

Imaging Endpoints in Brain Tumor Clinical Trials: Summary of a Jumpstarting Brain Tumor Drug Development Coalition Workshop

Objective:

To summarize the discussion sessions in the Brain Tumor Clinical Trials Endpoints Workshop 1 on Imaging held on January 30, 2014 in Bethesda, MD. The workshop was sponsored by the Jumpstarting Brain Tumor Drug Development Coalition with participation by the U.S. Food and Drug Administration (FDA).

Workshop Overview:

The Jumpstarting Brain Tumor Drug Development Coalition (which includes the National Brain Tumor Society, Accelerate Brain Cancer Cure, Musella Foundation for Research and Information, and Society for Neuro-Oncology) is sponsoring workshops to evaluate and improve the use of a variety of endpoints in brain tumor clinical trials, with a goal of advancing the development of treatments for glioblastoma multiforme (GBM). These workshops bring together key stakeholders (clinicians, researchers, industry, patient advocates, the National Cancer Institute, and the FDA) to discuss issues related to the development and use of endpoints, and to develop action plans that will ultimately lead to greater clarity and interest in the pursuit of clinical trials that can achieve FDA approval of new therapies. The discussions and action items that come out of these workshops are meant to inform and guide the neuro-oncology and clinical trial sponsor community.

The first of these workshops was held in Bethesda, Maryland, on January 30, 2014. This workshop's discussion was focused on the capability of neuro-imaging to accurately assess GBM response to therapies, and the use of current and emerging imaging-based endpoints in clinical trials. Specifically, stakeholders discussed how to overcome the variables in medical imaging, such as image acquisition parameters, that have hindered the ability to accurately assess brain tumor response to therapies, and how to best incorporate endpoints that rely on imaging into clinical trials.

The workshop began with three formal preliminary presentations to establish context for the day's discussions: "Efficacy Endpoints in Glioblastoma Multiforme Clinical Trials – A Regulatory Perspective," "Avastin & the Basis of Approval for GBM in 2009," and "Current Brain Tumor Imaging Protocols in Multicenter Trials." After the preliminary presentations, the workshop continued with four panel-led discussions. The panels consisted of experts in neuro-oncology and neuro-

radiology including FDA officials, industry representatives, and medical/academic researchers. After a short overview presentation by each panel, a two-part discussion period followed: (1) response by panelists to central questions posed by the moderator, and (2) facilitated audience question and answer. The panel discussions and central questions are summarized below.

Central Questions for Panelists
<p>Panel 1:</p> <ul style="list-style-type: none"> • Can current imaging technologies and criteria accurately determine if a tumor is shrinking or growing? • What is the confidence level of the brain tumor community with the current imaging techniques and criteria?
<p>Panel 2:</p> <ul style="list-style-type: none"> • What enhancements can be implemented with current technology to improve output while waiting for the new technologies to develop? • What technologies are being developed in the near future and distant future that will accurately determine if a tumor is shrinking or growing?
<p>Panel 3:</p> <ul style="list-style-type: none"> • What factors/considerations would permit overall response rate to be a valid endpoint for single arm trials to accelerate approval of promising therapeutics for GBM patients? • What are the critical factors that need to be addressed in a study to validate the imaging measurement/criteria? • What does a clinical trial look like that uses valid imaging as a primary endpoint?
<p>Panel 4:</p> <ul style="list-style-type: none"> • What are the clear priority action items (practical projects/tasks) to advance the field to improve the use of imaging to determine glioblastoma progression and response to drugs and biologics?

Summary of Panel Discussions:

Pros and Cons of Current Brain Tumor Imaging (Panel 1)

- The majority of participants agreed that radiographic response as measured through standard MRI correlates with tumor burden. Participants also determined that the tools to undertake single-arm, Phase 2 trials in some groups of therapies are available now.
 - A single scan showing a response may not be definitive, but this can be supported by confirming response on a follow up scan, and looking for a durability of response.
 - There are challenges in determining tumor response when using current contrast enhanced techniques, particularly in the context of non-cytotoxic/cytostatic drugs.
 - When agents affect vascular permeability, there are confounding effects, but evaluating non-enhancing disease could potentially lessen these.
- In recurrent GBM, changes in contrast enhancement generally assess changes in tumor burden. There are exceptions, and if these are diminished, demonstration of durable objective responses in a single arm study setting has the potential to represent benefit to patients.
- The field now sees the value of measuring the volume of the contrast-enhancing lesion, and feels this information is relevant to the assessment of tumor burden.

