Impact of imaging measurements on response assessment in glioblastoma clinical trials

David A. Reardon, Karla V. Ballman, Jan C. Buckner, Susan M. Chang, and Benjamin M. Ellingson

Center for Neuro-Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts (D.A.R.); Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, Minnesota (K.V.B.); Department of Oncology, Mayo Clinic, Rochester, Minnesota (J.C.B.); Department of Neurological Surgery, University of California – San Francisco, San Francisco, California (S.M.C.); Department of Radiological Sciences, David Geffen School of Medicine at UCLA, Los Angeles, California (B.M.E.); Department of Biomedical Physics, David Geffen School of Medicine at UCLA, Los Angeles, CA (B.M.E.); Department of Bioengineering, Henry Samueli School of Engineering and Applied Science at UCLA, Los Angeles, California (B.M.E.); Brain Research Institute, David Geffen School of Medicine at UCLA, Los Angeles, California (B.M.E.); UCLA Neuro-Oncology Program, David Geffen School of Medicine at UCLA, Los Angeles, California (B.M.E.)

Corresponding Author: David A. Reardon, MD, Clinical Director, Center for Neuro-Oncology, Associate Professor of Medicine, Harvard Medical School, Dana-Farber Cancer Institute, 450 Brookline Ave, Dana 2134, Boston, MA 02215-5450 (david_reardon@dfci.harvard.edu).

We provide historical and scientific guidance on imaging response assessment for incorporation into clinical trials to stimulate effective and expedited drug development for recurrent glioblastoma by addressing 3 fundamental questions: (i) What is the current validation status of imaging response assessment, and when are we confident assessing response using today’s technology? (ii) What imaging technology and/or response assessment paradigms can be validated and implemented soon, and how will these technologies provide benefit? (iii) Which imaging technologies need extensive testing, and how can they be prospectively validated? Assessment of T1 +/- contrast, T2/FLAIR, diffusion, and perfusion-imaging sequences are routine and provide important insight into underlying tumor activity. Nonetheless, utility of these data within and across patients, as well as across institutions, are limited by challenges in quantifying measurements accurately and lack of consistent and standardized image acquisition parameters. Currently, there exists a critical need to generate guidelines optimizing and standardizing MRI sequences for neuro-oncology patients. Additionally, more accurate differentiation of confounding factors (pseudoprogression or pseudoresponse) may be valuable. Although promising, diffusion MRI, perfusion MRI, MR spectroscopy, and amino acid PET require extensive standardization and validation. Finally, additional techniques to enhance response assessment, such as digital T1 subtraction maps, warrant further investigation.

Keywords: clinical trials, glioblastoma, imaging, MRI, response assessment.

On January 2014, the Jumpstarting Brain Tumor Drug Development Coalition, consisting of the National Brain Tumor Society, The Society for Neuro-Oncology, Accelerated Brain Cancer Cure and the Musella Foundation for Research and Information, in close collaboration with the United States Food and Drug Administration, sponsored a workshop to help evaluate response criteria and endpoints for neuro-oncology clinical trials. This manuscript summarizes the report of Panel 3 for this workshop which was charged with reviewing clinical trial design and its impact on imaging measurement of tumor progression and response to drug therapy.

Why is Imaging Important in the Context of Clinical Trials for Recurrent Glioblastoma?

Although overall survival (OS) remains the gold standard, evaluation of changes in tumor burden using imaging is critical for accurate interpretation of response to a particular therapeutic paradigm, particularly in the context of recurrent disease. Tumor shrinkage and delay of tumor recurrence, as measured by objective response rate (ORR) and progression-free survival (PFS), respectively, are potentially meaningful additional endpoints if they correlate with improvements in either OS or patient well-being. These associations have been frequently established for other malignancies, although the limited consistency has been limited historically for glioblastoma (GBM). One possible explanation for not observing these associations more regularly for GBM is that therapies evaluated to date, with the possible exception of bevacizumab, have been essentially ineffective. Another explanation may be linked with the remarkably adaptive capability of GBM tumors, particularly with regard to the emerging resistance to therapeutic intervention. Nonetheless, a wide spectrum of novel and promising therapeutics are currently in development for GBM patients, and additional innovative treatment strategies continue to emerge as scientific understanding of...
GBM pathophysiology advances. Radiographic endpoints including ORR and PFS continue to offer attractive, although imperfect, measures of potential antitumor benefit that may help guide development of these approaches forward. Furthermore, in the context of malignant glioma, an inherently and diffusely infiltrative and destructive class of tumors, ORR and PFS are particularly advantageous endpoints because therapeutics capable of shrinking tumors or prolonging PFS would be expected to preserve neurological function and overall quality of life.

In addition to these considerations, PFS and imaging response have a number of advantages over OS. First, these endpoints can be assessed rapidly, which saves time as well as resources. This feature is particularly advantageous given the rapidly increasing number of investigational agents and combinatorial regimens pending clinical evaluation. Second, PFS and imaging response are not impacted by crossover to subsequent therapy. Last, there is evidence to suggest an association between PFS, durable response, and OS in glioblastoma patients, suggesting that imaging response assessment may be a surrogate for clinical benefit as defined by overall patient survival.

