;0.001</td>
<td>6.4, 154%</td>
<td>78</td>
<td>0.005</td>
<td>9.4, 85%</td>
</tr>
<tr>
<td>4D-spherical</td>
<td>63</td>
<td>86</td>
<td>&lt;0.001</td>
<td>6.4, 154%</td>
<td>84</td>
<td>0.006</td>
<td>9.4, 85%</td>
</tr>
<tr>
<td>4D-anatomic</td>
<td>31</td>
<td>117</td>
<td>0.004</td>
<td>7.1, 102%</td>
<td>117</td>
<td>0.004</td>
<td>8.1, 65%</td>
</tr>
</tbody>
</table>
Most appealing is that computing these Days Gained scores requires no additional clinical data beyond what is routinely collected for patients with glioblastoma. However, the required volumetric segmentation of the tumor on 2 pretreatment and 1 posttreatment scan may present a bottleneck to clinical adoption. Although semiautomated tools exist that reduce segmentation effort, this task can be time consuming, especially when conducted on high-resolution MRI data. In future studies, we plan to investigate whether simpler measurements of tumor burden, such as those used in clinical response criteria (sum of products of maximally perpendicular diameters, for example) can be used as surrogates for a complete volumetric assessment of the tumor.

Previous studies have shown the linear nature of radial expansion of gliomas on MRI (17, 23); thus, it is not surprising that our minimal, linear model performs as well as the 4D-spherical and 4D-anatomic versions. However, while the minimal model may be optimal for computing Days Gained scores in a clinical setting, more detailed models will be required for analyzing tumor cell density profiles, spatial growth dynamics, or the interaction between a patient’s neuroanatomy and the growing tumor. For instance, in the context of customizing radiotherapy for an individual, spatiotemporal tumor dynamics must be resolved to optimize the radiation dose distribution (15, 33).

In this study, we computed each patient’s Days Gained score using the first postradiation T1Gd MRI scan. In the modern era of treatment with surgery and radiation combined with temozolomide (34), we have seen the rate of pseudoprogression as high as 50% of patients (4). Because every Days Gained score was measured during this period, we were concerned that the presence of pseudoprogression may lead to spurious results in our predictions. Rather, we found that all patients with pseudoprogression had Days Gained scores that were significantly higher than those in patients with true progression. These scores discriminated between the 2 groups, and cross-validation confirmed their high negative predictive value. Furthermore, even if pseudoprogression falsely lowered some Days Gained scores computed from the first postradiation scan, these scores remained prognostic for PFS and OS. We also recognize that pseudoresponse, a spurious decrease in contrast-enhancing tumor burden resulting from antiangiogenic therapy (35), could have affected Days Gained scores. However, this is unlikely in our patient population given that no patients received antiangiogenic therapy before the posttreatment scan used to compute their score.

In terms of discriminating pseudoprogression from true progression, our statistical analyses show that the power of Days Gained lies in its sensitivity and negative predictive value. Thus, progression observed in a patient scoring less than 90 Days Gained is highly likely to be a true increase in tumor burden rather than an artifact of treatment. Days Gained is less powerful at discriminating true progression from pseudoprogression in patients who score more than 90. In these cases, using Days Gained in combination with other tests and imaging modalities may help discriminate pseudoprogression.

We acknowledge that some of the shorter pretreatment intervals used to compute tumor growth rates might have introduced error into Days Gained scores. However, our study goal was to develop a metric of response that could be used with clinical data that are routinely collected in glioblastoma cases. We did not, therefore, seek to remove sources of error in our measurements if they were part of routine patient care but rather found a response metric that is prognostic for survival outcomes despite the presence of those errors.

We also acknowledge that a small fraction of our patients may not have received the current standard of care because they received treatment before the widespread adoption of the Stupp protocol; however, the vast majority of patients represent modern treatment cases. Our Cox proportional hazard analyses showed that Days Gained scores correlated with survival, indicating the independent value of the metric. There are many potential confounders in survival analysis in this population, including clinical trial enrollment and age; however, multivariate Cox proportional hazard analysis shows that Days Gained remains significant despite these confounders.

Our results show that Days Gained scores based on a simple linear growth model and computed at an early time point may
identify those patients who would benefit from changes in their course of treatment.

In future studies, we plan to investigate whether Days Gained scores remain prognostic of PFS and OS when calculated from later posttreatment time points, including during the postsalvage therapy period. For such a study, we will have to assess the impact that antiangiogenic agents (e.g., bevacizumab, cediranib) have on Days Gained scores. As addressed by the RANO working group, effective response criteria must account for the use of antiangiogenic therapies because they can reduce the permeability of tumor vessels and give false indications of response on T1Gd MRI (7). We also plan to include Days Gained in a prospective trial of treatment of glioblastoma.

Disclosure of Potential Conflicts of Interest
A. Lai has a commercial research grant from Genentech and is a consultant/advisory board member of the same. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions
Conception and design: M.L. Neal, A.D. Trister, R.C. Rockne, K.R. Swanson
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): A. Baldock, L. Guyman, T. Cloke, A. Lai, T.F. Cloughesy, M.M. Mrugala, J.K. Rockhill, R.C. Rockne, K.R. Swanson
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): M.L. Neal, A.D. Trister, S. Ahn, C.A. Bridge, R.C. Rockne, K.R. Swanson
Writing, review, and/or revision of the manuscript: M.L. Neal, A.D. Trister, S. Ahn, A. Baldock, C.A. Bridge, R. Sodt, A. Lai, T.F. Cloughesy, M.M. Mrugala, J.K. Rockhill, R.C. Rockne, K.R. Swanson
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): M.L. Neal, A.D. Trister, A. Baldock, C.A. Bridge, J. Lange, T. Cloke, R.C. Rockne, K.R. Swanson
Study supervision: A.D. Trister, R.C. Rockne, K.R. Swanson

Acknowledgments
The authors thank the generous and unwavering support of Ellsworth (Buster) C. Alvord Jr (1923–2010), who is gravely missed. May his scientific legacy live on through the work reported here. The authors also thank those who contributed to image analysis in this study: Kellie Fontes, Jacob Gile, Alex Kim, Chantal Murthy, and Clifford Rostomily.

Grant Support
This study is supported by McDonnell Foundation (J. Lange, R.C. Rockne, and K.R. Swanson); Brain Tumor Funders Collaborative (M.L. Neal, R.C. Rockne, and K.R. Swanson); NIH grants: NS060752 (M.L. Neal, S. Ahn, A. Baldock, R. Sodt, T. Cloke, M.M. Mrugala, J.K. Rockhill, R.C. Rockne, and K.R. Swanson), CA16437 (C.A. Bridge, L. Guyman, M.M. Mrugala, J.K. Rockhill, R.C. Rockne, and K.R. Swanson), and CA149237 (A.D. Trister); University of Washington James D. Murray Endowed Chair (K.R. Swanson); University of Washington Academic Pathology Fund (R.C. Rockne and K.R. Swanson).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 1, 2012; revised January 14, 2013; accepted January 17, 2013; published OnlineFirst February 11, 2013.

References


Response Classification Based on a Minimal Model of Glioblastoma Growth Is Prognostic for Clinical Outcomes and Distinguishes Progression from Pseudoprogression

Maxwell Lewis Neal, Andrew D. Trister, Sunyoung Ahn, et al.

Cancer Res  Published OnlineFirst February 11, 2013.

Updated version Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-12-3588

Supplementary Material Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2013/02/11/0008-5472.CAN-12-3588.DC1

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

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

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