- The current response criteria used in neuro-oncology is the Response Assessment in Neuro-oncology (RANO). RANO was created to address limitations in the former criteria, Macdonald. RANO adds additional areas of assessment, such as consideration of non-enhancing tumor, a requirement that the response to be confirmed and durable, as well as addressing the issue of pseudo-progression. RANO defines progression as a 25% increase in the sum of the perpendicular diameters of enhancing lesion and includes a significant increase in fluid-attenuated inversion recovery (FLAIR) on stable or increasing steroid dose. However, quantification of FLAIR related changes is difficult. At the time of its inception, there was a lack of consensus with regard to some components of the criteria, and it was felt that leaving judgment of these issues to the investigator's discretion was best. Application of the RANO criteria is a work in progress, and its validation and shortcomings will be determined by future studies. The RANO criteria may be revised to incorporate newer techniques that might improve quantification, such as contrast enhancement with T1 subtraction, to better determine response to therapy. Additionally, RANO criteria may specify unique response criteria for different treatment types.
- When using T2/FLAIR, there is a need to consider some of the other conditions that exist within the environment of the tumor that could lead to an evolving high signal, such as surgical gliosis, radiation gliosis, and edema.
- Under circumstances where it is unclear whether images represent progression or improvement, it is important to continue to collect imaging data to learn more and to enable better decisions to improve patient care and clinical trial execution. (Sometimes MRI result may not be clearly understood to show progression or response. However, that image can be valuable to review during later patient assessments. Therefore, it is best to continue imaging and not to remove patients from trials prematurely. Doing so could lead to censoring and the loss of data that could otherwise be meaningful.)
- Variability in imaging acquisition and interpretation needs to be reduced, and standardization is critical.
- The FDA is open to having discussions about data before companies commit to a go/no-go decision with regard to different therapies they are developing.

Emerging Techniques and Technologies in Brain Tumor Imaging (Panel 2)

- The field is striving for more accurate means of using imaging in clinical trials to measure the impact of a therapy against GBM.
- The general consensus is that current technology is able to measure the impact of a therapy (especially when the effect is large) and is reliable. The panel recommended introducing new technology systematically when it enhances the understanding of important aspects of tumor pathophysiology.

- A potential path forward to improve assessment of brain tumor imaging is the use of contrast enhanced T1-weighted subtraction maps, especially in the setting of anti-angiogenic therapies. Subtraction maps reduce variability in the context of regular cytotoxic therapy and help to reduce the error in identifying areas of contrast enhancement. It is suggested that pre- and post-contrast T1-weighted magnetic resonance (MR) imaging sequences be matched and standardized in future multicenter trials.
- Volumetric measurements provide additional benefits compared to 1D or 2D measurements. It was proposed that acquisition of volumetric T1-weighted images should be performed to quantify enhancing tumor burden.
- There is a need for the community to agree on a set of imaging techniques and specify their use in specific contexts. This will also require standardization and validation in terms of the clinical relevance of the imaging measurement. The FDA should be provide comments and identify concerns during the decision-making process.
- Determining the value of T2/FLAIR imaging was also identified as an area to be investigated further. Interpreting T2/FLAIR images can be difficult because the signal is not specific and may arise from gliosis or from tumor. However, follow-up of T2/FLAIR imaging with diffusion MR imaging may provide a means to differentiate tumor presence.
- Although valuable, diffusion MR imaging, perfusion MR imaging (specifically dynamic susceptibility contrast (DSC), and amino acid positron emission tomography (PET) require further investigation and validation to confirm their utility.
- There is unnecessary variability in the imaging of brain tumors. Four categories of imaging variability were identified: acquisition, processing, reading, and patient. There was a recommendation to standardize future protocols across multicenter trials, employ a central imaging review panel, and use automated/semi-automated segmentation. Other aspects of standardization include use of constant field strength, sequence parameters, contrast agent dose, and contrast agent timing. At the very least, the same scanner with the same protocol should be used for the same patient over time.

Trial Design and Its Impact on Imaging Measurement of Tumor Progression and Tumor Response to Drug (Panel 3)

- Several panelists stated that in order to be able to use radiographic response as an endpoint in a single arm study to justify expedited approval of a therapy, there would need to be a substantial, durable and reproducible objective response with demonstrable clinically meaningful benefit to the patient. It would be necessary to confirm the findings in a randomized study with survival as an endpoint.
- In the short term, single arm clinical studies with imaging endpoints may be used as a means to obtain an early signal of efficacy to drive development decisions.