Another important consideration with regard to imaging-based endpoints is the variability of historical benchmarks, which can serve as useful benchmarks when assessing new therapeutic interventions. For recurrent GBM, rates of ORR, PFS and OS have remained remarkably consistent across a variety of cytotoxic and biologically based therapies, excluding vascular endothelial growth factor (VEGF)/vascular endothelial growth factor receptor (VEGFR) inhibiting therapeutics. Tables 1 and 2 summarize outcome for phase ≥ II clinical trials conducted with cytotoxic agents and nonangiogenic, biologically based therapeutics, respectively, in recurrent GBM patients over the past 10–15 years. With rare exception, the ORR and PFS rates observed in these trials is ≤5%, with the only observed exception being trials where temozolomide was used at recurrence for patients who did not receive it as frontline therapy. Therefore, the benchmarks established by aggregate historical data for recurrent GBM patients provide a readily available and useful screening comparator for future clinical trials.

In summary, imaging-based endpoints of response are particularly relevant for neuro-oncology, and they can decrease the cost and time associated with a clinical trial, are not influenced by crossover, and can quantify effects of therapeutic regimens on tumor growth. Furthermore, historical benchmarks for recurrent GBM provide established yardsticks to gauge antitumor activity. On the other hand, as discussed in detail further in this supplement, accurate determination of progressive intracranial tumor has become increasingly challenging using traditional imaging modalities.

**Definition of Standard Imaging Endpoints in Clinical Trials for Recurrent Glioblastoma**

**Response and Objective Response Rate**

Response can be defined as a decrease in tumor size relative to initial tumor burden beyond some predefined threshold. Additionally, a confirmatory scan (see below) may also be necessary for verifying that true response has occurred and thereby exclude pseudoprogression. ORR can be defined as the proportion of patients treated with a drug that demonstrate response.

**Durability of Response and Need for Confirmatory Scans**

Although radiographic response following administration of a specific agent is always encouraging, agents that are associated with responses that are short-lived are unlikely to be of meaningful benefit to patients. If a measure of response duration, or the time from when response was noted to the time of progression, adds value to absolute response rate as it provides a measure of how long the tumor is controlled. Thus, the concept of response duration requires a confirmatory scan to ensure that the response has been sustained for follow-up evaluations.

**Progression-free Survival and Time to Progression**

Progression can be defined as an increase in tumor size relative to initial tumor burden beyond some predefined threshold. Similar to the definition of objective response, tumor progression may also include a confirmatory scan to exclude possible cases of pseudoprogression. Measures of the time from initial treatment to progression include time to progression (TTP) or PFS. An important distinction between these 2 measures is that TTP refers exclusively to time to tumor progression while PFS also includes time to death from any cause.

**Dynamic Growth Estimates for Quantifying Subclinical Benefit**

It is conceivable that a drug may have a subclinical benefit for patients by changing dynamic growth patterns of the underlying tumor. For example, the use of doubling time or growth rate based on contrast-enhanced CT has been used in a number of historic studies to quantify changes in growth rates at diagnosis, after therapies, and at tumor recurrence. Assessment of changes in dynamic growth patterns has been investigated, more so for low-grade than high-grade gliomas, but the underlying principles could be effectively applied to any CNS malignancy. In a recent analysis of 407 newly diagnosed adult low-grade gliomas, the velocity of diametric expansion was shown to be an independent, multivariable predictor for malignant transformation as well as OS, while others have shown that changes in velocity of volumetric expansion can predict outcome following radiotherapy or chemotherapy. Clinical benefit may be quantified in terms of survival gain with the addition of therapy, meaning how the altered growth rate was expected to result in an increase in survival relative to the pretreatment tumor growth trajectory. Careful assessment of the rate of growth (or response) of a tumor between consecutive scans may provide a measure for subclinical benefit or a means of differentiating true progression (or response) from pseudoprogression (or pseudoprogression). Nonetheless, the potential overall value of such changes will need to be evaluated in a well-designed, prospective trial.

**Need for Basic Standardized MRI Acquisition and Postprocessing Methodology**

The stability, accuracy, and reproducibility of brain tumor size measurements are intimately tied to MRI acquisition and postprocessing methodology details. In multicenter studies, the heterogeneity of MR scanners and parameters (eg, field strength, gradient systems, manufacturers, sequence parameters, etc.)
must be considered. It is well known that minor variations in hardware or sequence timing parameters may result in substantial changes in image contrast between tissues of interest, which can potentially confound interpretation of changes caused by therapy or the disease itself. Further, differences in postprocessing (eg, interpolation/smoothing, digital subtraction, etc.) or

**Table 1.** Representative recent phase ≥ II clinical trials for recurrent glioblastoma patients evaluating cytotoxic (chemotherapy) agents

<table>
<thead>
<tr>
<th>Agent</th>
<th>Patients (n)</th>
<th>Response Criteria</th>
<th>ORR (%)</th>
<th>Durability of ORR</th>
<th>PFS-6 (%)</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NABTC</td>
<td>437</td>
<td>Macdonald</td>
<td>7</td>
<td>Not stated</td>
<td>16</td>
<td>Lamborn, 2008</td>
</tr>
<tr>
<td>Carboplatin + thymidine</td>
<td>45</td>
<td>Macdonald</td>
<td>2.2</td>
<td>Not stated</td>
<td>16</td>
<td>Robins, 2002</td>
</tr>
<tr>
<td>BCNU</td>
<td>40</td>
<td>Macdonald</td>
<td>15</td>
<td>Not stated</td>
<td>17.5</td>
<td>Brandes, 2004</td>
</tr>
<tr>
<td>BCNU + TMZ</td>
<td>36</td>
<td>Macdonald</td>
<td>5.5</td>
<td>Not stated</td>
<td>21</td>
<td>Prados, 2004</td>
</tr>
<tr>
<td>Carboplatin + erlotinib</td>
<td>43</td>
<td>Macdonald</td>
<td>2.3</td>
<td>15 weeks</td>
<td>14</td>
<td>De Groot, 2008</td>
</tr>
<tr>
<td>CCNU</td>
<td>92</td>
<td>Levin</td>
<td>4.3</td>
<td>Not stated</td>
<td>19</td>
<td>Wick, 2010</td>
</tr>
<tr>
<td>CCNU</td>
<td>65</td>
<td>RANO</td>
<td>8.9</td>
<td>Not stated</td>
<td>25</td>
<td>Batchelor, 2013</td>
</tr>
<tr>
<td>Clororetazine</td>
<td>32</td>
<td>Macdonald</td>
<td>0</td>
<td>Not applicable</td>
<td>6</td>
<td>Badruddoja, 2007</td>
</tr>
<tr>
<td>Hydroxyurea</td>
<td>120</td>
<td>Macdonald</td>
<td>0.8</td>
<td>Not stated</td>
<td>5</td>
<td>Dresennam, 2010</td>
</tr>
<tr>
<td>Irinotecan</td>
<td>40</td>
<td>Macdonald</td>
<td>0</td>
<td>Not applicable</td>
<td>NR</td>
<td>Santisteban, 2009</td>
</tr>
<tr>
<td>Irinotecan</td>
<td>48</td>
<td>Macdonald</td>
<td>17</td>
<td>12–42 weeks</td>
<td>NR</td>
<td>Friedman, 1999</td>
</tr>
<tr>
<td>Irinotecan</td>
<td>40</td>
<td>Macdonald</td>
<td>0</td>
<td>Not applicable</td>
<td>0</td>
<td>Chamberlain, 2002</td>
</tr>
<tr>
<td>PCV</td>
<td>63</td>
<td>Macdonald</td>
<td>11</td>
<td>Not stated</td>
<td>32</td>
<td>Kappelle, 2001</td>
</tr>
<tr>
<td>PCV</td>
<td>83</td>
<td>Macdonald</td>
<td>3.5</td>
<td>Not stated</td>
<td>38.4</td>
<td>Schmidt, 2006</td>
</tr>
<tr>
<td>Procarbazine</td>
<td>113</td>
<td>Macdonald</td>
<td>5.3</td>
<td>Not stated</td>
<td>8</td>
<td>Yung, 2000</td>
</tr>
<tr>
<td>TMZ</td>
<td>112</td>
<td>Macdonald</td>
<td>5.4</td>
<td>Not stated</td>
<td>21</td>
<td>Yung, 2000</td>
</tr>
<tr>
<td>TMZ (7/7)</td>
<td>45</td>
<td>Macdonald</td>
<td>15.5</td>
<td>Not stated</td>
<td>43.8</td>
<td>Wick, 2007</td>
</tr>
<tr>
<td>TMZ (21/28)</td>
<td>33</td>
<td>Macdonald</td>
<td>9</td>
<td>Median: 30.4 weeks</td>
<td>30</td>
<td>Brandes, 2006</td>
</tr>
<tr>
<td>TMZ (21/28)</td>
<td>58</td>
<td>Macdonald</td>
<td>13</td>
<td>Not stated</td>
<td>11</td>
<td>Norden, 2013</td>
</tr>
<tr>
<td>TMZ (28/28) – PD off TMZ</td>
<td>29</td>
<td>Macdonald</td>
<td>11.1</td>
<td>Not stated</td>
<td>35.7</td>
<td>Perry, 2010</td>
</tr>
<tr>
<td>TMZ + thalidomide</td>
<td>43</td>
<td>Macdonald</td>
<td>7</td>
<td>Not stated</td>
<td>24</td>
<td>Groves, 2007</td>
</tr>
</tbody>
</table>

Abbreviations: BCNU, carmustine; CCNU, lomustine; NABTC, North American Brain Tumor Consortium; ORR, objective response rate; PCV, procarbazine, CCNU, and vincristine; PD, progressive disease; PFS-6, progression-free survival at 6 months; TMZ, temozolomide

**Table 2.** Representative recent phase ≥ II clinical trials for recurrent glioblastoma patients evaluating non-angiogenic, biologic-based therapeutics