- To achieve consensus on the validation and definition of a reasonable threshold for defining radiographic response with varying types of therapeutic interventions, the panel proposed the following priorities for designing clinical trials:
 - Minimize the false positive effects that are not obviously related to the therapy. Use the RANO criteria to help eliminate the proportion of patients with pseudo-progression that go into trials, as well as the effects of corticosteroids on imaging.
 - To give greater confidence that patients are actually deriving benefit from the therapy, it will be necessary to show a relationship between radiographic response and another clinical outcome, e.g. through evaluation of a patient-reported outcome, neurocognitive assessment, or neurological assessment.
 - Use of the most current imaging techniques will be required for validation. Different measures such as T1 weighted images, T1 subtraction, T2/FLAIR, perfusion imaging, and diffusion imaging should be included in the analysis.
 - To ensure accuracy and consistency, imaging thresholds of response and duration of response must be refined and standardized via retrospective and/or prospective analyses:
 - For retrospective analyses, data from properly designed randomized study trials are needed. Inclusion and exclusion criteria and cross validation are required. This approach may be of limited utility because imaging standards and technology are evolving so rapidly that it may become difficult to interpret the results of a retrospective analysis.
 - The panel preferred prospective analyses because the acquisition and processing of the images can be standardized from the start. It would also be possible to include additional imaging techniques and modalities with the goal of identifying the most reliable methods for determining tumor shrinkage and duration of response.
 - Data may be sourced from NCI and industry-sponsored trials. This would require an independent organization to hold that data in a way that protects both patients and trial sponsors.
 - In the context of single-arm studies, overall response rate and duration of response have less methodological challenges than endpoints that incorporate time to events. FDA does not consider time to event analyses of data from single arm trials interpretable for the purpose of supporting a marketing application.
- The panel considered validation of methods for imaging assessments of response and progression achievable. However, the path to validation would be greatly facilitated by the existence of better therapies capable of demonstrating more robust anti-tumor

responses.

Creating an Action Plan for Improved Imaging Measured Brain Tumor Endpoints (Panel 4)

- The panel stressed the importance of the need for more effective treatments and that challenges in imaging interpretation would be diminished by the existence of more effective therapies.
- Key areas that were prominent throughout the day were discussed. These included the importance of obtaining advice from the FDA throughout the process of trial design; collaborating with the community, industry, and investigators from the start; defining response criteria - especially if imaging is used as a measure of tumor response to the agent; demonstrating that the image of the tumor is sensitive and specific enough to show response to the treatment and ensure its reproducibility; and standardizing imaging acquisition to reduce variability and gain confidence in the measure.
- An important first step is to identify the factors that cause variability and then work on addressing methods to eliminate or reduce the impact of these factors. Standardization of imaging techniques through formulation and adoption of clear guidelines will aid in achieving this goal.
- Response criteria need to be standardized for both single center and multi-center trials. In addition, acceptable thresholds for defining response need to be confirmed.
- It was agreed that there should be one initial standard used by all stakeholders, but its application and relevance for different classes of therapies will be evaluated going forward.
- The RANO criteria are considered to be reasonably reproducible and accurate. However, two issues must be addressed: (1) applicability across different therapy classes, and (2) the value of FLAIR (in addition to contrast). RANO is an evolving set of criteria and will be modified as data from trials are assessed. The RANO working group is currently working to revise and update the criteria.
- The neuro-oncology field has evolved over the past five years. Scientific advances have been achieved and two large trials were conducted using bevacizumab, providing a wealth of information that can be used to move the field forward.
- Several action items for moving the field forward were identified:
 1. Examine the RANO criteria as a benchmark, and undertake independent review of RANO criteria.
 2. Refine and standardize the application of the RANO criteria.
 3. Understand the modifications to imaging benchmarks of response and progression across therapies with different mechanism of action.
 4. Determine acquisition standards.
 5. Attain consensus from key stakeholders on imaging acquisition standards and criteria for assessing tumor response and progression.

6. Constantly monitor what we can learn along this trajectory and provide useful feedback to industry and academic investigators to encourage and facilitate effective drug development.
7. Establish an explicit set of rules around data management.

Conclusion:

The workshop produced a number of suggested action items intended to represent a starting point for future work in the area of imaging brain tumor response to treatment in clinical trials. The Jumpstarting Brain Tumor Drug Development Coalition will continue to work with our advisors and the FDA to identify outstanding issues needed to move the community closer to more and better treatments. There was consensus that the RANO criteria should continue to serve as the standard response criteria in brain tumor clinical trials. However, the imaging elements of RANO would benefit from additional clarification, including a clearer definition of response tailored to the type of therapy under investigation. It was suggested that further research be conducted to understand the RANO criteria's evaluation of imaging based response rate, and the Coalition will aid this effort. Additionally, the Coalition will lead in the formation of a working group that will address and develop guidance materials for the standardization of image acquisition and analysis within brain tumor clinical trials. The goal is for this guidance to be adopted by all sponsors of brain tumor clinical trials and to demonstrate that the brain tumor research and clinical trial environment is innovative, adaptive, and ready for further investment by academia, pharmaceutical companies, and the government.