<table>
<thead>
<tr>
<th>Agent</th>
<th>Patients (n)</th>
<th>Response Criteria</th>
<th>ORR (%)</th>
<th>Durability of ORR</th>
<th>PFS-6 (%)</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMG102</td>
<td>61</td>
<td>Macdonald</td>
<td>0</td>
<td>Not applicable</td>
<td>15.0–17.9</td>
<td>Wen, 2011</td>
</tr>
<tr>
<td>Cilengitide</td>
<td>81</td>
<td>Macdonald</td>
<td>9</td>
<td>Median: 17 months</td>
<td>10–15%</td>
<td>Reardon, 2008</td>
</tr>
<tr>
<td>Cis-retinoic acid + celecoxib</td>
<td>25</td>
<td>Macdonald</td>
<td>0</td>
<td>Not applicable</td>
<td>19</td>
<td>Levin, 2006</td>
</tr>
<tr>
<td>Enzastaurin</td>
<td>174</td>
<td>Levin</td>
<td>2.9</td>
<td>Not stated</td>
<td>11</td>
<td>Wick, 2010</td>
</tr>
<tr>
<td>Erlotinib</td>
<td>54</td>
<td>Macdonald</td>
<td>3.7</td>
<td>Not stated</td>
<td>11.4</td>
<td>Van den Bent, 2009</td>
</tr>
<tr>
<td>Erlotinib</td>
<td>48</td>
<td>WHO</td>
<td>6.3</td>
<td>Median: 7 months</td>
<td>18.3</td>
<td>Yung, 2010</td>
</tr>
<tr>
<td>Erlotinib + sirol</td>
<td>32</td>
<td>Macdonald</td>
<td>0</td>
<td>Not applicable</td>
<td>3.1</td>
<td>Reardon, 2010</td>
</tr>
<tr>
<td>Fenretinib</td>
<td>23</td>
<td>Not stated</td>
<td>0</td>
<td>Not applicable</td>
<td>0</td>
<td>Pudavalli, 2004</td>
</tr>
<tr>
<td>Gefitinib</td>
<td>57</td>
<td>Macdonald</td>
<td>0</td>
<td>Not applicable</td>
<td>13</td>
<td>Rich, 2004</td>
</tr>
<tr>
<td>Imatinib</td>
<td>34</td>
<td>Macdonald</td>
<td>5.9</td>
<td>Not stated</td>
<td>3</td>
<td>Wen, 2006</td>
</tr>
<tr>
<td>Imatinib</td>
<td>51</td>
<td>Macdonald</td>
<td>5.9</td>
<td>&gt; 6 months</td>
<td>16</td>
<td>Raymond, 2008</td>
</tr>
<tr>
<td>Imatinib + hydroxyurea</td>
<td>33</td>
<td>Macdonald</td>
<td>9</td>
<td>Not stated</td>
<td>27</td>
<td>Reardon, 2005</td>
</tr>
<tr>
<td>Imatinib + hydroxyurea</td>
<td>231</td>
<td>Macdonald</td>
<td>3.4</td>
<td>Not stated</td>
<td>10.6</td>
<td>Reardon, 2005</td>
</tr>
<tr>
<td>Imatinib + hydroxyurea</td>
<td>120</td>
<td>Macdonal</td>
<td>1.7</td>
<td>Not stated</td>
<td>7</td>
<td>Dresennam, 2010</td>
</tr>
<tr>
<td>Lapatinib</td>
<td>17</td>
<td>Levin</td>
<td>0</td>
<td>Not applicable</td>
<td>Not stated</td>
<td>Thiesien, 2010</td>
</tr>
<tr>
<td>Temsirolimus</td>
<td>65</td>
<td>Macdonal</td>
<td>0</td>
<td>Not applicable</td>
<td>7.8</td>
<td>Galonis, 2005</td>
</tr>
<tr>
<td>Tipifarnib</td>
<td>676</td>
<td>Macdonal</td>
<td>7.5</td>
<td>Not stated</td>
<td>16.7</td>
<td>Cloughes, 2006</td>
</tr>
<tr>
<td>Vorinostat</td>
<td>66</td>
<td>Macdonal</td>
<td>3.0</td>
<td>Not stated</td>
<td>15.2</td>
<td>Galonis, 2009</td>
</tr>
</tbody>
</table>

Abbreviations: ORR, objective response rate; PFS-6, objective response rate
measurement techniques (e.g., bidirectional, unidirectional, volumetric, etc.) can also increase variability and uncertainty in lesion evaluation. Thus, there remains a significant need for tight control over sequence parameters in the context of multicenter clinical therapeutic trials in order to reduce measurement variability and increase accuracy of response assessment.

Possible Method for Defining a Clinically Meaningful Threshold for Progression for Estimates of PFS and TTP

As mentioned above, the imaging definitions for progression and response are defined relatively arbitrarily. One method of determining an optimal threshold for defining changes in tumor size that is clinically meaningful is to examine the effect a particular threshold has on imaging endpoint time-to-event parameters such as PFS and TTP. Conceptually, the optimal threshold for defining progression would result in a high correlation between these time-to-event imaging endpoints and objective measures of clinical benefit; however, these may be defined (Fig. 1). For example, the percent change in tumor size can be used as a continuous variable that is linked directly with PFS/TTP for each patient. The percentage change in enhancing tumor volume required for determining progression can then be adjusted, each patient’s PFS/TTP can be recalculated, and then all patient TTP/PFS information can be correlated with OS (as a measure of clinical benefit).

Fig. 2 outlines this process. Note that this strategy could also be used to optimize a threshold for determining response; however, measures of TTP/PFS or endpoints relating to progression will also include patients determined to have stable disease (not just responders), in which case this process will yield slightly different results.

Alternatively, we can explore use of the correspondence index, or c-index, as a measure of optimizing the threshold change in tumor size for determining progression. The c-index can be used to define an optimal threshold of change to define progression (rather than the arbitrary cutoff set at 25% increase in bidimensional product as specified by the Response Assessment in Neuro-Oncology [RANO] criteria or Macdonald criteria, for example) by examining the correspondence between progression and survival by continuous assessment or evaluation at set time points. For example, as outlined in Fig. 2, one may propose optimizing the threshold at a 3-month or 6-month evaluation of progression in order to have the highest correspondence with OS or some other objective measure of clinical benefit.

Clinical Trial Entry Criteria for Recurrent Glioblastoma

An important clarification provided by the RANO criteria was specification of the degree of radiographic worsening required to define progression in patients being considering for enrollment into clinical trials of salvage therapy. This clarification was included in the RANO criteria to achieve better consistency for patients enrolling into these trials and to specifically reduce the likelihood that patients with minimal evidence of radiographic worsening would be deemed progressive in order to obtain access to investigational therapy. Currently, the RANO criteria specify that patients must show at least a 25% increase in the sum of the products of perpendicular diameters of contrast-enhancing lesions or new enhancing lesions while on stable or increasing doses of corticosteroids to be deemed eligible for salvage therapy clinical trials. Furthermore, based on these criteria, clinical deterioration or increase in corticosteroid dosing alone would not be sufficient to indicate disease progression for entry into clinical trials.

Management of Pseudoprogression During Enrollment

A proportion of patients enrolled in clinical trials for recurrent GBM will have had pseudoprogression at the time of enrollment. Pseudoprogression may alter the interpretation or perception of subsequent drug response, given that pseudoprogression on MRI may improve spontaneously without therapeutic intervention and that such patients typically have a favorable survival. One obvious way to manage the proportion of patients with pseudoprogression is to limit the minimum time from the end of radiation therapy, when patients can be enrolled in trials for recurrent GBM. If we limit patients with disease progression to those who are at least 6 weeks from the end of radiation therapy, we can expect no more than 20% of all patients to have pseudoprogression, with longer intervals from the end of radiation therapy leading to smaller rates of pseudoprogression.

With regard to pseudoprogression itself, the list of therapies potentially associated with this phenomenon is increasing in neuro-oncology and includes radiation boost/reirradiation approaches, locally administered intratumoral therapies, and a wide array of immunotherapies including vaccines, immune checkpoint inhibitors, and T cell therapeutics. One possible approach for identifying patients with pseudoprogression in clinical trials for recurrent disease may be to evaluate several closely spaced scans (e.g., the scan that initially identified progression, another scan just prior to enrollment, and then one more scan just prior to initiation of treatment) to quantify changes in tumor volume in order to understand the basal rates of change prior to therapy. Timing for such sequential scanning, however, will require careful planning to minimize the patient’s risk of clinical deterioration. For patients with pseudoprogression, stability or improvement in tumor volume may be expected, whereas patients with true progression might show rapid increases in tumor volume consistent with continual, uninhibited tumor growth.

Minimum Tumor Size and Definition of Measurable Disease

As defined by the RANO criteria, measurable disease is defined as bidimensional contrast-enhancing lesions with clearly defined margins on MRI and 2 perpendicular diameters of at least 10 mm that are visible on 2 or more axial slices no farther than 5 mm apart without any interslice gaps. Additionally, the cystic or surgical cavity should not be measured in determining lesion size.

Specification of Known Prognostic Factors

Many prognostic factors may influence disease progression in patients with recurrent GBM. Comparison of outcomes between studies should be considered cautiously because of differences in factors between treatment groups including age, performance status, degree of resection, time from initial diagnosis, degree of prior treatment, extent of neurologic deficits, and tumor volume.
Fig. 1. Definition of tumor response to therapy. (A) Definition of time-to-progression (TTP) and/or progression-free survival (PFS). Tumors that are stable or not responding must show an increase in enhancement beyond a specific threshold from baseline to be deemed “progression”. The time from initiation of drug to progression is defined as TTP. (B) For “responding disease”, tumors must show a decrease in more than a specific threshold of change in size at some point during their therapy. This must be verified with a confirmatory scan 4 weeks following the scan with the largest response. This confirmation scan must not show tumor growth more than the threshold of progression, as defined from the smallest volume. Progression is then defined when the tumor grows to more than the threshold of change in size compared with the smallest volume. The “response duration” is defined as the time between the smallest volume, determined to be beyond the response threshold, and the time of progression. Landmark overall survival (OS) is determined from the time between the smallest volume, determined to be beyond the response threshold, and the time of expiration. (C) An example of “no response”, in which the tumor does not shrink beyond the optimal threshold. (D) Another example of “no response” is when the tumor shrinks beyond the optimal threshold, but the confirmatory scan increases more than the threshold defined for progression compared with the smallest volume. This is an unsustained response, and thus is not considered a responder.
Broad Categorization of Therapeutic Agents

Determination of response to therapeutics is intimately tied to their impact upon vascular permeability because GBM response assessment is currently dependent on contrast uptake as a surrogate for underlying tumor burden. Contrast uptake by malignant gliomas is directly dependent on vascular permeability and vascular surface area. Much of the recent and ongoing clinical research activity for GBM can be divided into 3 categories (Fig. 3):

- **Category 1**: agents that lack significant impact on tumor vascular permeability including traditional cytotoxic chemotherapeutics, proapoptotic agents, and therapies blocking key cell-signaling mediators; category 2, some agents that directly target tumor-associated vasculature and angiogenesis (eg, some inhibitors of VEGF/VEGFR, which tend to reduce contrast agent extravasation (ie, pseudoresponse); and category 3, agents that may impact tumor vessel integrity or alter various cytokines that could increase vascular permeability and lead to worsened contrast uptake independent of underlying tumor growth (ie, pseudoprogression). Possible examples of such agents include radiotherapy, vascular...
targeting agents, or some immunotherapeutics. Importantly, there are also therapeutic agents in which the distinction between these categories is unknown because the effect of such agents on tumor vascular permeability is unclear.

The impact of any therapeutic agent on tumor vascular permeability may be estimated by established techniques that assess changes in tumor vascularity such as dynamic contrast enhancement (DCE) perfusion MR imaging measurement of the transport coefficient, $K_{\text{trans}}$\textsuperscript{64}. It may be prudent to include such assessments as well as other exploratory imaging studies early in the development of investigational agents for GBM, such as phase I studies. Such assessments may help establish whether radiographic responses can be reliably ascertained by changes in contrast uptake as an accurate surrogate for antitumor effect. Furthermore, such assessments may help establish whether radiographic responses can be reliably ascertained by changes in contrast uptake as an accurate surrogate for antitumor effect. However, the distinction of specific changes in permeability that dictate a vascular agent via perfusion have not been established, and variability of DCE-MRI measurements of $K_{\text{trans}}$ may not be adequate to make this distinction. Ultimately, if impact on tumor vascular permeability cannot be confidently assessed, such therapeutics should be conservatively classified with those that are known to impact tumor vascular permeability.

### Modified Response Assessment Rubric

To account for the possibility of both pseudoprogression and pseudoresponse, a modified response assessment rubric can be considered based on the model used for the AVAglio trial (Fig. 4).\textsuperscript{65} Briefly in this model, the patient initially undergoes a pretreatment MRI scan (MRI [1]) prior to the first cycle of therapy. Following the first cycle of therapy, the patient receives additional MRI scans (MRI[N]). The model expands on that used in AVAglio by incorporating an algorithm for possible pseudoprogression.

#### Early Progression

If the lesion size has increased beyond a specific threshold between MRI Scan 1 and N, the patient is categorized as “early progression” and will be monitored for an additional time point and/or treatment cycle. After the next cycle of therapy (or another 4 weeks from Scan N), the patients undergoes a confirmatory MRI scan (MRI[N+1]). If the patient has an increase in lesion size from MRI Scan N, this is categorized as “true progression,” and the date of progression is the date of MRI Scan N is designated as the date of progression. If the patient has stable or decreasing size on MRI Scan N+1, pseudoprogression is confirmed, the new baseline for subsequent evaluation is MRI Scan N, and the patient continues on therapy.

#### Early Response

If the lesion size has decreased between MRI Scan 1 and N, the patient is categorized as an “early responder” and will be monitored for an additional time point and/or treatment cycle. After an additional cycle of therapy (or another 4 weeks from Scan N), the patient will undergo a confirmatory MRI scan (N+1). If the lesion has increased (indicating progression from MRI Scan N), this is considered an unsustained response or pseudoresponse. The date of progression for the patient will be MRI Scan N+1. Alternatively, if the lesion has not increased from Scan N, this is considered a durable response,” and the patient will continue on therapy. The date of progression is the time point of an increase in lesion size (from smallest lesion size) during the remainder of the study.

#### Stable Disease

If the lesion size has not increased or decreased beyond the set thresholds between scans 1 and N, the patient is considered stable. The patient will continue on therapy, and the date of progression is the time point of an increase in lesion size (from smallest lesion size) during the remainder of the study. For patients with significant neurological decline at the time of imaging progression...
as determined from Scan N, a confirmatory scan at time point N + 1 may not be possible or necessary. It is appropriate to define N as the progression time point for these patients.

Corticosteroid Use

Many neuro-oncology patients require systemic corticosteroids, such as dexamethasone, to decrease cerebral edema and improve associated headaches and other neurologic deficits. However, corticosteroids can significantly decrease tumor vascularity permeability and lead to decreased contrast uptake as well as diminished T2/FLAIR signal abnormality. Therefore, current imaging response criteria, including both RANO50 and Macdonald,66 preclude classifying a radiographic response for any patient who has received increased corticosteroid dosing prior to follow-up imaging. Specifically, both RANO and Macdonald criteria specify that a complete response or a partial response can only be assessed if the patient is on a stable or decreased corticosteroid dose. Patients who require an increase in corticosteroid dosing prior to a follow-up brain MRI should be classified as nonevaluable unless they satisfy criteria for either clinical or radiographic progression; they are appropriately classified as “progressive disease” at that time point. Thus, ORR for GBM patients excludes patients requiring increased corticosteroid dosing.

Utility of RANO/Macdonald Criteria for Single-arm Trials for Cytotoxic, Cytostatic, and Antineoplastic Agents That do not Modulate Vascular Permeability

The FDA regards overall radiographic response (defined as the proportion of patients with a Tumor size reduction of a predefined amount for a minimum time period) to be a valid endpoint for drug approval because such radiographic changes are felt to directly measure a drug’s antitumor activity.67 From a practical perspective, the radiographic response rate includes the sum of partial responses and complete responses and has traditionally excluded stable disease because stable disease may reflect the natural history of a given malignancy rather than actual therapeutic effect. Importantly, the overall value of a given radiographic response rate reflects not only the frequency of such responses but also their magnitude and duration as well as their association with symptom improvement.

From both a clinical and regulatory perspective, duration is a particularly meaningful aspect of radiographic response and is defined as the time from initial response until documented tumor progression. Several anticancer agents have been approved by the FDA based on a primary endpoint of ORR including denosumab for giant cell bone tumors,68 sutent for renal cell carcinoma,69 abraxane for metastatic breast cancer and non–small cell carcinoma,70 vismodegib for basal cell carcinoma,71 and oxaliplatin administered with 5FU/leucovorin for metastatic colorectal carcinoma.72 In their summary statements for these agents, the FDA noted duration of radiographic response to be an important consideration along with ORR frequency. In addition, many of these approvals were based on single-arm trials, indicating that a contemporaneous control arm may not be necessary for an ORR primary endpoint.

In an effort to improve the accuracy of brain tumor response assessment, the RANO criteria were carefully drafted by a multidisciplinary panel of experts in 2010.50 RANO was specifically developed to address key limitations of the Macdonald criteria,66 which were originally proposed in 1990 when radiographic responses were originally noted by CT scan in patients with 1p/19q co-deleted anaplastic oligodendroglioma tumors following procarbazine, CCN, and vincristine chemotherapy.73 The Macdonald criteria have served as the standard for radiologic response in neuro-oncology for the past 25 years. The RANO criteria were drafted to add several important considerations to the foundation provided by the Macdonald criteria.

Although the RANO criteria have been widely incorporated into neuro-oncology clinical trials and standard practice, prospective validation of whether these criteria provide a more accurate measure of tumor assessment than the historical Macdonald criteria has not been undertaken to date. Such a validation study will require a prospective comparison of response assessment measured by Macdonald versus RANO criteria in a setting where Macdonald criteria, which rely exclusively on measurement of contrast-enhancing tumor, are hypothesized to be deficient (eg, a trial evaluating an agent that directly alters vascular permeability). Given that the RANO criteria were developed to build upon response assessment as outlined by the Macdonald criteria, it is anticipated that RANO will be further modified as developing radiologic techniques (including some of those discussed by Panel 2 of this Workshop and reviewed in this supplement) gain widespread applicability and are in turn validated to provide more accurate response assessment capability.

Accurate and reliable radiologic assessment of tumor burden can be challenging for every solid tumor. As discussed by Panel 1 of this workshop and reviewed in this supplement, neuro-oncology patients, and in particular those with malignant gliomas such as GBM, are no exception. Areas of contrast enhancement and associated edema can be difficult to measure accurately on MRI, given their geometric complexity and variable growth patterns including extension along white matter tracts, ependymal surfaces, and neurovascular bundles. Several factors can also worsen MRI findings independent of tumor growth including seizures, stroke, hemorrhage, infection, and treatment-related pseudo-progression. Furthermore, corticosteroids, which are routinely used to decrease symptoms associated with cerebral edema, exert a potent antipermeability effect that can improve tumor-associated MRI findings independent of underlying tumor activity. Based on these concerns, the RANO guidelines appropriately preclude classifying partial or complete radiographic responses when patients have received increased corticosteroid dosing.

Use of ORR as a relevant endpoint for drug approval requires that radiographic response reflect therapeutic antitumor effect. ORR has been traditionally regarded as a reliable measure of therapeutic effect associated with cytotoxic agents because the underlying mechanism of action for these agents directly impacts tumor cell proliferation and/or survival, which in turn translates into shrinkage of responding tumors or growth of nonresponding tumors. Such changes in tumor size in GBM patients have been historically assessed with MRI by delineation of enhancing tumor mass as a surrogate for underlying tumor burden because tumor vessels in the macroscopic portion of the tumor are dysfunctional and leaky. Numerical thresholds for response and progression are in turn defined based on agreed-upon quantifiable parameters of enhancing tumors such as those outlined by RECIST.74 Macdonald criteria,66 or RANO.50
In contrast, agents that directly impact tumor vascular permeability, such as inhibitors of VEGF signaling or some therapeutics targeting other mediators of tumor angiogenesis, confound the ability to determine whether radiographic response reflects therapeutic antitumor effect. Such agents may diminish contrast uptake independent of a bona fide antitumor effect. In these circumstances, changes in enhancing tumor mass, as measured by currently utilized MRI techniques, provide an unreliable surrogate for underlying tumor burden. Similarly, accurate determination of tumor progression is limited.

Therefore, a principal consideration of whether radiographic response reflects a genuine antitumor effect for GBM is the class of therapeutic agent under evaluation. It is the consensus of Panel 3 that ORR and progression can be reliably assessed for therapeutic agents that do not directly modulate tumor vascular permeability. In contrast, agents that directly modulate vascular permeability, such as VEGF/VEGFR inhibitors, preclude the ability to accurately assess tumor burden and determine progression as functions of enhancing tumor mass with currently available assessment techniques. In summary, presently employed radiologic techniques and response assessment criteria can be reliably used to assess ORR as a measure of antitumor activity for GBM therapeutics, except for those that impact tumor vascular permeability.

**Consensus Statement – Objective response rate is an appropriate endpoint for single-arm trials of cytotoxic, cytostatic, and antineoplastic agents that do not directly modulate vascular permeability**

The rate of ORR, as assessed by currently available imaging techniques, is a valuable and rapidly assessed endpoint for single-arm studies in recurrent GBM patients that can support accelerated drug development of promising new agents. As previously discussed, the historical benchmarks of ORR for traditionally cytotoxic agents and biologically based therapeutics (excluding those directly targeting VEGF/VEGFR), provide well-established and consistent comparator data to support the evaluation of promising, similarly classified therapeutics in future single-arm and consistent comparator data to support the evaluation of those directly targeting VEGF/VEGFR, preclude the ability to accurately assess tumor burden and determine progression as functions of enhancing tumor mass with currently available assessment techniques. In summary, presently employed radiologic techniques and response assessment criteria can be reliably used to assess ORR as a measure of antitumor activity for GBM therapeutics, except for those that impact tumor vascular permeability.

The following factors are important considerations for ORR rate as an endpoint for GBM trials:

1. The class of therapeutic agent evaluated should exclude those that directly impact tumor vessel permeability.
2. The study population is limited to patients with recurrent glioblastoma and in whom recurrence is consistently defined by established parameters such as those specified by the Response Assessment for Neuro-Oncology (RANO) criteria;
3. The duration of radiographic response is an important indicator of meaningful antitumor effect and associated clinical benefit.

In addition, factors that add further value to a durable ORR rate include:

1. Increased rates of overall tumor shrinkage in study participants, as reflected by waterfall or spider plots;
2. Reproducibility across studies and investigators;
3. Confirmation by an independent review panel of expert neuroradiologists;
4. Correlation with overall survival (this must be validated in an appropriately randomized clinically trial);
5. Correlation with additional measures of clinical benefit (ideally, this would also be validated in an appropriately randomized clinically trial).

**Additional Efforts of High Potential Value**

Going forward, a critical priority to enable more effective use of imaging-based endpoints for neuro-oncology clinical trials includes a mandate for neuroradiologists and imaging scientists to generate guidance on optimized, standardized parameters for routine MRI sequence acquisition and processing. An effort to bring together imaging scientists, radiologists, neurosurgeons, radiation oncologists, and neuro-oncologists to develop a standardized MRI protocol is currently underway. Once defined, widespread implementation of such protocols will facilitate informative comparisons across imaging datasets along with more ease of interpretation by regulation agencies.

Another important effort, to be undertaken in parallel, should focus on assessing the clinical utility of advanced imaging techniques. A general strategy to evaluate the potential value of such imaging approaches will likely include a 2-step process. The first step should include a retrospective evaluation to assess the clinical merit of a specific advanced imaging technique using existing imaging and clinical datasets. If such a retrospective analysis suggests that a specified advanced imaging technique offers potential value for outcome assessment, the second step might involve a prospective evaluation of such a technique within a planned clinical trial. This step would provide an opportunity to proactively interrogate whether the data generated by the advanced imaging technique provide significant incremental value above the data generated by standard/routine imaging techniques. In our supplemental material, we provide an example of a 2-step strategy using T1 digital subtraction maps, a promising approach for evaluating response in patients undergoing treatment with therapeutic agents that directly alter vascular permeability.

**Conclusions/Future Considerations**

1. The discussion generated by this workshop provides a framework to further address neuroimaging challenges and prioritize strategies for moving forward.
2. A forum building on the RANO working group experience should prioritize standardization of acquisition and post-processing parameters of routinely performed MR imaging in order to generate uniform imaging standards for incorporation into clinical trials and ultimately into daily practice.
3. Retrospective and prospective validation of RANO response assessment should be undertaken, including separate evaluations for agents that are known to alter, as well as not alter, tumor vascular permeability.
4. Retrospective and prospective evaluation should also be undertaken to further evaluate the impact of additional advanced imaging techniques including diffusion, perfusion and T1-subtraction mapping approaches. Separate evaluations should be performed for agents that are known to alter, as well as not alter, tumor vascular permeability.
Supplementary Material

Supplementary material is available online at Neuro-Oncology (http://neuro-oncology.oxfordjournals.org/).

Funding

This work, summarizing discussion led by Panel 3 of the “Jumpstarting Brain Tumor Drug Development Coalition and FDA Clinical Trials Neuro-Imaging Workshop” held on January 30, 2014, was supported by general funding provided by National Brain Tumor Society, The Society for Neuro-Onco, Accelerated Brain Cancer Cure and the Musella Foundation for Research and Information.

Conflicts of interest Statement. Reardon – Speaker’s Bureau: Genentech/Roche; Merck/Schering; Advisory Board/Consultation: Genentech/Roche, Merck/Schering; Amgen, Novartis, EMD Serono, Stemline Therapeutics, Momenta Pharmaceuticals, Midatech; Research: Incyte; Ballman –; Buckner –; Chang – Research: Novartis; Merck/Schering; Ellingson – Genentech/Roche (Consultant & Research Grant); Siemens Healthcare (Consultant & Research Grant).

References


